Exploration of Relations between Belief-Tracking and Motor Processing using a new Ecologically-Valid Helping Task for Adults

By

Christopher N. Maymon

A thesis
submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Victoria University of Wellington
2018

Acknowledgements

I am indebted to many people for their support, friendship and guidance over the last three years. First, and foremost, I am extremely grateful to my supervisor Associate Professor Jason Low. Thank you for your persistent faith in my abilities and your enduring patience as my understanding grew, often slowly and through error. I am certain this dissertation would not have been possible without your unwavering encouragement in the face of adversity.

I would also like to extend my gratitude to Dr. Stephen Butterfill for his insight and helpful feedback. In addition, I extend my thanks to Dr. Hannes Rakoczy for his evaluation of the present experiment, and also to Dr. Alia Martin for her thoughts about how to more clearly organize these experiments for presentation.

Next, I must recognize my colleagues in the ToM lab. I am forever grateful to Katheryn Edwards, an intellectual giant whose counsel and comradery steadied my resolve on numerous occasions. I also wish to thank Schyana Sivanantham, Claudie Peloquin, Radhika Patel-Cornish, Tui Shaw, and Cong Fan for tolerating my unconventional, borderline-obsessive occupation of our shared workspace.

Outside of Victoria University, I am grateful to my dear friends and creative collaborators Kalani Muller and Pani Stavrou-Wilson, whose respective artwork is displayed herein in the form of the procedural figures (see figures 2, 3, and 9) and the state-analysis figures (see figures 8, 10 11 and 14). I also would like to thank my friends and colleagues: Nathan Beattie, Thomas Robinson, Holly Neilson, Summer Aykroyd, Ceara Nicholls, Buaphrao Raphiphatthana, Andrea Almeida, Jamie Cattell, Maggie Kycia, Jared Corston, Laina Isler, Sumaya Lamb, James Boag and Erica Douglas. I am profoundly grateful to my friend and partner, Jessica Randall, who freely chose to accompany me during the final and most strenuous chapter of this dissertation.

Finally, I offer my heartfelt gratitude to my family. In particular, I am indebted to my mother Kathleen Maymon, whose love and encouragement has been the de facto engine of my work on more occasions than she knows. Dad, Kelsey and Brendan, thank you for your continuous support and efforts to stay in close contact with me from the other side of the world. Finally, I wish to express my thanks to my beloved grandparents: to my late grandfather Russell Maymon, and my grandmother Nana Peg for our many spirited conversations. And to my Nana Claire and Papa Bob who have always been an anchor of encouragement and confidence in my abilities.

To my mentor and grandfather, Robert Nicholson

Abstract

Three experiments investigated efficient belief tracking as described by the two-systems theory of human mindreading (Apperly & Butterfill, 2009) whereupon mindreading implies the operation of a flexible system that is slow to develop and cognitively effortful, and an efficient system which develops early but subject to signature limits. Signature limits have been evidenced by children's and adults' difficulty anticipating how someone with a false belief (FB) about an object's identity, will act. In a recent investigation of signature limits, erroneous pre-activation of the motor system was detected when adults predicted the actions of an agent with an identity FB, suggesting that efficient mindreading and motor processes are linked (Edwards & Low, 2017). Moreover, young children differentiated between true and FBs about an object's location, but not identity, as revealed by the object children retrieved in an active helping task (Fizke et al., 2017). The aim of the present thesis was to provide new evidence of signature limits in adults, and of the recent conjecture that efficient mindreading and motor processes interact. In helping tasks, participants' interpretation of another's actions is crucial to how they coordinate their helping response. Therefore, an ecologically valid helping task was adapted to investigate the proposed interface between efficient mindreading and motor processes. The present work measured adults' eye movements made prior to helping, and their helping actions across a set of distinct directional full-body movements (around which side of a desk they swerved, which compartment they approached, toward which compartment they reached, and which object they retrieved). In this way, it was possible to investigate whether gaze direction correlated with full-body movements and whether adults' gaze differed when the agent's FB was about an object's location or identity. Results from Experiment 1 indicated that efficient belief tracking is equipped to process location but not identity FBs, and that - in the location scenario - gaze direction correlated with the immediate stage of participants' helping action (the direction they swerved). To investigate this correlation further, Experiment 2 drew upon research suggesting that temporarily tying an observer's hands behind their backs impaired their ability to predict the outcome of hand actions (Ambrosini et al., 2012). Results showed that tying adults' hands behind their back had a negative effect on their gaze behavior and severed the correlation between gaze and swerving, suggesting that the link between efficient mindreading and motor processes is fragile. Experiment 3 tested an alternative interpretation for Experiment 2's findings (that restraining participants' hands applied a domain-general distraction, rather than a specific detriment to belief tracking) by tying up participants' feet. Results were ambiguous: the gaze behavior of participants whose feet were tied did not differ from those who were unrestrained, nor from those whose hands were bound. These findings support the two-systems theory and provide suggestive evidence of a connection between efficient mindreading and motor processes. However, the investigation highlights new methodological challenges for designing naturalistic helping tasks for adult participants.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Mindreading: Conceptual Change vs. Early Competency

1.1 Introduction

Research in the field of mindreading has been associated with several seemingly contradictory findings. In the following sections, a chronological account of several studies is presented, the results of which have motivated competing theories of human mindreading. In addition, the following chapter sets up a framework for critically evaluating mindreading accounts themselves, as well as the many controversies which have clouded the interpretations and reliability of experimental findings.

1.1.1 False Beliefs

Human social behavior depends on an ability to read each other's mind, or to make inferences about the normative reasons that explain or predict others' actions. These normative reasons give rise to dispositions about the world, collectively referred to as 'mental states' (Premack & Woodruff, 1978). Mental states motivate one to act and guide those actions toward an outcome. Research into human mindreading has primarily focused on those mental states which are *propositional*: they describe an agent's stance in relation to objects and events in their environment (e.g., 'X knows *that*, thinks *that*, or remembers *that*). *Belief* is a noteworthy mental state as it is capable of motivating future actions in service of information which may have been true at one time but is now false. For this reason, the appreciation of others' beliefs has long been heralded as the *sine qua non* of mature mindreading.

A mindreader can predict that another person (referred to hereafter as 'an agent'), whom has obtained a *false* belief (hereafter: FB), will act in accordance with that belief, and by doing so their actions will typically result in failure. Recognizing that someone is acting on a FB situates a mindreader in a place of responsibility. One could respond by alerting the mistaken agent of their error or to help bring their belief in line with reality by guiding them to some evidence proving its untruth. Doing so necessarily employs an array of cognitive operations such as: constructing a representation of the contents of another's mind as resulting from certain perceivable changes in the environment, holding another's belief in mind whilst keeping it separate from one's own knowledge, and predicting another's future actions on the basis of their FB. Most importantly, understanding FB requires an appreciation for the potentially counterfactual nature of beliefs, being that beliefs can be at once valid (vis-à-vis the information known to the agent) and invalid (vis-à-vis reality).

1.1.2 Change-of-Location FB Tasks

FB tasks have been the chosen workhorse for the assessment of mindreading abilities. However, there are many ways in which someone can acquire a FB. For instance, many street performers and magicians capitalize on FBs that result from an onlooker's failure to track a target object which changes location. In a classic change-of-location FB task, participants are shown the following scene: a protagonist, 'Maxi', stores a bar of chocolate in a kitchen cupboard. Maxi leaves the room, and, in his absence, his mother moves the chocolate from the cupboard and into a kitchen drawer. Children are asked to predict where Maxi will look for his chocolate when he returns (Wimmer & Perner, 1983). Most children under the age of 4-years will fail to consider Maxi's FB and predict that Maxi will go toward the cupboard, where the chocolate actually is.

How robust is this finding? In a meta-analysis of 178 studies, Wellman, Cross and Watson (2001) confirmed that children's performance on standard FB tasks dramatically improves between 30-months (with ~80% of children failing) and 56-months of age (with 74.6% of children passing). The studies included in the meta-analysis cover a variety of FB narratives (including different protagonists, different target objects, whether the object is moved in order to deliberately trick the protagonist) and various formats of the critical test question (e.g., 'where will the agent look?', 'what will the agent think/believe?', 'what will the agent say?'). While the age effect was substantial across every analysis, performance was also facilitated by the following task variables: (1) an emphasis on deception when moving the target object, (2) the child's active participation in moving the object from one location to another, (3) a depiction or explicit re-stating of the protagonist's mental state and (4), the removal of the object from the scene rather than storing it in another location. In addition, country of origin had a significant effect on performance. For example, younger (40-month-old) children from Cameroon perform significantly better than 40-month-old children from western countries (Avis & Harris, 1991). However, despite relative differences in the onset of their shift in performance, they show parallel developmental trajectories and confirm that the development of FB understanding is universal (see also the meta-analysis by Liu et al, 2008).

Wellman and colleagues concluded that from around their fourth birthday, children undergo a ‘conceptual shift’ from thinking of others’ minds as synchronous with the child’s own knowledge (egocentrism), to thinking of others’ minds as reflections of one’s own reality informed by their distinct experiences and perspective.

1.1.3 Unexpected Contents (Appearance vs. Reality)

Another format for FB testing addresses a child’s understanding that beliefs can be informed by an object’s appearance which also may not reflect reality. In the same manner that an object may be believed to be in one location, when it is actually somewhere else, FBs can also arise from a misrepresentation of an object with two aspects. When presented with a series of objects with two aspects (e.g., an object that is really a sponge but resembles a rock), 3-year-olds show little appreciation for this ‘appearance-reality’ distinction. However, 4-year-olds are able to verbally distinguish what it really is from what it appears to be (Flavell et al., 1983).

Combining the appearance-reality distinction with the methodology from Wimmer and Perner's (1983) FB task, Hogrefe and colleagues (1986) designed a novel paradigm which takes advantage of the natural tendency for certain containers (e.g., a matchbox) to hold specific objects (matches). Researchers first asked children what they thought was in the matchbox, to which children uniformly reported its typical contents. It was then revealed to them that the box held chocolates, not matches. After the matchbox was closed once more, a new character (Maxi) enters the scene. The researchers then asked the children to report what Maxi thinks is in the box. Only 20% of 3-year-old children reported Maxi’s FB (‘he thinks there are matches in the box’), whereas 75% of 4-year-olds correctly reported Maxi’s FB. Much like the pattern of results found in change-of-location FB tasks, 3-year-olds’ responses were not random, but instead reported the information that they themselves know is in the matchbox. It is important to acknowledge that the change-of-location and unexpected contents FB tasks differ greatly in methodology, suggesting that results cannot be explained by appealing to extraneous task features inherent to either procedure.

1.2 Conceptual Shifts at four years of age

Children younger than 4 years of age systematically fail to predict other peoples’ actions and desires, when those actions and desires are warranted by a FB (Wimmer & Perner, 1983; Astington & Gopnik, 1991). They fail to select arguments which explain the behavior of someone who has a FB (Bartsch & London, 2000), and they fail to retrodict their own past FB when explaining the reasons for their own behavior (Gopnik & Slaughter, 1991). What are children acquiring on their fourth birthday which cues them to consider the beliefs of other people? Perner (1991) proposes that passing the FB task signposts the emergence of meta-representation (representations of an agent's representation of the world) and, consequently, an appreciation of the influence of propositional attitudes (i.e., a disposition that an agent holds in relation to an object; Fodor, 1978) on the behavior of other people. It is thought that children younger than 4-years either lack the cognitive resources necessary for generating and assigning representations with multiple levels of recursion (e.g., 'I know that she knows that…') or that the requisite cognitive mechanisms have yet to come online. This is not to say that younger children do not express *any* understanding of other people's mental lives. Toddlers frequently use mental state words to communicate a person's desires (e.g., "wants", "hopes", "wishes") or to explain that person's behavior (Repacholi & Gopnik, 1997). However, concurrent use of 'belief' terminology emerges later.

FB understanding is not the only meta-cognitive ability which undergoes a conceptual change at 4-years. This is also the age at which children begin to succeed on 'Level-2' perspective-taking tasks (Flavell, Everett, Croft & Flavell, 1981; Masangkay, 1974). Level-1 perspective-taking tasks require an understanding about *whether* someone has seen an object, whereas level-2 perspective-taking requires an understanding of the *particular way* an object is seen by somebody (i.e., that an object could look like it is something else when viewed from a different vantage point). It has been argued that level-1 perspective-taking can be achieved by mentally drawing an uninterrupted 'line of gaze' between an agent's eyes and a target object, a strategy that eschews cognitively taxing meta-representation (Michelon & Zacks, 2006). It is arguably for this reason that even non-human animals, such as great apes and corvids, are able to coordinate their actions in such a way as to capitalize off of what a conspecific cannot see (Hare et al., 2001; Clayton, Dally, & Emery, 2007). By contrast, it is difficult to imagine a strategy aside from meta-representation which could permit someone to appreciate how an object appears from a different frame of reference. Accordingly, non-human animals show no appreciation of level-2 perspective differences, and only by 4-years-of-age do children make explicit predictions about an agent's actions on the basis of that agent's subjective point-of-view.

It is also around this age that children begin to understand that a single entity can be referred to using two (or more) names or titles. The ability to assign multiple identities to a single entity is referred to as 'intensionality' (Russell, 1987). One body of evidence has shown that children struggle with intensionality (Apperly & Robinson, 1998, 2001; Kamawar & Olson, 2009; Sprung et al., 2007) and, more specifically, that children begin to grasp intensionality only after they master the FB task (Perner, et al., 2002). More recent evidence has revealed that children's performance on a simplified intensionality task is strongly correlated with their performance on a standard FB task, suggesting that these abilities may emerge in tandem (Rakoczy, et al., 2015).

1.2.1 Challenges to the Conceptual Shift Account

Two recent avenues of research have cast doubt on the 'conceptual shift' account. First, evidence from non-verbal FB tasks has revealed that infants may be sensitive to the mental states of other people (Kovács, Téglás, Endress, 2010; Onishi & Baillargeon, 2005; Surian, Caldi & Sperber, 2007). Second, studies using adult participants have shown that they do not always take others' perspective into account when it is relevant to do so, even in relatively simple circumstances (Epley, Morewedge & Keysar, 2004; Keysar, Lin & Barr, 2003). In the following sections, several relevant findings from non-verbal tasks are outlined and adults' performance on mindreading tasks are discussed later, in Chapter 2.

1.2.2 Violation of Expectancy FB Tasks

Recently, researchers have developed new FB tasks devoid of direct verbal questions. In a landmark study, Onishi and Baillargeon (2005) recorded 15-month-old infants' reactions to watching a FB scenario unfold. Infants were first familiarized to the following pattern: an agent puts a toy in one of two colored boxes (green or yellow), then a curtain is raised, occluding the agent from view. When the curtain is lowered, the agent places one hand on the box where she previously hid the toy and reaches in with the other hand. Then she pauses. Next, they watch one of four 'belief-induction' trials wherein the agent sees the toy either remain in the green box or move over to the yellow box. When the curtain is raised this time, the toy either moves to the other box (thus falsifying her belief about its location) or moves and then returns to the original box (thus preserving her true belief). In the following 'test trial' the agent is shown reaching into either one box, or the other.

Here, the researchers make use of a natural tendency for people to look longer at events which the viewer finds surprising. If infants can represent the belief of the agent, they should look longer when the agent acts in a way which is incongruent with the agent's belief (whether false or true). This methodology is known as the Violation of Expectation (VOE) paradigm. Onishi and Baillargeon (2005) found that infants looked significantly longer when the agent acted contrary to her belief. The authors interpreted the results as evidence that by 15-months-of-age, infants expect that an agent's actions are a consequence of their beliefs. Their interpretation constitutes a "rich" interpretation of infant's looking time, as it argues that nothing short of bona fide representational mindreading could explain infants' looking behavior. However, "lean" interpretations of the data have challenged these assumptions, noting that infants' looking patterns can be explained by understanding behaviors rather than mental states (Perner & Ruffman, 2005; Ruffman et al., 2012; Ruffman, 2014; Low & Perner, 2012).

A few VOE FB tasks have reported similar looking patterns with even younger infants (Surian, Caldi & Sperber, 2007; Luo, 2011). The youngest age group that has been reported showing differential patterns of looking comes from a sample of 7-month-olds (Kovács, Téglás & Endress, 2010). In this task, infants observe an agent who watches as a ball rolls behind an occluder and either stays there or rolls off-screen. Then the agent leaves, and in his absence, the ball either leaves, stays, or returns. The agent then returns and the occluder drops, either revealing a ball or revealing nothing. Since the occluder hides the ball from the view of both the agent and the infant, researchers could compare infant's looking times following outcomes that were not surprising to the infant nor the agent, surprising only to the agent, surprising only to the infant, or surprising to both the agent and the infant. They found that infants looked longer following the outcome: 'no ball revealed behind the occluder' when only the agent held a FB that there was a ball behind the occluder, compared to when neither the agent nor the infant believed that there was a ball behind the occluder. The authors interpreted this data as providing support for the rich interpretation: that infants were calculating the agent's belief and expected her to act in accordance with that belief. These findings levy a strong challenge to both the conceptual shift account of human mindreading, and other long held assertions about socio-cognitive development (e.g., that mastery of 'belief concepts' presupposes linguistic competence; Milligan, Astington, & Dack, 2007; Hutto, 2008). The main interpretation acknowledges a meaningful division between implicit and explicit mindreading. While looking time measures demonstrate implicit knowledge of belief, answering direct questions about how other's actions are motivated by beliefs requires an explicit awareness of such concepts. However, before unpacking the theoretical and developmental implications of VOE studies, it helps to consider the evidence from additional non-verbal measures, such as anticipatory looking and helping tasks.

1.2.3 Anticipatory Looking

The eyes behave in a number of predictable ways. Not only do they tend to fixate for longer at events which violates one's expectations, but research has also shown that from around 1 year of age, infants follow the gaze of other people (Moll & Tomasello, 2004) and make online predictions of certain actions by looking to the endpoint of that action (Ambrosini, et al., 2011; 2013). Predictive saccades also target the outcome of actions that have not yet begun, so long as the viewer has been familiarized to a consistent pattern of behavior. This methodology is referred to as Anticipatory Looking (hereafter: AL). Unlike the VOE paradigm, which measures a retrodictive response and therefore can only indicate that the subject did *not* expect something to happen (but not what they would have expected or what they specifically found surprising), studies measuring subjects' AL in a FB scenario reveal their *a priori* expectations. By analyzing the location participants tend to fixate prior to the start of an action, researchers can make more specific inferences about the observer's expectations.

AL was first used to measure FB understanding in young children (ages 2.5 - 4.5 years; Clements & Perner, 1994). By comparing children's AL immediately following a verbal prompt ("I wonder where she's going to look?"), they tested whether younger children's incorrect verbal responses to direct FB questions are indicated by their predictive eye movements. Instead, Clements and Perner found that verbal responses dissociated from AL. Children were shown two different FB scenes, as well as true belief control conditions, adapted from the Sally/Anne FB task used by Wimmer and Perner (1983). In one of the FB scenarios, a protagonist (Sam, a mouse) stores his cheese in one end of a V-shaped tunnel. Sam then travels down to the middle of the tunnel, and children are told that he has fallen asleep. While Sam is asleep, a second character (Katie, also a mouse) enters and moves Sam's cheese to the other end of the V-shaped tunnel. After Katie leaves, children are told that Sam has just woken up and is looking for his cheese. Children then hear the 'anticipation prompt' ("I wonder where he's going to look?") followed by a brief pause wherein children's AL was recorded by a video camera. Thereafter, children are directly asked the following question: “Where do you think that Sam will look for his cheese, first?” Results showed that children’s verbal responses aligned with the predictions from the conceptual shift account: children younger than 4-years tended to select the actual location of the cheese, while 4-year-olds tended to answer in accordance with Sam’s FB. However, their AL responses painted quite a different picture. Children in the younger age group (2 years 5 months) looked at the real location (in line with their incorrect verbal responses), but those children in the 2 years 10 months age group (and older age groups) showed AL toward the location which aligned with Sam’s FB. Here, the same children seem to possess some implicit appreciation for what an agent would do when laboring under a FB, while remaining anchored to the actual location of the object when directly questioned.

In another AL FB task, Southgate, Senju and Csibra (2007) manipulated whether the object was first placed in a new container and then removed from the scene (all while the agent was looking away) or removed from the scene immediately. 25-month-old infants showed correct AL in either condition, suggesting that while an object’s actual location may hold some influence over verbal responses, predictive gaze is immune. Surian and Geraci (2012) tested 17-month-olds who watched an animated shape (a triangle) chasing another shape (a circle) around a screen. At one point, the circle travels through a Y-shaped tube which leads to two boxes. The triangle pauses at the bottom and watches as the circle exits one arm of the tube and disappears inside the nearby box. In familiarization trials, the triangle pursuits the circle into the box. However, in the test trial, the triangle either remains at the bottom and watches as the circle either exits the box and moves into the other box (TB condition), or the triangle turns around and moves off-screen and does not watch the circle change places (FB condition). When the triangle had a FB about which box the circle disappeared into, infants looked to where the triangle would go according to its FB. Alternatively, when the triangle saw which box the circle disappeared into, infants looked to the box where the triangle was hidden. Similar patterns of looking were reported from FB tasks testing 18-month-olds (Thoermer, et. al., 2012), 3- and 4-year-olds (Low & Watts, 2013; Low et al., 2014; Ruffman et al., 2001), 6- to 8-year-olds (Senju, et al., 2010), and adults (Schneider et al., 2012; Senju, et al., 2009). Taken together, AL studies lend support to the notion that FBs are implicitly processed ontogenically earlier than explicitly understood.

Although it can be said that young children simultaneously fail in their verbal predictions while they look toward the correct location, it does not follow that their looking necessarily implies unconscious knowledge. To shed light on the confidence of children's verbal predictions, Ruffman and colleagues (2001) asked children to bet tokens on where they predicted that the agent would go. Results showed that younger children who made correct anticipatory looks, despite making incorrect verbal predictions, were highly confident in those predictions. Specifically, 94% of the children (mean age: 3.40 years) who showed this pattern, bet all of their tokens on their verbal prediction.

It is relevant to note, however, that recent efforts to replicate the results from several AL tasks have been unsuccessful (Kulke et al., 2017; Burnside et al., 2017; Grosse-Wiesmann, et al., 2017). In one study, a sample of elderly participants (Mean age = 71.43 years), were tested using the original stimuli from two AL studies (i.e., Senju, Southgate, White and Frith, 2009, and Surian and Geraci, 2012), resulting in failure to replicate the reported effects from either study (Kulke et al., 2018). This replication crisis casts a disquieting shadow of doubt over implicit mindreading. One possible explanation is that original studies may have suffered a type-1 error resulting from insufficient sample sizes (e.g., *n* = 20, Southgate, et al., 2007). Alternatively, implicit mindreading may simply be less robust than previously thought, or just as robust but challenging to detect. Whatever the case, solving this crisis will require that further attempts to replicate AL studies are undertaken, and may require a reappraisal in the way that AL is measured or interpreted.

1.2.4 Helping Tasks

The majority of non-verbal FB tasks measure predictive or reactive eye movements. However, the eyes are not a perfect window into cognition. They move in a number of automatic ways (e.g., Land, Mennie & Rusted, 1999) but can also be consciously operated, even for communicative purposes. In addition, the amount that saccades, smooth eye pursuits, and head movements can be used to infer one's state of mind is a subject of lasting debate (Krauzlis, 2004). Therefore, there is good reason to consider whether one finds corroborating evidence from alternative non-verbal FB tasks. Following from studies showing that children as young as 14-months-of-age spontaneously provide help to others (Warneken & Tomasello, 2006) researchers have designed FB tasks which prompt children to interact with an agent who is laboring under a FB.

Buttelmann, Carpenter and Tomasello (2009) tested 16- and 18-month-olds in the first active helping FB task. Infants first watch an agent place a toy in one of two boxes. Then the agent leaves the room. While the agent is away, an experimenter enters the scene and moves the object from box A to box B. Then she showed the infant that the boxes can be locked using a pin. She locks both boxes and leaves. When the agent returns, she kneels at the box where she last stored her toy and begins pulling on the lid struggling to open it. 18-month-old infants did not help her attain her *proximal* goal (i.e., the goal of opening box A, as evinced from her actions, not her intentions) but instead went over to box B, unlocked it, and retrieved the agent's toy. 18-month-old infants also chose to help her in different ways depending on what she had previously observed: if the agent watched while her toy was moved from box A to box B, infants were more likely to help her achieve her proximal goal by showing her how to unlock the empty box A. 16-month-old infants were also significantly more likely to help an agent with a FB retrieve the object in box B. However, when the agent remained in the room during the switch, 16-month-olds were split on whether to open box A or box B. Buttelmann and colleagues judged that these results provide clear evidence that 18-month-olds understand FBs, and that extraneous task demands (e.g., lacking the executive resources needed to inhibit their desire to approach the box with a toy inside) may have prohibited 16-month-olds from providing an equally clear pattern of behaviors.

In an extension of their original active helping task, Buttelmann, Suhrke and Buttelmann (2015) tested how infants help an agent who has a FB not about an object's location, but rather about an object's dual-function. 18-month-old infants were first introduced to an object, and shown that it has two functions (e.g., a sponge which has been painted to look like a rock, a pen that is shaped like a branch, a box with the appearance of a book, and a duck toy attached to a brush). The rock/sponge condition will be used to illustrate the procedure. First, infants watched an experimenter place the rock/sponge on a high shelf. Then, an agent enters the scene, announces that she wants something, and then struggles trying to reach the rock/sponge. Then, she looks to the experimenter and asks for help, to which the experimenter shrugs and pretends to be busy. The agent looks saddened, and as she gives up, the experimenter removes an occluder which, to the infants' surprise, reveals two objects: a regular looking sponge and a regular hard rock. Then the experimenter asks the infant to help the agent get the object that she wants. The idea is that, if the agent was not present when the experimenter showed off the objects' dual-function (i.e., a FB trial), then infants should not expect that she knows that the rock is also a sponge and should offer her the rock. If the agent was present for this demonstration (i.e., a TB trial), then infants should expect that she is looking for a sponge and offer the agent the sponge.

When the results from all four types of objects are presented together, they support this prediction: in the FB conditions, 64.6% of infants chose to offer the object matching the appearance of the object while in the TB conditions, 66% of infants chose to offer the object matching the true function of the object. However, when the results from each object are analyzed separately, the pattern becomes far less clear. Only when reasoning about a duck/brush object do infants make significantly different selections based on the agent's belief, with 66.7% of infants offering the duck in the FB condition, while 20% of infants doing so in the TB condition. When the agent has a FB about a rock/sponge, 83% of infants offered the rock, but so do 43% of infants when the agent has a TB. When the agent has a FB about a branch/pencil, 80% of infants offer a branch, and 60% of infants do the same when the agent has a TB. Finally, when the agent has a FB about a book/box, 36% of infants offer the book, and 22% do the same for an agent with a TB.

These results have been heralded by some, as evidence that infants can reason about FBs about an object that has two identities. However, it is not clear whether an object having two functions offers the same conceptual challenge as an object having two identities. In a study comparing children's answers to questions about either multiple identities or multiple function, 3 to 5-year-old children performed significantly better on questions about dual-function (Perner, Mauer & Hildenbrand, 2011). Specifically, children demonstrated an understanding that a key that open's one box (a snake's cage) can also open another box (the lion's cage), signposting competency in their understanding that one object can serve two distinct functions. However, these same children struggled to understanding that the key that opens the snake's cage is the *same* key that opens the lion's cage, the appreciation for which is pivotal to the appreciation of multiple identities. This disconnect leads one to maintain a skeptical stance when applying the results from dual-function tasks to a discussion about the understanding of multiple identities. It is more likely that the task from Buttelmann and colleagues (2015) taps a conceptual structure having more in common with an appearance-reality distinction than an identity task.

One methodological limitation of the Buttelmann helping task, is that infants are only permitted/prompted to help the agent after the agent has made a failed attempt to either open box A, or to reach for the high shelf. This leaves open the question about if and when during the procedure, infants planned their helping actions in light of the agent's FB. Knudsen and Liszkowski (2012) tested whether infants make forward inferences about an agent's erroneous action and intervene *before* help is solicited. 18- and 24-month-old infants were seated at a table across from an agent. A row of upside-down cups was arranged on the table with an object hidden inside one of them. First, the agent began deliberately looking for an object under each cup and responded positively when they found the object underneath. Next, the agent says that she needs to leave the room and stores the object under one of the cups and leaves. A second experimenter appears and switches the object into the cup on the opposite side of the table, and leaves. When the agent returns, she walks in a predesignated route toward the table. During her walk, the number of times that the infant pointed to the cup where the object had been moved was recorded. In this condition, 43% of 18-month-olds and 62.5% of 24-month-old infants made spontaneous pointing responses to the location housing the cup. However, when the agent had watched the switch phase (TB condition) or had previously expressed a different goal (cleaning the cups, rather than looking for an object under them), infants produced significantly fewer pointing responses. Importantly, unlike the Buttelmann studies which reported high attrition rates, Knudsen and Liszkowski (2012) did not exclude any infants on the basis of their responses. These results provide clearer evidence that infants make predictions that an agent with a FB will act in error and respond to this prediction by directing their actions to indicate the location hiding the object the agent is looking for.

Allen (2015) raised a crucial point about the underlying assumptions made when interpreting results from helping tasks. He notes that, in the Buttelmann et al. (2009) study, it is assumed that infants who correctly help the agent open the box with the toy inside, do so by having first reasoned about the agent's FB. Allen stresses the importance of considering alternative non-mindreading explanations for what would otherwise be considered positive evidence for FB reasoning. For example, since the Buttelmann et al. (2009) study's results depend on infants behaving differently in TB and FB scenarios, it is important to be sure that there are not alternative plausible reasons to act differently in these two conditions.

One alternative reason for infants' differential actions could be that, in the FB condition, they are reasoning about behavior instead of belief. If they are operating on the heuristic that people tend to look for objects where they last put them, then this reasoning alone could motivate them to open box B (Low & Perner, 2012). Another alternative is that, in the FB scenario, the second agent behaves in a 'sneaky' manner, which may have an influence on the underlying social context. Since children routinely learn by playing games, this emphasis on *hiding* the toy in order to trick the agent may trigger familiar 'hide-and-seek' heuristics.

To demonstrate some the non-mindreading control conditions that are necessary before the results from Buttelmann et al. (2009) can be interpreted as revelatory of mindreading, Allen (2015) tested 3-5-year-old children in a replication of the FB condition from Buttelmann et al. (2009), and two control conditions (*clairvoyance* condition and *hands-full* condition). The clairvoyance condition follows the procedure of the FB condition, however, when the agent returns to the room, she first goes to box B, where the object has been moved. In the hands-full condition, the agent and child engaged in a hiding game where they filled one box (box A) with toys and left box B empty. Then the agent leaves the room to get more toys to hide. While she's away, a second experimenter suggested that they move the toys from box A to box B. Unlike the other conditions, this experimenter did not show any of the deceptive behaviors indicating that they were playing a trick on the agent. When the agent returns, her hands are full of toys. She then goes to the box she thought was empty (box B) and fails to open it.

Results showed that the majority of children (84%) helped the agent by opening the box with the toy inside. However, the majority of children (82%) also opened this box in the clairvoyant condition. This casts doubt on the original findings, which stated emphatically that children must understand FBs in order to pass this test. If the agent's belief was central to the children's decision, they should have reasoned that the agent in the clairvoyant condition does not know that the box is not empty and should therefore help open the empty box. In the hands-full condition, 52% of children chose to open the box with the toy inside. When split into three groups based on age, a clear developmental trend emerged. 88% of 3-year-olds opened the box with the toys, whereas 43% of 4-year-olds and 25% of 5-year-olds opened this box. This finding shows that it cannot be taken for granted what young children (let alone infants) are thinking especially in the absence of adequate control conditions testing non-mindreading alternative interpretations.

1.2.4.1 Replication of FB helping tasks

Recently, Priewasser and colleagues (2017) attempted to replicate the findings from Buttelmann, Carpenter and Tomasello (2009). Although they were able to replicate the pattern reported in the original FB condition (i.e., infants tended to help the experimenter open the box with the toy), they could not replicate the opposite pattern in the TB condition. Instead, children seemed to be unsure of what to do and consequently showed no difference in their selection between the empty or full box. This is problematic for Buttelmann's interpretation, as success in the FB condition (and not the TB condition) leaves open the possibility that infants made their choice by attending only to the agent's goal of retrieving the toy, rather than by taking the agent's belief into account. According to Buttelmann et al. (2009) "[infants] must understand *why* the person is doing what he is doing in terms of his beliefs about the world – in order to interpret his goal correctly and so to provide the appropriate help (pg. 338)." This implies that belief understanding is a necessary antecedent of goal attribution. However, if beliefs are pivotal to infants' interpretation of an agent's goals, then infants should help the agent open the empty box in the TB condition, in the same way that 14-month-old infants provide instrumental help to others who are struggling to achieve similar goals (Warneken & Tomasello, 2006, 2007).

To be sure, there are a number of reasons why one would be confused by the agent's actions in Buttelmann and colleagues' (2009) TB condition. First, it is highly unusual that someone would abandon a toy that they previously wanted, in favor of an empty box. Perhaps Priewasser and colleagues' infants had reason to doubt whether the agent remembered that the object was just moved to box B and were attempting to remind him. Or perhaps infants in Buttelmann's TB condition inferred that the agent must no longer be interested in the toy, or that the toy must belong to someone else. In an altogether different vein, perhaps infants did not reason about beliefs at all, but instead about *reasons* for actions. If so, then infants may have chosen the box containing the object because that is the only object that they have good reasons for seeking.

To distinguish between these alternative interpretations, Priewasser and colleagues (2017) ran a follow-up study and adapted the helping task to include a third box. In what the authors call the "old" FB and TB conditions, the procedure is identical to Buttelmann et al., (2009) with the exception of the presence of the third box. In the "new" FB and TB conditions, when the agent returns she instead goes to the third box and tries to open it. The authors reasoned that, if the infant is truly taking the agent's belief into account, then in the new FB trial, they should help the agent open the third box, as opposed to box 2, where the toy really is. The agent has no reason to believe that the object is in box 3, since the last place she saw it was in box 1. Therefore, a mindreading infant should not expect that she is trying to find her toy; she must be trying to open box 3 for some other reason. However, if the infants are not reasoning about the agent's beliefs, but instead thinking more like teleologists, and therefore evaluating the goal of the agent's behavior in light of her having good reasons for her actions (see Perner, Roessler, & Priewasser, 2018; Perner & Esken, 2015) then they should open box 2 and retrieve the toy regardless of whether the agent looks in box 1 or box 3 when she returns, because there is no good reason for the agent to open box 3.

Results were in line with the teleological prediction. Infants in the "new" FB condition chose to open the box with the toy, rather than help open box 3. In addition, infants in the "old" FB, and "old" TB condition almost never chose to approach the third box, and results replicated those from their first experiment: infants chose box 2 more than box 1 in FB conditions but were equally likely to choose either of these two boxes in the TB condition. These results cast doubt on Buttelmann's interpretation, that infants who appropriately helped the agent with a FB, did so because they used the agent's belief to explain their behavior.

Crivello and Poulin-Dubois (2017) ran a conceptual replication of the Buttelmann et al. (2009) study. They targeted two additional problems. First, the original study reported a high attrition rate: 54% of infants did not show any helping behavior and were excluded from the final analysis. To minimize attrition, Crivello and Poulin-Dubois administered the task on a table, rather than on the floor, in order to shorten the distance infants would need to travel to make a choice (an arm reach rather than walking or crawling to a box). This had the desired effect; only 28% of infants failed to produce a helping response. Second, the original study may have been underpowered (*n* = 12 in the FB condition) thereby inflating the risk of committing a type-1 error. Crivello and Poulin-Dubois tested a much larger sample (*n* = 41) and found that only 37% of infants helped open the full box in the FB condition (by contrast, the original study reported that 72% opened the full box in the FB condition). Furthermore, in a second experiment with even higher statistical power (*n* = 91) and a matching TB condition, infants did not perform above chance in either condition (56% and 38%, respectively), and responses in either condition did not differ significantly. Regardless of the agent's belief, infants tended to choose the box containing the toy.

Although conceptual replications are useful for testing whether original results were influenced by extraneous task variables, it is equally important that strict replications are conducted to confirm the reliability of the original finding. Powell and colleagues (2017) ran a strict replication of the Buttelmann, et al., (2009) FB condition using a high-powered sample (*n* = 94) of 17-20-month-old infants. They observed above chance performance, with 62.8% of infants selecting the full box. Taken at face value, this evidence replicates the pattern of responses in the FB task of the original study. However, there remains at least two problems. First, it is not clear why infants pass this task when it is set up on the floor, but not when the it is set up at a table. Second, without a clear difference in infants' behavior from a TB condition, it cannot be determined that infants are not simply more likely to approach a box with a toy in it.

Where do these mixed results leave us regarding interpretations and directions for future research? In the hunt for a resolution to the debate between a "rich" and "lean" interpretation of infants' non-verbal responses, Buttelmann and colleagues' (2009) results are problematic (Allen, 2015). Much like other non-verbal measures of FB reasoning (e.g., VOE and AL studies), the failure to replicate original effects suggests that active helping tasks may not provide a platform for detecting stable effects. Alternatively, this criticism may only apply to Buttelmann and colleagues' (2009) active helping task, rather than to *all* possible active helping tasks. Unlike VOE and AL FB tasks, FB helping tasks are relatively few in number. These findings are meaningful to the developmental literature, as they serve a descriptive purpose in measuring infants' and young children's sensitivity to others' goals.

1.2.5 Summary

Next, the discussion turns to one main line of interpretation about infants' and older children's mindreading abilities. On the one hand, children younger than 4 years show robust failure to appreciate others' FBs on a variety of verbal tasks (Wellman, Cross & Watson, 2001). On the other hand (all replication matters aside), multiple non-verbal measures show that infants make predictive eye movements which reflect the understanding of others' behavior. In the following section the "rich" interpretation of this data is presented, which argues that FB reasoning comes online early in ontogeny.

1.3 The Early Mindreading Account (EMA)

When describing the Early Mindreading Account (hereafter: EMA), I will routinely group together multiple theories which fall more or less under the same umbrella (e.g., Baillargeon, Scott & He, 2010; Leslie, 1994). In summary, the EMA views certain evidence of non-verbal FB tasks as litmus tests for mindreading. According to this account, mindreading emerges at as early an age as there is positive evidence from non-verbal FB tasks. At present, that would imply that as early as 7-months of age, humans possess an operational (and representational) mindreading ability, as this is the youngest age group reported showing successful looking in a VOE task (Kovács, Téglás, & Endress, 2010). To begin, I will focus on the cognitive system(s) which give rise to mindreading, as described by the EMA.

Baillargeon, Scott and He (2010) describe the cognitive mechanisms which develop to form a mature modular mindreading system. This system is the product of two subsystems which perform different specialized functions for belief reasoning and only one of which is part of our core cognition. Subsystem-1 (SS1) is the innate one, and it performs two functions necessary to attribute rudimentary mental states: processing motivational states and reality-congruent states. The former allows an infant to appreciate the motivation behind an agent's actions (e.g., she has the *goal* of obtaining her toy). Assigning an agent, a goal allows one to predict their behavior: the agent will act in service of that goal. The latter allows the infant to keep track of the information that an agent does or does not know. This operation has been described as a *masking* mechanism. When an agent is distracted by a phone call, or has her vision occluded by a curtain, the masking mechanism of SS1 allows the infant to separate shared information from privileged information to which only the infant had access. An infant who is operating only on SS1, would be sensitive to an agent's goal (say, retrieving her toy), and also the information of which the agent is ignorant (that the toy has been moved in the agent's absence). This information alone could license an implicit sensitivity to mental states.

Could an infant with only this mechanism succeed in non-verbal FB tasks? Although SS1 provides a mechanism suitable for one to be sensitive to an agent's preferences, goals and informational access, it does not allow for the abductive reasoning that permits a mindreader to judge how an agent will act, based on that collection of information. To explain further, an example from Scott and Baillargeon (2009) will be borrowed. Consider an agent who has demonstrated a preference for a toy. Of the available hiding spots, only the infant participant knows where the agent's toy is hidden. The agent was not present when it was hidden, and so, an infant operating only on SS1 can mask the actual location of the toy, thereby providing an accurate representation of what the agent knows (or rather, does *not* know). Now, consider the same circumstance except that the agent has a FB that her toy is in the last place where she stored it. The agent is not aware that her toy was moved while she was away. Here, SS1 is not suitable for representing the agent's FB. Masking the actual location of the object does not specify where the agent falsely believes her toy is. In this case, Subsystem-2 (SS2) is needed.

SS2 is described as a *decoupling* mechanism. When an infant watches an agent develop a FB, two distinct representations of the events in the scene are created in the infant's mind: the infant's knowledge and the agent's belief. In order for the infant to predict what an agent with a FB will do, the infant needs to decouple from their default mode of representation (i.e., biased toward egocentric predictions) and in its place adopt a representation of the agent's representation. This allows the infant to appreciate reality-*in*congruent informational states, thereby permitting the FB understanding that facilitates their responses to VOE, AL and helping FB tasks. Given this explanation, SS2 must come online as soon as infants show differential looking or helping behavior in non-verbal FB tasks.

1.3.1 Performance Demands & Competence Shifts

The EMA postulates that children's failure to provide correct verbal responses in standard FB tasks is caused by artefacts of the FB question, overwhelming their limited executive functioning and/or language skills. Executive functions (hereafter: EF) include a series of mental skills which collectively allow for the regulation of attention and facilitate control over our actions. EFs are integral to problem-solving (Zelazo et al., 1997) and have long been studied in relation to mindreading (Perner & Lang, 1999). Homing in on its relation to mindreading, the main source of disagreement in mindreading research is the degree to which EF are necessary for the *emergence* of mindreading, or merely the *expression* of mindreading (Moses, 2001; Russell, 2013). How can one determine which of these accounts is more likely? One way to do so would involve testing children on verbal and non-verbal FB tasks, as well as a battery of tasks which tax their EF in specific ways. If EF are critical for the emergence of FB understanding, then individual differences in EF should correlate with performance on non-verbal and verbal FB tasks. If, on the other hand, EF are only relevant for the expression of FB understanding then EF should only correlate with verbal FB tasks.

The relation between mindreading and EF is clear in one regard: children's performance on verbal FB tasks and EF are significantly correlated (Perner & Lang, 1999; Moses, Carlson & Sabbagh, 2005; Carlson, Zelazo & Faja, 2013; Garon, Bryson & Smith, 2008; Devine & Hughes, 2014). However, this does not help in distinguishing the emergence account from the expression account. As it currently stands, the evidence that mindreading is either facilitated by EF or processed silently in the child's mind until EF skills develop, is mixed.

1.3.1.1 The Expression Account

If children have the cognitive abilities necessary to represent mental states then there may be ways of simplifying the verbal FB task in order to diminish EF demands, thereby allowing children to express their conceptual knowledge. One way in which that can be done has already been discussed: eliminate verbal responses. In a correlational investigation of linguistic abilities and FB understanding, Low (2010) tested 3 and 4-year-olds in an AL FB task, along with a battery of verbal FB tasks, and assessments of verbal abilities, complement syntax understanding, and cognitive flexibility. He found that performance on the AL FB task was correlated with verbal FB reasoning. However, of the remaining variables tested, performance on verbal FB tasks was correlated with performance on measures of verbal ability, complement syntax, and cognitive flexibility (and this correlation was maintained even when AL FB competence was controlled for). None of these correlations were found linked to performance on the AL FB task. Taken together, this pattern of results suggests that while non-verbal FB understanding is linked to verbal FB understanding, they are far from equivocal. The relationship between AL FB understanding and verbal FB understanding was also reported in a longitudinal study where 18-month-olds' individual differences on the AL FB task predicted their verbal FB understanding at 48 months of age (Thoermer et al., 2012). However, these findings were recently challenged by a cross-sectional study of preschool children's FB reasoning, EF and language abilities (Grosse-Weismann et al., 2017). Results revealed a dissociation between performance on verbal FB tasks and AL FB tasks: 3-year-olds fail verbal FB tasks and 4-year-olds pass; however, both age groups pass AL FB tasks. In addition, while performance on the verbal FB task correlated with syntax ability and EF, such correlations were not found with the AL FB task, and the AL FB task and the verbal FB task were not correlated. These results favor the expression account, but for verbal FB reasoning only. Non-verbal FB tasks appear to be disconnected from EF skills, as well as from verbal FB tasks.

Another methodology for testing the expression account comes from verbal FB tasks which have been adjusted in order to reduce the EF demands placed on the child. Another way to reduce EF demands is by removing the object from the scene rather than transporting it from one box and into another, thereby removing the prepotent response to report the location where the agent *should* go. In addition, simply generating a response to the critical test question may pose its own extraneous challenge to young children's processing abilities, as they are required to interpret the question, hold it in their working memory, and produce a response. Setoh, Scott and Baillargeon (2016) tested 2.5-year-old children on a verbal FB task which aimed to reduce the combined processing demands resulting from these task artefacts. After the protagonist (Emma) leaves her apple inside one of two boxes and exits the scene, a second character retrieves the apple from the box and removes it from the scene altogether. Note that the critical test question: "Where will Emma look for her apple?" is still accompanied by a choice between two boxes. In addition, children are asked two practice questions in the lead up to the critical test question, which primes their responses to "where" questions.

In this simplified verbal FB task, 2.5-year-old children performed significantly above chance, with 78% of children pointing to the box that the protagonist falsely believed contained her apple. Even when children were split into two groups according to age (33-month-olds and 30-month-olds) both groups performed significantly above chance and did not differ significantly from each other. The authors interpret this finding as evidence that FB understanding was there all along, but it is only when extraneous demands are removed from the standard FB task that children reveal their underlying mindreading competency. Three follow-up experiments explored the specific contribution that low inhibition and practice trials made to children's ability to express their FB understanding. Originally, the practice questions revealed two objects (e.g., an apple and a banana) and asked, "where is Emma's apple?" This question not only prompts children to respond to "where" questions, but also trains children to make a choice between two possible options. In experiment 2, the practice questions were instead accompanied by a single image of the correct object. 33-month-olds still performed above chance, but 30-month-old infants now performed at chance level. Experiment 3 and 4 focused on 33-month-olds. In experiment 3, the authors were interested in whether they benefitted from two practice trials, or if only one would suffice. Children now performed at chance levels, indicating that the cumulative effect from multiple practice trials facilitated their success. In experiment 4, children were prompted by the same practice trials as in the first experiment, however, the second character did not remove the object from the scene but placed it in the other box. Now, 33-month-olds performed significantly below chance.

Setoh and colleagues interpreted their results as evidence that FB reasoning develops earlier than 4-years, and that the 'conceptual shift' at this age is actually a shift from insufficient to sufficient EF resources which facilitate the expression of that reasoning. However, Rubio-Fernandez, Ettinger and Gibson (2017) point out three criticisms of this data. First, they note that reducing processing demands in this way opens their results up to solutions which circumvent FB understanding. For example, in the two practice trials children are prompted to point to where an object is in order to prepare them for the critical question: "where will Emma look for her apple?" It is assumed that these practice trials have the specific effect of priming "where" questions and choices between two options, but it is perhaps more likely that they prime children to point toward a location that is associated with an object. In the test question, children are faced with a choice between the last location the object was placed, and a location which never housed the object. Since there is no competing reason for choosing the latter location, children may have pointed toward the former simply because that's where the apple was most recently stored. Rubio-Fernandez and colleagues also argue that these results are not consistent with an EMA account. The EMA predicts that both response generation and inhibitory control lead to increased processing demands. Consequently, reducing either of these demands should result in improved performance. However, Setoh and colleagues observed below chance performance when response generation demands were reduced (Experiment 4). These children were no better at responding to a standard FB tasks with or without practice trials. Furthermore, Rubio-Fernandez and colleagues note that the EMA could have claimed support for their account no matter how children performed in Experiment 4. If instead they performed at, or above chance, the EMA would conclude that response generation places heavy demands on children's EF.

In a rebuttal, Scott, Setoh and Baillargeon (2017) state that their prediction is informed by evidence showing that 3-year-olds perform at chance in low-inhibition verbal FB tasks (Wellman, Cross & Watson, 2001) and they perform below chance on high-inhibition verbal FB tasks despite several practice "where" questions (a finding that was confirmed by the results from Experiment 4; Yazdi et al., 2006). Thus, their account specifically predicts that young children should only perform above chance when both of these processing demands are controlled for. In response to Rubio-Fernandez and colleagues' criticism regarding confounding priming effects from "where" questions, Scott and colleagues ask why this should only be the case when two "where" questions are asked, since the results from experiment 3 (where only one practice question is asked) show that children revert to chance performance. This rebuttal does not refute the challenge as it remains possible that young children are only misguided when they detect a consistent pattern of "where" questions. Furthermore, this evidence does not address another alternative hypothesis: that FB understanding relies on sufficient EF skills not only to express that understanding but to facilitate it. In the next section, this competing account of EF for FB understanding, is unpacked.

1.3.1.2 The Emergence Account

Perhaps it is not the case that a child possesses FB understanding but lacks the EF skills to communicate their reasoning. Instead, perhaps the rate at which EF abilities mature permit proportionately abstract conceptual thinking. After all, there are executive demands which are inherent to FB reasoning, but also separate from executive demands taxed by the expression of that reasoning. For instance, representing an agent's mental state (in either a verbal or non-verbal FB task) involves inhibition. One must inhibit the way that the environment is seen by the observer in order to consider the way it is perceived by an agent. Another EF skill (working memory) is also taxed by non-verbal FB reasoning, as one cannot appreciate that an agent has a FB about, say, an object's location, without recalling where that object was when the agent last saw it. Therefore, an 'emergence account' posits that the development of EF skills is necessary for the acquisition of FB concepts themselves. The emergence account is mainly supported by longitudinal studies showing that younger children's individual differences on EF tasks predicts later performance on verbal FB tasks (Carlson, Mandell & Williams, 2004; Hughes & Ensor, 2007). The emergence account is also supported by evidence showing that children with autism spectrum disorder (ASD), who show general deficits in EF, are also impaired on FB tasks (Hughes, 1998; Moses, 2005).

In a recent study, Carlson, Claxton and Moses (2015) investigated the EF-FB reasoning relationship in order to evaluate the emergence and expression account. They reasoned that the expression account would predict that EF skills would correlate with FB tasks, but only when the FB task poses considerable demands on EF abilities. By contrast, the emergence account would predict that EF skills will correlate with all FB tasks, regardless of the relative EF task demands. Carlson and colleagues (2015) tested preschool aged children on four tasks where mental state attribution was necessary. Half of the tasks were 'high EF tasks': a verbal change-of-location FB task and a verbal appearance-reality task. The other half were 'low EF tasks': a 'think-know' task where children must reason about which of two characters to trust, either one who 'thinks' an object is in one place, or one who 'knows' an object is in a different place, and a 'sources of knowledge' task where children are shown an object (either by feeling it, seeing it or being told about it) and are then asked to explain *how* they knew what the object was. Children were also tested on a series of EF tasks, and tests of verbal ability. Results revealed that individual differences in the EF tasks correlated with *both* high and low EF mental-state tasks. This finding supports the emergence account and casts doubt on the expression account. However, since their results are correlational, it is not possible to determine where the causal power lies. Thus, these results leave open the question of whether EF skills are responsible for mental state reasoning, or if thinking about other's mental states assists children's development of EF abilities.

In a recent meta-analysis of 102 studies on the relationship between mindreading and EF, Devine and Hughes (2014) found that only a hybrid 'expression-emergence' account could provide a suitable explanation for the variation in effect sizes across studies. That is, the strength of the association between FB reasoning and EF was found to vary according to the type of FB task (lending support to the expression account). However, the strength of association with EF was the same for first- and second-order FB tasks, the latter of which involves multiple levels of recursive thinking (i.e., '*he* knows that *she* knows that…') and should therefore make more taxing demands of EF. In addition, the results from the available longitudinal studies (n = 10) revealed an asymmetry in the direction of association: early EF skills predicted later FB reasoning more strongly than the reverse (lending support to the emergence account). Taken together, these results do not allow for either of the competing accounts to be ruled out and can be best explained by proposing a hybrid combination of the emergence and expression accounts. This conclusion is challenging to incorporate into an argument for or against the EMA account's explanation that children's limited EF skills explain how infants pass non-verbal FB tasks while 3-year-olds fail standard verbal FB tasks. At this time, the available evidence cannot rule out the EMA's expression account or the competing emergence account, however, the nature of the relationship between EF and FB reasoning suggests that the link between the two is more complex than an EMA account predicts.

1.4 Lower-level Alternative Hypotheses for non-verbal FB understanding

Next, the focus shifts to alternative theories explaining the difference in children's performance on non-verbal FB tasks. While the EMA views success on non-verbal FB measures as highly continuous with success on verbal FB tasks, the mirror image of that is the interpretation that looking data are nothing like verbal responses and should not be conflated. These theorists favor a leaner and more cautious interpretation of non-verbal data.

According to the EMA, infants' eye movements in non-verbal FB tasks reflect the inner-workings of a psychological reasoning system. Others maintain a skeptical stance toward the assumption that infant's non-verbal success is necessarily the result of a cognitive system operating specifically for that purpose. Instead, eye movements may be the result of domain-general lower-level statistical learning processes (Perner & Ruffman, 2005; Ruffman, et al., 2012; Ruffman, 2014). According to such low-level explanations, humans possess an early developing capacity for statistical learning, and this allows them to cluster and categorize behaviors. As infants observe a wider variety of behaviors, they begin to codify behavior-outcome relations into 'behavior-rules': heuristics governing reliable patterns of behavior.

Behavior-rules allow for the prediction of future behaviors, granting the observer sensitivity to atypical behavior. Critically, this model operates independent of any appreciation for the mental states motivating behaviors. According to the 'behavior-rules' account, it is at least equally plausible that infants predict an agent's subsequent behavior (or detect an agent's atypical behavior) by applying previously learned rules. From an information processing standpoint, the behavior-rule account may present a more optimal strategy than the appreciation of mental states, as it bypasses the need to make cognitively taxing inferences about the internal propositional attitudes motivating an agent to act. For example, in Onishi and Baillargeon's (2005) FB task, it cannot be determined from looking time whether infants' expectations were informed by a meta-representation of her belief, or by applying a heuristic informed by experience which expects her to follow the behavior-rule: 'people tend to look for things in the last place that they saw them' (Perner & Ruffman, 2005).

The behavior-rule account can explain positive evidence from several non-verbal FB tasks. Nonetheless, the 'behavior-rules' account has been hotly debated not only within the context of infant mindreading studies (Traüble et al., 2010; Surian & Geraci, 2012; He et al., 2011) but also in studies of animal mindreading (Penn & Povinelli, 2007; Lurz, 2009; 2011; Heyes, 1998; Karg et al., 2015). Evidence from looking data presupposes that such behaviors are necessarily indicative of some degree of conceptual understanding. Without the use of additional explicit measures, it is not clear whether the participant predicts the behavior of an agent on the basis of a single step: from knowledge about relevant invariant features of the agent ➔ the agent's subsequent behavior, or on the basis of multiple steps: from invariant features of the agent ➔ the agent's mental states ➔ the agent's subsequent behavior (Povinelli & Vonk, 2006). This issue has been termed a 'logical problem' because for every non-verbally measured mindreading hypothesis, there exists a complimentary behavior-reading hypothesis.

Low and Wang (2011) argued that research on infants' FB tracking has, to some extent, addressed the behavior-rule challenge. Re-examining the claim that behavior-rules are computationally simpler than mental state reasoning, Low and Wang devised a formula to compare the number of behavior-rules versus the number of 'mindreading' rules in a given FB task, finding that the number of 'mindreading' rules were fewer. Low and Wang argued that children's success across multiple belief-use situations (i.e. what the agent will do) by multiple belief-formation situations (i.e. how the false belief is produced) would require stacking on more and more behavior-rule explanations than mindreading-rule explanations. Ruffman (2014) challenged Low and Wang and suggested that the formula does not account for infants' ability to make rule generalizations (e.g. someone who says an object is in X also tends to point to X and to look for the object in X). If infants can generalize these rules across contexts, the number of behavior-rules required shrinks below those required for mindreading rules. Thus, until positive evidence from a non-verbal FB task emerges, which presents subjects with a novel circumstance, as well as a novel mode of perception such that a behavior-rule could not have been previously generated to account for the agent's behavior (for further discussion see Lurz, 2011), the behavior-rule hypothesis will remain a valid competing theoretical explanation for infants' success across different non-verbal FB tasks.

1.4.1 Behavior-Rules: An unfalsifiable claim, or the null hypothesis?

In a critique of the behavior-rule account, Scott (2014) points out that this account remains subject to serious theoretical shortcomings. First, the behavior-rules account necessarily relies on post hoc explanations. It can explain all possible patterns of data since all behaviors that an agent could make in a FB task can be generalized to similar behaviors observed in others. Scott (2014) goes on to argue that the behavior-rule account does not provide a cogent theoretical framework for generating testable predictions. Although most non-verbal FB tasks present an agent whose actions can be interpreted according to multiple conventional behavioral patterns, a behavior-rule account does not make *a priori* statements about which rules will be used to predict subsequent behaviors.

Scott (2014) notes that because of the post hoc nature of behavior rule explanations, Ruffman (2014) makes different claims about which statistical pattern infants expect an agent to act upon. For example, in one study (Scott & Baillargeon, 2009) an agent repeatedly demonstrates that she has a goal: she wants to hide her key inside a penguin toy that is always presented in two pieces, as opposed to an identical 1-piece penguin toy. In a FB condition, an experimenter assembles the 2-piece penguin during the agent's absence. Next, the experimenter puts a box over the 1-piece penguin and puts a transparent box over the assembled 2-piece penguin. Infants showed longer looking when the agent returned and reached toward the transparent box housing the assembled 2-piece penguin. The behavior-rule explanation for this pattern of looking is that infants looked longer because they had never seen the agent reach for an intact penguin in previous trials (a new agent-object relation). Furthermore, they did not show similarly longer looking when the agent reached toward the other box because the agent acted the same way as in previous trials, i.e., she reached away from the intact penguin (an established avoidant agent-object relation).

This explanation is problematic for two reasons. First, Scott (2014) argues that one could also suggest that infants should have expected the agent to reach for the transparent box because in all previous trials, the agent always reaches for a visible penguin. Second, the behavior-rule interpretation does not adequately explain all patterns of looking. For example, in another VOE FB task (Surian, Caldi & Sperber, 2007) an agent watches her preferred food as it is placed behind one of two barriers (TB). Infants look longer when the agent approaches the non-preferred food hidden behind the other barrier which, according to Ruffman (2014), reflects the fact that infants had never seen the agent approach this food before. However, when the agent is ignorant as to which barrier hides which food, infants spend equal time looking regardless of which barrier she approaches. Scott (2014) claims that, according to the behavior-rule account's explanation of their differential looking time in the TB condition, infants should look longer when the agent approaches the non-preferred food in the ignorance condition as well, as this too is a novel agent-object relation. Therefore, *prima facie* behavior-rules presents a more theoretically defensible claim compared to rich interpretations of non-verbal data; however, behavior-rules are unequipped to explain the full breadth of looking patterns and have not yet opened itself to opportunities for future investigations by generating testable hypotheses.

One could argue that the burden of proof does not fall on the behavior-rule account (or any other alternative non-mentalistic interpretations), as their position represents the null hypothesis vis-à-vis domain-specific mindreading hypotheses. Scott (2014) may be committing a logical fallacy of false equivalence when she suggests that a behavior-rule account is not a "fair alternative to the mentalist account" (p. 302). Hypothetically, evidence against a behavior-rule account's specific predictions about which rule an infant would base their expectations off of during a particular circumstance, would not constitute evidence in support of a domain-specific mindreading account. Nor does it refute the behavior-rule account more generally.

1.4.2 Domain-general reasoning and perceptual novelty: A falsifiable claim

A further objection to the EMA account is made by Heyes (2014). She remains unconvinced that infants have demonstrated evidence that their psychological reasoning goes beyond a representation of "colors, shapes, and movements, rather than as actions on objects by agents" (p. 6). Heyes suggests that all existing non-verbal FB tasks using infant samples can be explained using the following low-level processes. First, Heyes notes that infants look longer at events that they find perceptually novel for various reasons (e.g., any novel configuration of color, shape and movement), and this novelty effect can also be triggered by features of the task which are not manifestly novel but are associated in some way with the appearance of other objects (e.g., by virtue of being a similar shape or hue). Second, because infants struggle to hold information in their working memory, it cannot be assumed that just because they have observed familiarization trials, they will be able to keep that information in mind to compare against subsequent test trials. Third, infants are prone to 'retroactive interference' effects from salient events. These interferences can effectively wipe clean an infant's memory of all preceding information in the task and can be caused by something as banal as an agent re-entering a scene. Given these three limitations, it is impossible to design a coherent FB narrative that controls for all potential opportunities to detect novelty, so as to be sure that longer looking can be taken to reflect genuine responses to an inconsistency between the agent's FB and her subsequent action.

Scott and Baillargeon (2014) have raised objections to Heyes' (2014) novelty framework. They argue that Heyes' account ignores 20 years of research collectively showing that infants represent external events in terms of agents who operate on principles such as rationality, efficiency and consistency, rather than simply an arrangement of colors, shapes and movements (for a review see Baillargeon et al., 2016). Scott and Baillargeon also argue that Heyes makes unfounded claims about the fragility of infants' ability to hold events in mind, as there is no supporting evidence for the hypothesis that a character entering the scene will have such a highly disruptive effect, or that infants' memories fade over a brief delay. Much like the behavior-rule account, this account cannot adequately explain why infants look longer at one event than another. Since there is no existing task that can remove all opportunities for novelty detection, infants have reason to detect novelty in every condition. This leads to a feature that distinguishes Heyes' alternative interpretation from the behavior-rule account. Where the behavior-rule account may be difficult to challenge, Heyes' novelty hypothesis does imply predictions about infants' looking behavior. Since no existing condition rules out all opportunities for novelty, Heyes would predict that all VOE studies should report no difference in looking time. Thus, unlike the behavior-rule account (which can explain why infants would look longer in one condition over another) Heyes' account is challenged by each VOE study that successfully rejects the null hypothesis.

1.5 Conclusions

A considerable body of non-verbal FB tasks has revolutionized the understanding of FB reasoning. However, the sophistication that is attributed to early mindreading (i.e., the degree of continuity between implicit and explicit mindreading) has divided the field. Some interpret the results from these tasks as evidence that FB reasoning comes online before the first birthday (Kovacs, Téglás & Endress, 2010; Luo, 2011). Many proponents of this view disagree about the role that EF play in FB reasoning. Still, others defend the 4-year 'conceptual shift' account by maintaining a skeptical stance toward non-verbal findings altogether and claiming that they can be adequately explained in terms of lower-level associations (Ruffman, 2014; Heyes, 2014).

Given that there is an enduring absence of scientific consensus despite the abundance of evidence amassed from various FB tasks and measures, it is appropriate to consider novel frameworks that can be used to produce alternative, empirically testable theories. One way to do this is by challenging the assumptions made by previous theories. The interpretations of non-verbal data that have been discussed thus far operate on an assumed dichotomy: infants are mindreaders, or they are not. However, in a recent review of the empirical literature, Rakoczy (2017) rejects this assumption, pointing out that "more than behavior-reading can be much less than full-fledged [mindreading]" (p. 698). In the next chapter, an alternative theory of human mindreading is presented, which provides an explanation for those conceptual abilities spanning the gap between statistical learning and full-blown mental state ascription.

CHAPTER 2

A Tale of Two-Systems

2.1. Introduction

Human mindreading has manifested itself as a puzzle from both the cognitive and developmental perspectives. On the one hand, infants appear to demonstrate a procedural or implicit sensitivity to other's mental states. On the other hand, it takes years for this implicit knowledge to inform children's verbal responses when directly questioned. As discussed in the previous chapter, cognitive theories favoring a unified domain-specific system responsible for mindreading in its entirety face empirical challenges. In the following chapter, it is shown how fundamental qualities of mindreading make contradictory demands. One way to resolve these contradictions is by framing mindreading abilities as reflective of multiple cognitive systems.

2.1.1 Efficiency

Apart from understanding how infants pass some FB tasks despite 3-year-olds failing on others, the mindreading community is coupled with another puzzle. One also needs to consider how we can read others' minds in situations which place demands on processing speed. For example, in many competitive team sports, players immediately and repeatedly take the perspective of their opponents in order to outmaneuver them. In Soccer, a formidable defender is one who can tell if the opposing player is about to dribble, say, left with the ball, or merely feign left but ultimately switch and dribble to the right. Players do not have time (or the spare cognitive resources) to represent their opponents full-blown mental states (i.e., as one does in a conversation where mental states are represented and evaluated). It does not behoove the soccer defender to take note that their opponent has a perspective difference about other equally detectable changes in the environment: say, a plane flying in the sky behind him, or the marching band lining up in the distant corner of the pitch. However, it is imperative that the soccer defender is sensitive to more relevant perspective differences (e.g., that unbeknownst to the opponent, another defender is approaching from behind him). In this case, keeping the opponent blind to the teammate's approach may be the key to outmaneuvering him and stealing the ball.

An efficient process is just as reliant on neglecting a host of irrelevant information, as it is on processing relevant information swiftly and accurately. It is not clear if or how a single cognitive system could efficiently represent and reason about abstract concepts, like beliefs, without appealing to operations that are known to be cognitively effortful. This problem is of great importance because historically, mindreading has been considered a sophisticated, flexible and effortful cognitive ability (Harris, 1994; Gopnik & Meltzoff, 1997). There is evidence that we may never fully master it (see Apperly, 2011, pp.86-107), and that even those who can mindread, sometimes fail to use it when it is useful to do so (Keysar, Linn & Barr, 2003). For example, when neurotypical adults were asked to predict where an agent with a FB will look for her violin, those adults who were informed about the real location of the violin were more likely to make errors than those who did not know where the violin had been moved to (an effect known as the 'curse of knowledge'; Birch & Bloom, 2004). Yet findings from AL and VOE tasks suggest that even infants can efficiently anticipate other peoples' behaviors reflecting their mental state. This presents a problem for the 'one-system' theories of human mindreading, and resolving it requires an explanation for how a unified mindreading system could achieve cognitive flexibility in such fast-paced situations.

2.1.2 Flexibility

Under looser temporal constraints, humans can flexibly reason about others' minds. For example, a courtroom could be described as an arena for the careful and measured evaluation of one's past mental state, as it relates to the crime they are alleged to have committed. In this arena, a group of people are presented with various pieces of physical or testimonial evidence which jointly contributes to a representation of a defendant's intentions, informational access, or other content revelatory of their mental processes. Representing crimes in this format is often lengthy and laborious. And so too are explicit judgements about mental states, or overt predictions of behavior on the basis of mental states (Apperly & Butterfill, 2009). Indeed, there is a measurable processing cost to thinking about beliefs, or at least in the preparing of a response to questions which tap belief reasoning (Apperly, Back, Samson & France, 2008).

Back and Apperly (2010) reported evidence that humans take longer to process FB (and TB questions) compared to matched reality questions. Using a method pioneered by Apperly and colleagues (2006), adults were shown a series of images portraying an event sequence where one agent obtains either a false or true belief about the location of an object. Participants were then shown an image depicting reality (the two locations with the object visible inside one) or depicting the agent's belief (a picture of the agent with a thought bubble containing the two locations with the object visible inside one). Participants were asked to press one of two keys, indicating whether the reality picture or the belief picture, was accurate. It is also important to note that in this experiment, participants were not directly instructed to track the agent's belief, merely the location of the object. Results showed that adults were significantly faster to respond following TB event sequences than FB event sequences, and faster to respond to final pictures portraying reality than those portraying the agent's belief. In a second experiment, adults were instructed to track where the agent believed the object was located. Again, adults were faster to respond to TB trials than FB trials. However, adults responded faster to pictures portraying the belief of the agent than to pictures portraying reality. This suggests that, although reasoning about FBs is more effortful than thinking about TBs, current realities or past realities, this FB processing cost can be circumvented when one is primed to attend to other's beliefs. The authors conclude that the way in which humans mindread does not fit the profile of an automatic system (see Bargh, 1994). However, humans can make speeded and spontaneous inferences about beliefs when primed to look out for such information.

2.2 Contradictory Demands: a two-systems rationale

Efficiency and flexibility are both desirable qualities, but rarely are they found in tandem. This is because efficiency and flexibility make contradictory processing demands. Cognitive efficiency is achieved when a system operates on a restricted type of information which can be categorized and operated on using minimal resources. On the other hand, cognitive flexibility is the result of a system which can take variations in context into account. Flexibility is a rare quality to find in modular, domain-specific cognitive systems but rather tends to be achieved through integration with a large network of cognitive domains. Doing so is consequently taxing on one's limited cognitive resources. In light of evidence showing that people sometimes mindread rapidly and unconsciously, and at other times deliberately and at the cost of executive resources, there are two plausible explanations. Either one of these characteristics is false, or this apparent contradiction reflects a divide in the means by which people variously engage in reading other people's minds (Apperly & Butterfill, 2009).

2.2.1 Signature Limits: an empirical testing ground

If humans use two systems to mindread, how could these systems be empirically distinguished? One way to tease apart two processes is by considering their respective processing limits (often referred to as a system's 'signature limits'). A signature limit of a system is a pattern of behavior that the system exhibits which is both defective, given what the system is set up to handle, and peculiar to that system (Carey, 2009). If an efficient system for mindreading has signature limits which are not shared by the flexible system, there should be circumstances where each system's respective outputs produce a dissociation. This requires answering two follow-up questions: what are these circumstances and what specific dissociation would a two-systems theory expect to observe?

One way for an efficient mindreading system to operate is by circumventing the representational demands characteristic of propositional attitudes like full-blown beliefs. Butterfill and Apperly (2013) propose that an efficient mindreading system *tracks* (rather than represents) a limited range of relational attitudes (rather than propositional attitudes) that are associated with beliefs. These relational attitudes operate by keeping track of what an agent has or has not registered. Registrations are belief-like in the sense that they facilitate accurate predictions about an agent's actions in relatively straightforward FB circumstances. However, they are also simplistic relative to full-blown beliefs such that, under more conceptually challenging FB circumstances, an efficient system will make signature calculation errors. Specifically, registrations capture the relation between three features of an environment: an agent, an object and its location. It has been proposed that these relations permit registrations to adequately understand change-of-location FB scenarios, as these circumstances can be understood by merely tracking the agent's relation to an object as being in one location or another. Since in FB change-of-location tasks, an agent registers a target object in box 1, and does not register the object when it is moved to box 2, an efficient mindreading system could predict that an agent will act on her last registration and therefore approach box 1. However, moving an object while another person is absent is not the only way that one may develop a FB. One could also develop a FB about an object that can be mis-identified when its surface features allow it to take on multiple appearances. Such an object would be said to have multiple identities or aspects which it can be known under.

Consider the main character(s) in the 19-century gothic novella *The strange case of Dr. Jekyll and Mr. Hyde* (Stevenson, 1886). In this classic tale, Dr. Jekyll and Mr. Hyde are the same person in the sense that they occupy the same body and yet they are two distinct personalities with polarized behavioral traits. One could say that both Dr. Jekyll and Mr. Hyde are *aspects* of the same entity, which one could call {Dr. Jekyll/Mr. Hyde}. When Gabriel (Jekyll's friend) sees Mr. Hyde, he does not believe that he is also looking at Dr. Jekyll (and vice-versa). Therefore, Gabriel has a FB about the identity of the entity {Dr. Jekyll/Mr. Hyde}. Even though it is true that while seeing Mr. Hyde, Gabriel also has a FB about the location of Dr. Jekyll (as he must falsely believe that Dr. Jekyll must be elsewhere) this belief presupposes that Gabriel connects Dr. Jekyll with Mr. Hyde in a meaningful way. What Gabriel's error necessarily stems from is the *quantity* of persons that Gabriel is seeing. It is the appreciation for the fact that an agent has this numerical FB on the basis of an object's aspectual appearance that Butterfill and Apperly (2013) consider beyond the purview of an efficient mindreading system. This would imply that efficient mindreading should reveal signature limits when attempting to predict the behavior of an agent who has such an identity FB, because registrations can only encode relations between an agent, an object, and its location; they cannot support the *way* an agent sees an object. Understanding that an agent has a FB about an object's identity cannot be 'stripped down' to a sufficiently minimal format: they are inherently propositional.

2.2.2 Location and Identity

The crucial difference between location and identity FBs, is that when an agent develops a FB about an object's identity, the FB stems from an error about the number of objects in the scene. Therefore, it may be more appropriate to refer to it as *numerical* identity. Identity FBs come in two varieties. The first (and the most frequently tested) is a FB based on a numerical *expansion*. An agent laboring under an expansion identity FB would be facing a single entity which can be either (1) accurately referred to under at least two titles (e.g., 'Peter does not know that the firefighter is also Mr. Müller'; Perner, Mauer & Hildenbrand, 2010) or (2) appears to be (at minimum) two distinct and separate entities. Dr. Jekyll and Mr. Hyde represent a case of expansion identity: they embody the same physical material yet present their distinct personas through visually distinguishable aspects.

The second type of FB about an object's numerical identity results from the opposite type of quantity error, which one could call numerical *compression*. An agent laboring under a compression identity FB would be faced with two different entities which share a visual likeness and therefore can be mistaken as one entity (supposing that the agent has never seen both entities at once). In the 1961 film, *'The Parent Trap'* two re-united identical twin sisters capitalize on compression identity FBs by swapping lives in order to force their divorced parents to contact one another. Compression identity FB tasks are absent in the mindreading literature. Therefore, it is one of the goals of the present thesis, to present a compression identity FB task in order to provide further evidence for the view that numerical identity FBs incur signature limits on efficient processing.

2.2.2.1 Testing FBs about Object Identity

Apperly and Butterfill (2009) hypothesized that efficient belief tracking is not qualified to ascribe FBs about an object's identity but can effectively track FBs about an object's location. In order to test this hypothesis, Low and Watts (2013) presented two videos to a group of 3-year-olds, 4-year-olds and adults. The first video was a standard FB-location sequence based on the AL task from Southgate et al. (2007). Three responses were measured: participants' AL, their verbal responses to a direct FB prediction question ("Which box will the woman look in?") and a control question to make sure that critical information could be recalled ("Which box did the monkey put the ball in *first*?").

Next, participants watched a FB-Identity task. Two boxes are shown, which have two openings (covered by streamers): one facing the camera and the other facing the center of the table. Behind the boxes is a short green wall fitted with two openings that are covered by flaps which open outward from the other side of the wall. In the familiarization trials, both boxes are lifted allowing participants to see the objects hidden inside. Each box houses either a red or a blue object. After the boxes are lowered, an agent appears behind the green wall with his head visible above the wall. His head moves as he tracks each toy while they move from one box to the other. Note that at no point does the agent ever see both objects at the same time. Once each matching toy has switched boxes, lights illuminate from the windows at the same time as a tone is heard. After a brief (1750 ms) delay, the agent reaches through one of the windows, into the corresponding box and retrieves the toy inside. In all subsequent familiarization trials, the agent always chooses the same color toy regardless of what type of object it is. This is done to demonstrate that the agent has a preference for objects of a specific color.

After viewing four familiarization trials, participants watch a FB-identity test trial. The boxes are lifted, however, this time there is only one object (a dog-robot toy) under one of the boxes. The boxes are lowered, and the agent appears. The participants watch as the dog-robot crosses from one box to the other. Next, the dog-robot exits the box from the opening facing the camera. Because of where he is positioned, the agent would not be able to see that the dog-robot toy has exited from the front opening. The dog-robot then spins 180° and reveals that it has one side that is colored blue and another side that is colored red. The dog-robot then returns to the box and (with, say, its blue side facing the participant) crosses back to the other box. If the agent has preferred blue objects in all previous trials, then he should reach into the now-vacant box, driven by his FB about the identity of the dog-robot object (i.e., from his vantage point, he had seen a blue dog-robot go into that box). If participants can track such beliefs, then it is expected that they would look toward the empty box during the 1750ms pause following the light and sound signal. Alternatively, a two-systems prediction would maintain that anticipatory eye movements reflect a mindreading process that cannot process beliefs based on how an agent sees the dog-robot (only whether he has seen it or not) and thus, children and adults should look to the full box instead. While this type of belief tracking should overwhelm the efficient system, verbal responses about the agent's subsequent action are not predicted to be informed by an efficient system and therefore should not suffer these same limits. Instead, verbal responses should reveal a developmental trend as children's flexible mindreading system improves. Therefore, a two-systems account would predict that while 3-year-olds should answer incorrectly in either condition, they should show accurate AL in the FB location condition, but not in the FB identity condition. 4-year-olds should show the same pattern of looking as 3-year-olds; however, their verbal responses should significantly outperform 3-year-olds. Finally, because the two-systems account predicts that signature limits are in place throughout the lifespan, adults are predicted to show the same pattern of AL as 3- and 4-year-olds, while their verbal reports should be accurate across conditions.

Low and Watts (2013) found that, in their FB-identity condition, 3- and 4-year-olds erred both in their verbal predictions (13% correct) and their first AL (6% looked to the correct location). However, in the FB-location condition, verbal responses improved as a function of age (3-year-olds: 31% correct; 4-year-olds: 75% correct) while first AL were at ceiling level for both 3-year-olds (94% correct), and 4-year-olds (100% correct). The same pattern of looking found with 4-year-olds was found in adults, who showed below chance first AL in the identity-FB condition (25% correct), but above chance first AL in the location-FB condition (90% correct). This effect was also found when the authors compared total duration of AL to the correct location (within the 1750 ms window) by calculating a differential looking score (DLS) pioneered by Senju, Southgate and Frith (2009). Irrespective of age, participants spent a fractional amount of time looking to the correct window in the identity-FB condition compared to the location-FB condition.

Wang et al. (2015) acknowledged one limitation of the Low and Watts (2013) identity-FB procedure: the dog-robot only reveals his dual-identity *after* he has made his first crossing. This meant that participants would need to retroactively update their ascription of the agent's belief, which could have placed excessive demands on participants' executive processing. If so, anticipatory looks may have reflected the exhaustion of executive resources, rather than a genuine difference in their ability to predict the actions of an agent laboring under an identity FB. Wang et al. (2015) designed modified versions of the Low and Watts paradigm, where the dog-robot reveals its dual-identity to participants before crossing from one box to the other. They replicated the findings from Low and Watts, lending support to the view that efficient mindreading suffers signature limits not applicable to flexible mindreading. Furthermore, Wang et al. (2015) were the first to show evidence of signature limits in efficient belief processing using a non-western sample of children (from the Semai tribe of Peninsular Malaysia) and adults (from Mainland China).

While these results align with the predictions from a two-systems account, some have levied criticism that Low and Watts's FB-identity task may be more challenging for reasons that have little to do with beliefs. When the participant watches the dog-robot toy cross from one box to the other, they must keep in mind that, from the agent's perspective, the dog-robot looks red. The challenges herein are two-fold: first, there is a continuous inhibition cost incurred as participants must inhibit the color that they can see from their perspective before appreciating what color it is from the agent's side. Second, participants need to mentally rotate themselves in order to step into the agent's shoes and imagine what it looks like from the other side. This re-orientation component is not an inherent feature of identity FBs so Low and Watt's results may reflect a difficulty with spatial rotation (which is known to be cognitively challenging: even expert Tetris players prefer to physically operate a button to rotate shapes when planning their placement rather than rotate shapes in their heads; Kirsh & Maglio, 1994). Therefore, it is difficult to say from this study alone whether these results reflect genuine processing limits or extraneous task variables.

Most notably, Low and Watts' findings were further challenged by a recent investigation into the reproducibility of several studies using AL as their dependent measure (Kulke et al., 2018). The original purpose of the study was to explore signature limits, which they began by replicating four AL FB location conditions to later serve as control conditions. These conditions included those from Schneider, Bayliss, Becker, and Dux (2012), Surian and Geraci (2012), Southgate, Senju, and Csibra (2007), Senju, Southgate, White, and Frith (2009), and the FB-Location control condition from Low and Watts (2013). In their first experiment, Kulke and colleagues were only able to fully replicate the effect reported by Low and Watts (i.e., adults looking more to the empty *belief congruent* box instead of the full box). However, upon further analysis, the researchers discovered two confounding variables in the original videos. In the first two familiarization trials, the agent always places the object in the box on the belief congruent side, and this was also the side used as the 'correct' side during the test trials. In addition, immediately prior to the onset of the AL measurement window, when the agent turns back around to face the participant, they turn their head so that the agent faces the belief congruent side as they do so. When these two confounds were removed, the original pattern of looking more toward the belief congruent side, was no longer found. Instead, looking did not differ from chance.

Based on these findings, it is at least as likely that the Low and Watts paradigm may only reflect perceptual side biases in participants looking patterns, rather than signature limits. In defense of signature limits, however, similar results reported by Low and Watts (2013) have been found by other studies using varying ages, populations and paradigms (e.g., Fizke, Butterfill, & Rakoczy, 2013; Low et al., 2014; Mozuraitis, Chambers, & Daneman, 2015; Edwards & Low, 2017; van der Wel, Sebanz & Knoblich, 2014; Fizke et al., 2017).

2.2.3 Evidence to the Contrary: Limitless Infants

There are studies which have reported evidence challenging the claim that there are signature limits on efficient belief reasoning. Scott, Richman, and Baillargeon (2015) have reported evidence of flexible implicit mindreading in 18-month old infants using a complex VOE

FB task. In this task, infants watch a sequence of events wherein one agent (‘the thief’) implants a FB in another agent (‘the owner’). Scott and colleagues argue that if a limited efficient mindreading system is the mindreading system that infants are using, then its limits should be exposed by a sufficiently complex FB task, such as when an agent’s FB is caused by a duplicate object’s deceptive appearance.

Infants were familiarized to the following pattern: the owner displays a preference for toys that rattle (as demonstrated by her placing them into a lidded box close to her). Any toys which do not rattle, she tosses into the bin on the far side of the table. All of this is passively observed by the thief. Rattling or not, the owner always places the toy on the tray in front of the thief before temporarily leaving the room. If the toy is a rattling one, the thief always picks it up and plays with it during the owner’s absence, taking care to always return it to the tray when a knock on the door is heard, signaling the owner’s return. If the toy is not a rattling one, the thief will leave it where it is, thereby conveying the thief’s preference to the participants.

Scott and colleagues set up three experiments which differed in regard to the events in the test phase. In the main experimental condition (the “deception” condition), the test trials follow this narrative: the thief attempts to steal a rattling toy while the owner is away. To do this, she takes the rattling toy from the center tray and stores it in her front pouch pocket. Then she reaches into the bin and quickly replaces the toy with one of the non-rattling toys which had been discarded by the owner during an earlier familiarization trial. The main difference between test trials was whether the thief replaced the rattling toy with either a non-rattling toy of a matching color, or a non-rattling toy of a non-matching color. According to the EMA account, infants possess a psychological reasoning system that is as sophisticated as the one adults possess. Therefore, an EMA account would predict that infants would look longer in the non-matching trials, as the infants should have expected the thief to use a matching toy as a decoy for the one she stole. Results supported their prediction. In their second experiment, they replicated this finding with a new group of infants and the same pattern was found.

The authors interpreted these results as evidence that infants expected the thief to choose a toy which would successfully deceive the owner and were surprised when she did not do so. In order for infants to have this expectation, first they must have reasoned about (1) the thief’s preference of toy, (2) that there must be some consequence if the thief is still holding the toy when the owner returns, and (3) how the available objects could be arranged in order that deception is achieved. This would further imply that infants have the capacity to reason about how an agent with an identity FB will act (i.e., that the owner will mistakenly place the matching toy in the box, not the bin), and that infants possess enough proficiency in their FB reasoning so as to infer that an agent will come to develop an identity FB when faced with a particular complex set of circumstances.

The researchers compared this difference against another condition where the owner brings in a non-rattling toy in the test trial, which the thief replaces with either a non-rattling matching toy or a non-rattling non-matching toy. They predicted that in this silent control condition, infants would not find it surprising if the thief uses a non-matching toy any more than a matching toy, because the thief has already broken protocol by picking up a non-rattling toy for the first time in the procedure. Results supported this prediction: infants showed similar looking in response to both versions of the silent control condition.

In their third experiment, the scenario was adjusted to include the owner returning to the scene and either acting in line with her FB (by storing the matching non-rattling toy in the lidded box with the rattling toys) or acting in a way which suggests she does not have a FB (by tossing the matching non-rattling toy in the bin with the other non-rattling toys). The authors predicted that infants would look longer when the owner discarded the non-rattling toy. In addition, an *alerted* condition was contrived where the owner returns early and sees the thief holding both matching toys in her hands. Despite the owner watching, the thief continues to finish stealing the rattling toy and leaves the non-rattling toy in the center tray. Infants then saw the owner either toss the toy in the bin or store it in the lidded box. The authors predicted that infants should look equally when watching either outcome, since even though the owner was alerted to the thief's plan, she could not know from watching this whether a rattling or non-rattling toy was the one left on the tray. Results were in support of the authors' predictions. In the standard deception conditions with the added outcome trials, infants looked longer when the owner tosses the non-rattling toy in the bin and is not deceived by its appearance. And when the owner returns and watches the thief steal one of the toys (yet does not know which one has been stolen) infants do not show differential looking regardless of where the owner puts the toy.

Prima facie, these findings are in conflict with what the two-systems account would predict. Infants' efficient mindreading should not be equipped for adequate predictions on the basis of how an object is seen by another agent (let alone the causal antecedents which give rise to FBs about how an object is seen). However, upon more careful scrutiny, it is possible that this task fails to meet the criteria for an identity FB task. Identity FBs are essentially numerical. An agent believes that there are multiple different objects, yet there is only one object with multiple aspects; or an agent believes there is one object when there are actually multiple identical objects. The latter FB can only be made when the agent has not already seen that there are multiple identical objects. Is the owner mistaken about the number of objects in the scene? Perhaps not. Here, the owner had already registered each of the identical objects in the familiarization trials beforehand. Therefore, it is possible that infants' efficient system attributed to the owner's a FB but about the location of these two objects, not the identity of the object. By comparison, imagine that one were to ask the agent in the Low and Watts (2013) identity FB task how many dog-robot toys there are in the scene. His response would be: "two: a red one and a blue one" and he would be incorrect. However, if one were to ask the owner from Scott and colleagues (2015) how many green toys there are (supposing that the matching toy is green, in this case), then she would answer: "two: there is a rattling green toy on the center tray, and a non-rattling one in the bin from earlier." Notice that the owner would technically be incorrect because the rattling toy is actually in the thief's pocket, and the non-rattling toy is on the center tray. However, she is not wrong about the number of identical objects that are present in the scene. Therefore, it is possible that a crucial property of identity FBs (i.e., numerosity) was missing from the narrative, and in its absence the owner's FB was processed as a FB about two objects' respective locations.

Scott and colleagues (2015) claim that the two-systems account would make the opposite prediction from what the results show. In the deceived condition, the authors explain that a minimal mindreading account would expect the owner to discard the matching non-rattling toy, because infants would "expect her to register the toy on the tray for what it really *was*, the matching [non-rattling] toy" (p.48). This prediction implies that infants would know *for sure* which object is which (i.e., that infants could not become confused about which object was which, in the scene). Although it is true that a minimal mindreading system would not be able to encode how the *agent* could see an object as something it is not, this does not imply that a minimal mindreading infant could not also become confused about the object in view. Moreover, in a sufficiently complex narrative such as Scott and colleagues' (2015), it is likely that an infant *would* find it challenging to keep two identical objects separate in their mind. Therefore, it is not the case that a two-systems account would predict that infants would always expect an agent to treat an object as what it really is, if the infant could also be unsure about what the object in question really is. That being said, it is true that the pattern of looking found in Scott and colleagues' (2015) study, is in conflict with what a two-systems researcher would predict.

To be sure, the notion that infants can keep track of such a multifarious network of objects, preferences, behaviors and goals is extraordinary, as infants should have trouble interpreting such a complex narrative. However, it remains worthwhile to conduct additional experiments which rule out the possibility that infants may have attributed to the owner something computationally simpler than a numerical identity FB. Another reason to remain skeptical of Scott and colleagues' (2015) interpretation is due to the questionable rationality of the thief's behavior. The authors argue that infants are surprised because, in the non-matching trials, the thief acts irrationally (i.e., the agent could more easily deceive the owner by using a matching toy). A wide range of evidence from visual habituations tasks has shown that infants expect that agents will act in ways that are rational, efficient (i.e., they will obtain their goal through the most direct means) and/or consistent (Csibra, 2008, Csibra et al., 2003; Southgate, Johnson & Csibra, 2008; Scott, Baillargeon & Bian, 2016). According to this body of evidence, infants should find the thief's behavior at least inefficient when the thief chooses to steal a rattling toy by engaging in risky and time-sensitive deception. If the thief remembers that there are matching non-rattling toys in the bin, so too should the thief remember that there are three rattling toys in the lidded box to her left. If she wants a rattling toy, she could save the trouble of looking through the bin for a matching non-rattling toy, and instead simply remove one from the lidded box (for further discussion about the rationality of the thief's actions, see Low et al., 2016).

Similar alternative interpretations can be applied to the findings from Scott and Baillargeon's (2009) two-penguin task (see section 1.4.1). This task also claims to provide evidence that 18-month-olds form appropriate expectations about an agent who has a FB about an object's identity. Infants learn that an agent has the goal of hiding his key. The agent's options are: a penguin toy which is split through the middle (appropriate for hiding a key), or a penguin toy which is whole. When the split penguin is assembled, the two penguins are identical. At the start of each familiarization trial, the split penguin is presented in two-pieces. The agent watches as the split penguin is put back together, and following a light and sound cue, the agent consistently chooses the split penguin, and stores his key. In the test trial, the penguins are presented in a novel way: for the first time, the split penguin is constructed before the agent appears and placed into a transparent container, while the one-piece penguin is placed inside an opaque container. The authors reasoned that if 18-month-old infants were operating on a representational system for implicit mindreading which can take into account the way that the agent will see this scene (i.e., that the penguin in the transparent container is the one-piece penguin and the split penguin should be in two-pieces underneath the opaque container), then they should look longer when the agent chooses the constructed split penguin. This is what they found. However, this experiment may also fall short of *necessarily* tapping into infants' processing of numerical identity FBs. Here again, the agent is well aware of the number of penguins in the scene; what she is mistaken about is where they are. There have also been some concerns about whether the agent could not be seen to have been making a distinction between object types rather than object identities *per se* (see Butterfill & Apperly, 2013, c.f., Scott, Richman & Baillargeon, 2015).

What these results more clearly show is that infants know how an agent's behavior will differ based on having witnessed certain information. However, it may yet be premature to assert that infants were reasoning about identity FBs when forming their expectations about the agent's behavior. In a discussion about non-verbal tasks for measuring FB understanding, Zawidzki (2011) points out that, in Scott and Baillargeon's (2009) penguin study, the agent has already been made aware of the fact that the split penguin can be constructed to look like the 1-piece penguin. When the agent sees, for the first time, that there are suddenly two boxes available, and only a whole penguin is visible, she should believe that the split penguin could be behind either (or neither) of these boxes. In summary, not all misunderstandings which arise as a result of mistaking identical objects, are necessarily identity FBs. If objects of a likeness are seen together, they are treated as a type, and distinguished based on their other qualities. Furthermore, an agent who has been informed about the number of identical objects in the scenario cannot subsequently be said to have developed a FB about numerical identity. Finally, it should be noted that these results from either study have yet to be externally replicated. In other VOE studies, replication efforts have been mixed. For example, while there have been replications of Onishi and Baillargeon's (2005) findings (e.g., He, Bolz & Baillargeon, 2011; Song, Onishi, Baillargeon & Fisher, 2008) attempts to find the same results in independent labs have been largely unsuccessful (e.g., Yott & Poulin-Dubois, 2016; Yott & Poulin-Dubois, 2012; Powell et al., 2017, Dörrenberg, et al., unpublished; reported in an open data set by Kulke & Rakoczy, 2018). Thus, these results should be taken as cautionary until they have been replicated by independent labs.

What this discussion hopes to demonstrate is that the essential criteria demarcating identity FBs have not been fully pinned down, leaving room for alternative interpretations of the available evidence. However, there has been a recent movement in the literature to correct this (Oktay-Gür, Schulz, & Rakoczy, 2018; Fizke et al., 2017). Furthermore, it is worthwhile to treat VOE studies as only capable of signposting competency; they do not reveal the underlying cognitive processes (Haith, 1998). If one wants to uncover the component processes of these abilities, one needs to go beyond eye gaze measures.

2.3 Beyond Eye Gaze

Since studies measuring eye gaze dominate the evidence both for and against signature limits on efficient mindreading, and since interpretations of eye gaze data present numerous opportunities for alternative explanations, there is a demand for corroborating evidence from studies which do not rely on eye gaze. In one recent study, Fizke, et al. (2017) designed an active helping task for use with children (2- and 3-year-olds) which tests their responses to an agent who has a FB which the authors claim *necessarily* results from a numerical discrepancy.

Children were assigned to either an identity FB (or TB) condition, or a location FB (or TB) condition, which was designed to replicate the FB condition from Buttelmann et al, (2009). In the location FB condition, an agent (Susi), leaves her blue ball in one of two boxes, before she exits the scene. In her absence the ball moves from box 1 to box 2. Susi re-enters the scene and tries to open box 1 but fails to do so. Susi then turns to the child and asks, "Can you help me?" In the identity FB condition, children are first introduced to a toy which could be transformed to take the appearance of either a carrot or a bunny (using a zipper, similar to a reversible jacket). Once children are made aware of the toy's transformative abilities, the toy is reset to reveal its bunny aspect. Then the FB sequence begins. Susi says, “I have to leave, I will put the bunny in this box” and she stores it in a box (note, that in the FB identity condition, only one box is used). In her absence, the experimenter removes the bunny, transforms it into a carrot, and leaves it on the other side of the table in full view. Susi then returns and tries to open the box, failing to do so, and then asks the child for help. In the TB control conditions Susi re-enters the scene before the experimenter moves the toy and while watching the toy moved or transformed, and verbally indicates that she understands what has happened.

The 2-systems account would predict that young children can track FBs about location, but not FBs involving numerical identity. Therefore, Fizke and colleagues predicted that: children should open box 1 more often in location TB than in location FB but should not open box 1 more often in identity TB than in identity FB. Results supported this prediction: while children were more likely to open box 1 in the location TB condition rather than location FB condition (12 children vs. 6 children; $p = .037$) they were just as likely to open box 1 in the identity FB or TB conditions (15 children vs. 12 children, $p = .166$).

This study is the first of its kind to compare young children’s helping responses on matched FB location and FB identity tasks. In addition, unlike other FB identity tasks, (e.g., Low & Watts, 2013; Low, et al., 2014) Fizke and colleagues’ (2017) task presents no mental rotation demands and presents an agent who has a FB which arises necessarily from numerical identity, since Susi never sees the transformation of the bunny into a carrot in the FB condition. These results casts doubt on the findings reported by Buttelmann, et al. (2015), who claimed that infants could correctly reason about identity FBs when choosing how best to help the agent.

Recently, this pattern was confirmed when Oktay-Gür, Schultz and Rakoczy (2018) measured 3-6-year-old children on a battery of explicit and implicit FB tasks. Each of these tasks had closely matched versions either requiring the appreciation for an object’s aspectuality or not. On the explicit tasks, children performed significantly above chance and performance was strongly correlated, suggesting that aspectuality does not overwhelm or challenge children’s abilities to consciously consider FBs. However, in a helping task, children’s performance on FB and TB tasks was significantly different depending on whether aspectuality was involved. In non-aspectual tasks, TB and FB conditions differed significantly: children helped the protagonist open the box with the toy only when she had been out of the room when it had moved. However, in closely matched aspectual tasks no significant difference (p = .29) could be found between performance on FB and TB versions. It is also noteworthy that Oktay-Gür and colleagues acknowledge that their dependent measure (the child's spontaneous behavior) allows for irrelevant or otherwise ambiguous responses. For example, the child may appear to choose both boxes simultaneously. In a secondary analysis, ambiguous choices were excluded but the effect was still observed: only in non-aspectual tasks did children successfully give the protagonist the object she was looking for. These findings lend support to the notion that young children and infants operate on a minimal mindreading system which is unequipped to comprehend that an agent may develop an identity FB as a result of an object's aspectuality.

2.4 Evidence from Adults

Studies using adult participants occupy an insular niche in the mindreading literature. The reason for their relative scarcity is that, on most FB tests adults show ceiling performance (Apperly, 2011). Sometimes a group of adults are used in studies measuring infant's non-verbal performance, to test the effect across separate measures (e.g., Kovács, Téglás, & Endress, 2010). In other cases, adults with Autism Spectrum Disorder (ASD) have provided insight into the online processing of FBs. In one FB study testing high-functioning adults with ASD, researchers observed significantly poorer AL, despite ASD adults performing just as well as neurotypical adults on verbal FB questions (Senju, et al., 2009). One interpretation of this finding is that the flexible mindreading system improved over time, but the efficient mindreading system was resistant to change. This interpretation fits with the two-systems view as the efficient system should be encapsulated from other domains of cognition and therefore improvements and deficits to domain-general cognitive systems should not inform the output of the efficient system. Another way to test adults while avoiding ceiling results is by testing the speed and accuracy of responses to different mindreading judgements (e.g., Samson et al., 2010, Qureshi et al. 2010) or by coupling the FB task with a concurrent unrelated task in order to mask the FB manipulation (e.g., Schneider, Not & Dux, 2014) or to test how well their mindreading endures under a high working memory load (e.g., Schneider, Lam, Bayliss, & Dux, 2012).

In the quest to empirically distinguish efficient and flexible mindreading, it is crucial to identify signature limits. To this end, adults are not only an appropriate group to test, but preferred, as researchers can relatively safe to assume that neurotypical adults are mature mindreaders. If they still show signature limits in their non-verbal performance on FB tasks nonetheless, this would provide stronger evidence of distinct processes, compared with studies using young children, where evidence of signature limits may be explained by noting potentially overwhelming extraneous task variables.

Other researchers have used adults to investigate online processing using continuous measures. van der Wel, Sebanz and Knoblich (2014) designed a FB task where adult participants were instructed to point (using the cursor of a computer mouse) to one of two objects, which always ended up behind one of two occluders. An agent could be seen in the bottom corner of the screen and left the scene at critical times during the movement of the objects, which permitted participants to recognize that sometimes the agent had obtained a FB about the location of the objects. Half of the participants were told that they needed to track the agent's belief, and the other half were not informed that the agent was task relevant, thereby allowing the researchers to examine the effect of an agent's belief on either explicit or implicit processing. A sound cue prompted users to move their cursor, but after the cursor had moved upward 50 pixels, the occluders disappeared revealing the real location of the objects. By analyzing the curvature of their mouse trajectory (e.g., the area-under-the-curve) as well as participants' latency to begin moving the mouse, researchers were able to test the hypothesis that others' beliefs are processed automatically, and whether drawing explicit attention to others' beliefs impacts performance.

The authors found that the agent's belief had a significant influence on participants' mouse trajectories, even when they were not told that the agent's belief was relevant. However, only when participants were explicitly informed to keep track of the agent's belief were reaction times slowed by the incongruence between their and the agent's belief. Taken together, these results suggest that adults encode the beliefs of others, even when they have no reason to do so. These results also lend support to the two-systems account, as the slowing of reaction times in the explicit group implies the operation of a controlled and flexible mindreading system. The absence of this effect in the implicit group implies that such controlled operations are exclusive to conscious processing and do not affect efficient mindreading processes. This implies the operation of an online system which processes others' beliefs (or belief-like states) without consciously realizing it and, presumably, without expending a notable amount of cognitive resources. However, this online system does not hold the same jurisdiction over more conscious decisions. Only when beliefs are explicitly attended to do they incur a measurable processing cost.

Not all research has supported the assumption that adults, being fully matured mindreaders, will always deploy their mindreading abilities when it is in their best interest to do so. Keysar et al. (2000) asked adults to take part in a ‘director task’ where the participant is directed by a confederate to move a series of objects around in a grid. The grid is designed so that several of the spaces housing objects are occluded from the director’s side but are visible to the participant. This presents opportunities for the director to reference objects that may be ambiguous without taking into account what the director is able to see. For instance, there was a one-inch candle hidden behind an occluder, as well as a two- and three-inch candle which were not occluded. If the director’s instructions are to “move the small candle to the right” the correct response would be to move the two-inch candle to the right and to leave the one-inch candle where it is. Interestingly, adults often considered the one-inch candle the referent of the request or reached for it initially.

In a follow-up study, Keysar, Linn and Barr (2003), ran the same procedure but instead of having an object simply placed behind an occluder, they instead showed the adult the object, placed it in a bag, and into one of the occluded spaces. In addition, the participant and the director switched places for a practice trial to emphasize to the participant which objects in the grid were not visible from the director’s vantage point. The aim here was to make this game as easy as possible: all participants have to do is ignore the object that is hidden in the bag. The only challenge present was that the object in the bag was a homonym of another object that was mutually visible (e.g., a cassette *tape* hidden in the bag, and a roll of *tape* mutually visible in the grid). Therefore, in the critical instruction: “move the tape one space to the right”, participants simply needed to ignore the bag with the cassette tape and move the roll of tape to the right.

Surprisingly, 71% of adults wrongly moved the bag, at least once during the task. The researchers also captured participants’ eye movements during the task and found that 92% of participants showed at least one fixation to the bag between the time when the director named the object ‘tape’ and the moment when the participant began reaching for an object. In a second experiment, the director was shown a certain object (e.g., a ball) being placed into a bag, but when the director’s view was blocked, the participant observed the ball replaced by the cassette tape. Now, the director is aware of the bag, but has a FB that it contains a ball, meaning that their instruction to “move the tape…” could not possibly be in reference to the cassette tape hidden inside the bag. Still, 24% of adults incorrectly moved the bag, at least once.

Keysar and colleagues’ (2000, 2003) findings are consistent with evidence that adults’ belief reasoning abilities suffers from egocentric biases and cannot be used automatically (Epley, et al., 2004; Birch & Bloom, 2007; Samson et al., 2010). These results support the notion that belief reasoning is cognitively demanding, and casts doubt on the findings from looking time studies using complex narratives with interlocking mental states (e.g., Scott et al., 2015; Scott & Baillargeon, 2009). If adults struggle to interpret beliefs during arguably simple circumstances, it is puzzling that 18-month-olds kept up with the thief’s intention to cause another person to have a specific FB through deceptive switching (Scott et al., 2015).

2.5 Theoretical Challenges to the Two-Systems Account

Although the two-systems account can explain certain puzzling results, other theorists have noted problems in the literature for which the two-systems account does not provide a convincing explanation. One aspect of the two-systems account which has been only briefly mentioned throughout the discussion thus far is *encapsulation*. If one proposes the operation of two distinct modular processes acting in parallel and sometimes outputting different responses or expectations, then at least some of the information available to one process must not be available to the other (Fodor, 1983). The efficient mindreading system is said to achieve such efficiency because it is largely encapsulated from other cognitive processes. Because of this encapsulation, it is also relatively inflexible and therefore subject to signature limits. Carruthers (2016) brings up evidence that calls this proposed encapsulation into question. For example, the finding from Qureshi et al. (2010) show that performing a concurrent EF task disproportionately slowed adults’ ability to resolve perspective inconsistencies, suggesting that efficient mindreading is impacted by EF and therefore, is not encapsulated from EF. Furthermore, Christenson and Michael (2016) add to this line of discussion by providing examples of cognitive systems in other domains that appear to operate both flexibly and efficiently without relying upon informational encapsulation (e.g., language comprehension). Westra (2017) further challenges the two-systems account by noting it may be wrong to suppose that implicit belief tracking is rudimentary and insensitive to context. Analogous to the two-systems’ implicit/explicit distinction, Westra uses the level-1/level-2 distinction in human perspective-taking to show that level-1 perspective-taking (i.e., knowing *what* somebody else is looking at) operates extremely rapidly (e.g., 10-15 ms; Kingstone et al., 2004) but is nevertheless un-encapsulated from long term memory (Weise et al., 2012). However, the two-systems account does not argue that the efficient mindreading system is completely encapsulated from other processes; the difference between the efficient and flexible mindreading systems, is a matter of each system's *degree* of encapsulation.

Carruthers (2016) also questions the underlying logic of the two-systems theory. He argues that even if it is true that infants cannot represent the aspectuality of beliefs, and if it is true that representing the aspectuality of beliefs is necessary in order to say that one has a concept of belief, it does not follow that infants *must* be operating on one system during infancy and develop a more flexible and distinct system later. Instead, Carruthers (2016) argues that it is equally plausible that a single system may become more conceptually sophisticated through incremental learning. This view is supported by research showing that increased exposure to social activities in early childhood can help children pass FB tasks at a younger age (Hughes & Leekam, 2004; Rosnay & Hughes, 2006).

While the competing view: that a one-system account is more parsimonious than a two-system account, seems rational, adopting this view may inherit its own set of empirical and theoretical challenges, and some from the same results which are challenging to the two-systems explanation. For example, if incremental learning explains how a mindreading system develops, it becomes challenging to explain how infants can understand the social situations in Scott et al., (2015). Moreover, the one-system accounts on offer (Baillargeon et al. 2010; Leslie et al. 2004; Carruthers, 2013) may not be so dissimilar to two-systems accounts. In each case, an early developed belief reasoning system is online albeit missing some crucial 'belief selection' mechanism which allows the child to choose between their and another's belief when responding to verbal questions about FBs. This mechanism comes online later and, when it does, it unlocks this ability for a variety of contexts. Since this mechanism seems to come online around the fourth birthday for the majority of children, this seems to fit better with a discontinuous shift rather than a continuous and gradual enrichment of their skills. Furthermore, Carruthers (2017) also states that his one-system account could be described as a two-systems account. With these theoretical characteristics in mind, it is difficult to tell whether it is really the idea of multiple systems which is driving the present disagreement, or if the debate is really between different sets of two-systems accounts.

2.6 Conclusion

A two-systems account of human mindreading seems fruitful in providing an adequate explanation of both the wealth of evidence claiming that infants and non-human animals can mindread (albeit minimally), and the even more robust finding that humans fail standard FB tasks before their fourth birthday. Signature limits on conceptual flexibility implicate non-verbal behaviors when faced with aspectuality demands, while verbal responses to those same conceptual challenges are unaffected.

Although it would be premature to suppose that the mindreading community has come to any form of consensus about the debates discussed in the present chapter, much of the evidence continues to converge in favor of a division of labor between efficient and flexible mindreading. In the following chapter, a new conjecture is introduced: that efficient belief reasoning interfaces with motor processes (Butterfill & Sinigaglia, 2014). The evidence surrounding the use of motor representations to determine the goals or intentions of other people are explored, and (most interestingly) the conditions in which motor representations fail to operate reliably are investigated.

CHAPTER 3

Motor Representations for Action Prediction

3.1 Introduction

This thesis, so far, has revolved around theories of human mindreading with special emphasis on one- vs. two-systems accounts. The present state of the field is one locked in contentious debate about the unity or dis-unity of mindreading abilities, a debate echoed across numerous fields of cognitive science including reasoning (Sloman, 1996), and social cognition (Keren & Schul, 2009). Such arguments make pivotal first steps toward a comprehensive account; however, disagreements of this nature often go unresolved. It has been almost a decade since Apperly and Butterfill proposed a two-systems account of human mindreading (and since Scott and Baillargeon have proposed the EMA), and despite a wealth of investigation, theorists are no closer to reaching a consensus about the degree of unification between cognitive mechanism(s) underpinning mindreading abilities.

This chapter necessarily sidesteps the present lack of consensus, in order to investigate some of the more speculative properties of the efficient mindreading system proposed by Apperly and Butterfill (2009) and supported by recent empirical evidence. An account of human mindreading is incomplete without considering how it can influence actions, and in the following chapter the motor cognition literature is unpacked in the hopes of uncovering if and how processes for action production may interact with cognitively efficient belief reasoning.

3.1.1 A Theoretical Disclaimer

I should pause here to emphasize that much of the work presented in this chapter combines the findings from several distinct methodologies, not all of which necessarily overlap. I want to preface this literature review by noting that there are serious philosophical challenges levied against the position I intend to defend, and a clear empirical test of those challenges has not been proposed (although there have been *some* philosophical musings about the conditions which may be testable; see Butterfill & Sinigaglia, 2014). Where the argument relies on conjecture or the presumption of an untested outcome, I will endeavor to announce their absence, and proceed with cautious speculation.

3.1.2 Unraveling Actions: A selective pressure

My aim is to show that there is evidence suggesting that human mindreading interfaces, at least partially, with motor representations. However, the defining qualities of a ‘motor representation' are difficult to pin down, and the subject of much philosophical debate (Jeannerod, 1997, 2006; Butterfill & Sinigaglia, 2014; Mylopoulos & Pacherie, 2017). In the present thesis, when 'motor representations' are referred to, the view advanced by Butterfill and Sinigaglia (2014) is intended: they are representations of an action outcome which, unlike propositional attitudes, are represented motorically (i.e., represented by the motor system). The theoretical risks and benefits for siding with such a view are unpacked in section 3.9.

It is far from controversial to suppose that the ability to infer the meaning behind one's actions predated alternative means for extracting mental states (e.g., through the use of linguistic or gestural communication). Inferring the goals of others' action by analyzing their ongoing actions would provide early hominins with enormous advantages. For instance, cooperative hunting (necessary for besting large game) would have been impossible without an efficient means of interpreting and predicting the other hunter's actions (Skyrms, 2004; Stiner et al., 2009). Thus, it is plausible that the analysis of body movements provided our ancestors with a rough-and-ready translation of another's goals, at least.

3.2 Mindreading in Motion

Mindreading is inextricably tied up with the coordination of movement. And on its own, the coordination of movement is dependent upon anticipation. If I notice a cup of coffee begin to teeter off the edge of a table, I can anticipate that it will fall, and I may thrust my arm out to steady it before that happens. The fast reaching action carries weight, but my body anticipates this transfer of balance, so it automatically initiates action plans that adjust and correct the kinesthetic disequilibrium. This automatic process ensures that my body resists following my arm and I avoid falling on my face.

The payload for accurate mindreading is the prediction of others' action. If one can accurately predict the actions of an agent, then one has an understanding of the agent's action. So far, the focus has been on cognitive theories of belief reasoning underpinning mindreading. However, ascribing beliefs is not the only way in which an observer may predict the behavior of an agent. For instance, when watching another person grasp a cup of coffee, one can anticipate that their next move will be to bring the cup to their mouth, and in doing so, one need not reason about that person's beliefs. They only need to anticipate the likely goal of the unfolding action.

To keep concepts from overlapping, first a note about terminology.

3.2.1 Goals and Intentions

In the upcoming chapter, there are often references to inferences about *goals*, but these should not be conflated with mental states. Mental states are propositional attitudes, such as intentions and beliefs. Mental states take a sentence-like format and make actions intelligible by way of ascribing to them causal explanations (Apperly, 2011). Goals are not propositional attitudes; they are outcomes to which actions are directed; an *end point* (Ruffman et al., 2012). An observer who has only a concept of ‘goals’ could accurately infer the outcome of an agent’s presently unfolding action. Knowing an agent’s intentions or beliefs, however, would permit an observer to predict how an agent will act after the presently unfolding action completes. For example, suppose that a patron at a bar suddenly stands up, grabs their barstool, lays it down, and begins stomping it to pieces. To an observer, the goal of the patron’s behavior is obvious; his goal is to bring his foot downward and into the stool. However, what the observer cannot know from the patron’s actions is the intentions his actions subserve. The observer has no idea what beliefs are motivating his actions, or even what he will do after each stomp. Perhaps he is simply exercising some pent-up fury and will stop after he completes his current stomp action. Perhaps he intends to destroy the stool in order to fasten a weapon from the splintered pieces. These intentions are not translated by goals alone (insofar as goals are operationalized in the present work).

Do anticipations of an agent’s likely goal relate to anticipations informed by belief tracking? In the following chapter, it is asked whether both processes for generating predictions share a connection with motor representations underpinning action execution. Doing so will require engaging the evidence both for and against an intersection between the system responsible for generating motor representations and the system responsible for the tracking of goals or mental states. Next, it is shown how researchers can experimentally manipulate one’s access to their motor representations. Finally, a prediction is offered: that such manipulations would produce a corresponding effect on efficient mindreading, revealed through patterns of predictive eye gaze and altercentric goal-directed movements. This plan will first require a brief departure into motor cognition.

3.3 Specialized Neurons for Observing and Performing Actions

Traditionally, action coordination is thought of as a top-down serial process starting with the transduction of afferent sensory signals into common neural code, then routing that information to frontal executive areas, resulting in an action plan that directs certain muscle movements toward a certain goal. Although this portrayal may be intuitively alluring, recent neurophysiological and neuropsychological evidence question whether this process is strictly ‘top-down’. For example, it has been shown that the motor cortex (namely, the inferior frontal gyrus, the premotor cortex and the inferior parietal lobule, all of which are crucial to action *execution*) plays a unique role in the understanding of other people’s actions, by producing an internal simulation of the observed action.

A series of experiments in the late 1980s using single-neuron recording in macaque monkeys revealed a collection of specialized neurons in the ventral premotor cortex (PMv; area F5) that discharge when the monkey grabbed a nut and tried to open it, and when the monkey observed an experimenter grab a nut and try to open it (Rizzolatti et al., 1988; Gentilucci et al., 1988). From the point of view of these specialized neurons (dubbed ‘mirror neurons’) there is no difference between performing and observing an action. Because single neuron recording is a risky technique for testing human participants, direct evidence that a mirror neuron system exists in humans has yet to emerge. However, a wealth of studies using various neurophysiological and neuroimaging measurements have converged on the conclusion that the human brain predicts and reacts similarly to action observation and action execution (see Rizzolatti & Craighero, 2004).

Interestingly, mirror neurons discharge when the subject observes actions involving objects which differ in appearance but are similar in the way one would act upon them (Murata et al., 1997; Raos et al., 2006). Unlike neurons tasked with the representation of the strictly visual characteristics of objects (e.g., color, shape, and size), mirror neurons discharge in a manner which suggests that they rely on a more generalized representation of objects as they relate to achieving a particular goal. Instead of firing in response to, say, all objects shaped like a cup, this class of neurons would fire in response to any object which requires the type of grip one would use to hold a cup. These neurons treat someone reaching to grab a cup and someone reaching for an apple as alike, but they would distinguish a reach toward an apple from a reach toward a peanut.

These findings offer two intriguing implications worth mentioning. First, since these neurons fire in response to particular types of movements that are appropriate for specific outcomes, it is possible that mirror neurons are responsible for identifying the goal of an action. This would suggest that some goals are efficiently, if not automatically, extrapolated from observing an action. However, this is not to say that mirror neurons necessarily offer the *only* means by which goals are extrapolated from visual information, nor does it imply that mirror neurons can always determine the goal of a perceived action (especially if the specific action-outcome procedure being observed is unfamiliar to the observer; see Casile, Caggiano, & Ferrari, 2011). The second intriguing implication stems from the fact that, during observation, mirror neurons fire in a way that is analogous to the execution of one's own action. This suggests that the motor system responsible for coordinating action is not merely the recipient of higher-cortical instructions. Instead, the motor system appears to play an equally instrumental role in the perception, analysis and anticipation of other people's goal-directed actions. If this is the case, then it is at least plausible to suspect that people also make use of their motor system when engaging in mindreading, as movement is often revelatory of an agent's mental state. In the following sections, a balanced and empirically supported description of the mechanisms which recruit mirror neurons for action understanding is presented, with special emphasis on one popular yet controversial hypothesis: the direct matching hypothesis.

3.4 The Direct Matching Hypothesis

To what extent does action understanding owe itself to specialized mirror neurons? A few theorists have postulated that mirror neurons perform an 'action mirroring' function, and that this 'action mirroring' *constitutes* action understanding (Rizzolatti, Fogassi, & Gallese, 2001; Gallese et al., 2004; Gallese & Goldman, 1998). But how exactly is 'action understanding' operationalized? Since the present work focuses on minimal mindreading processes, a comparatively minimal definition of action understanding is that it is essentially 'goal ascription'. This view is postulated by Butterfill and Apperly (2016) who define goal ascription as "the process of identifying an outcome to which an observed or anticipated sequence of bodily configurations and joint displacements are directed". By analogy, when watching someone's arm reach toward a series of objects, action mirroring helps deduce the arm's most likely target. It would follow that the reliable success of such an operation would depend on the action mirroring generating a motor representation which captures the observed actor's kinematic actions. One theory asserts that the motor representation generated for understanding another's movement and the motor representation generated for producing one's own future actions, are not just sufficiently similar to each other; they are a *direct* match.

The direct matching hypothesis is accomplished by virtue of a three-step process. Collectively, this process is an analytic process of an actions means, which facilitates an identification of an action's most likely end. First, upon viewing an agent's goal-directed action, an observer's motor system is activated and facilitates an internal motoric imitation of the action. This imitation is accomplished by mapping the observed actions onto the motor repertoire of the observer. Second, by mapping the ongoing action onto the motor repertoire of the observer, the observer's past experiences allow for an identification of the likely goal of the action. By way of simulation, the observer therefore adopts the agent's goal as their own. Third, this goal is projected onto the agent.

The direct matching hypothesis is supported by evidence from two main disciplines: studies of motor activation during action observation, and behavioral studies measuring predictive eye movements while watching an unfolding action. The latter is also informed by studies where certain brain regions are experimentally disrupted through the use of transcranial magnetic stimulation (TMS). The direct matching hypothesis is challenged by evidence demonstrating motor activation following the observation of actions that are not yet incorporated into the observer's motor repertoire, are performed by a non-human actor, or are simply impossible for observers to produce. This order prescribes the structure of the evidence presented in the following chapter.

3.4.1 Evidence from hand-gaze coordination

The direct matching hypothesis has received empirical support from behavioral research comparing gaze-hand coordination during action observation and action execution. In a seminal study, Flanagan and Johansson (2003) measured the gaze-hand coordination of anticipatory eye movements participants made when watching another person completing a simple block stacking task, as well as their gaze-hand coordination in pro-active eye movements: saccades made in anticipation of the participant's own planned actions whilst performing the task. They reported two main findings. First, participants looked in anticipation to the target of the actor's reach rather than track the actor's hand. Second, the coordination between gaze and hand movements when observing someone else perform the action was strikingly similar to the gaze-hand coordination made when participants performed the action themselves. That is, participants tended to direct their eyes to the contact site (where blocks were picked up) and the landing site (where the blocks were placed) prior to the arrival of the index finger of the actor or their own index finger, and the timing of these respective anticipations were highly similar. To explain this continuity between the coordination of hand and eye movements, the authors suggested that, for action production and action observation, the eyes of an actor or an observer are driven by matching motor programs. According to this explanation, when one begins to produce a reaching action, their gaze-hand coordination is determined by a specific motor representation generated by the motor system, which organizes muscle contractions in order to complete the unfolding action. Moreover, when watching others, motor representations are also generated for the purpose of coordinating an internal simulation of an actor's movement.

3.4.2 Evidence from predictive looking in adults

One limitation of Flanagan and Johansson (2003) is that the sequence of events is shown several times and thus, becomes highly predictable to the observer. This leaves open the question of whether motor representations help to drive eye movements to the target of a reach when that target is not already known. In addition, if motor representations are used to help predict other peoples' actions, and if eye movements are driven by motor representations, then one should find that eye movements should make use of certain motor cues (such as a pre-shaped hand reaching toward multiple objects) and thus be able to fixate the most likely target of a reach. Ambrosini, Constantini and Sinigaglia (2011) tested this prediction by showing adults a series of videos where an actor reaches toward two objects (one that would require a whole hand prehension to grip, and one which would require a precision pincer grip using the thumb and index finger). Since the actor's reaching destination varied according to their grip, the outcomes were less predictable than the videos from Flanagan and Johansson (2003). Adults' eye movements were captured while watching the videos, and then scored for 'accuracy' (i.e., the percentage of trials where the correct target was fixated prior to the hand reaching it) and 'arrival time' (i.e., when the eyes first reached the correct target relative to the movement of the hand). Ambrosini and colleagues (2011) found that accuracy was significantly higher when participants watched a pre-shaped hand reaching, compared to control trials where participants watched a hand that was not appropriately pre-shaped (i.e., held in a fist; ‘no-shaped’ trials) move toward one of the two objects. Participants’ gaze also arrived at the correct target faster when watching a pre-shaped hand compared to the ‘no-shaped’ trials. According to the authors, the difference in accuracy and speed of predictive eye movements fits with the predictions of the direct matching hypothesis. When participants watched a pre-shaped hand moving, they efficiently interpreted these motor cues in relation to the unknown outcome of the unfolding action. It is the immediacy of this effect on the observers’ predictive saccades which argues against the alternative explanation, that observers use more sophisticated top-down processing, based on contextual information, in order to make the same prediction.

3.4.3 Evidence from TMS studies

In another study, Constantini et al. (2013) showed adults the same videos used by Ambrosini et al. (2011), and measured adults’ predictive eye movements after undergoing transcranial magnetic stimulation (TMS). TMS selectively suppresses the excitability of targeted cortical brain regions. If other people’s actions are represented motorically in the mind of an observer, then selectively suppressing regions of the motor cortex should degrade or disrupt their use of motor representation for action prediction. This should result in a reduction of the accuracy and speed of an observer’s predictive looking. Since previous neurophysiological and brain imagining research has shown that viewing other people’s hand actions activates the ventral premotor cortex, specifically (PMv; Rizzolatti et al., 2001; Sinigaglia, 2010) Constantini and colleagues (2013) measured gaze arrival times after receiving TMS-induced suppression of the PMv. Results revealed that suppressing the PMv selectively impaired gaze arrival times when viewing a pre-shaped hand. In control trials, gaze arrived earlier in pre-shape trials compared to no-shape trials (replicating the effect found by Ambrosini, Constantini and Sinigaglia, 2011). However, after undergoing suppression of the PMv, gaze arrival times were the same when viewing pre-shaped or no-shaped hand. Evidently, when the PMv is offline, observers do not benefit from the motor cues communicated through a hand’s pre-shaped posture.

These results are supported by another TMS study conducted by Elsner and colleagues (2013). After TMS pulses were directed toward regions of the motor cortex dedicated to the production of hand actions, participants’ predictive gaze was slower relative to baseline.

However, delivering TMS pulses to regions of the motor cortex dedicated to control of leg movements did *not* slow predictive gaze. This finding provides stronger evidence in support of the direct matching hypothesis, as it demonstrates that action mirroring is somatotopic with respect to the motor cortex regions recruited for making online action predictions.

3.4.4 External constraints and prediction delays

TMS studies show that motor representations can be stalled by suppressing the corresponding regions of the motor cortex responsible for executing the specific action being observed. If motor representations are stalled by disruptions to the matching neural motor regions, are they also stalled by applying external restraints to the specific effector used to produce that action? Ambrosini, Sinigaglia and Constantini (2012) tested this hypothesis by measuring pro-active gaze shifts while participants' hands were tied behind their backs (so that they themselves could not perform the same action they were observing). The videos were the same ones used by Ambrosini et al. (2011). Results indicated that having one's hands tied slowed both the initiation of pro-active saccades, as well as the arrival of those saccades to the correct target object. The authors interpreted the delay in saccadic activity as evidence that motor representations are stalled when the observer is not in a position to perform the observed action themselves.

Taken together, these studies suggest that motor representations are generated during action observation, and then recruit effector specific regions of the motor system. In terms of their function, motor representations improve an observer's ability to efficiently interpret configurations of movement and facilitate forward inferences about ongoing actions. These findings generate several testable hypotheses. For instance, if momentarily limiting one's capacity to perform an action is enough to impact adults' predictive gaze, then infants should not be able to predict actions which they lack the dexterity and fine motor skills to perform themselves.

3.5 Evidence from infants

Research on action understanding in infants has provided insight into the development of action prediction abilities and their relation to motor cortex activation. In terms of response variables, the field is presently populated by two classes of infant data: behavioral and electrophysiological. Behavioral responses include patterns of eye movements, such as directional saccadic shifts or visual habituation (i.e., looking time). The electrophysiological measures use electroencephalography (EEG) to record electrical activity in the brain. These two classes of data do not serve interchangeable purposes. Behavioral techniques are typically used to test whether infants can infer the goal of certain actions (or whether the moving object they are observing is perceived as a goal-directed agent), whereas electrophysiological techniques are used to test when infants engage their own motor system while observing certain actions. In the next sections, evidence from behavioral techniques will be discussed. Electrophysiological measures are discussed later in section 3.6.3.

3.5.1 Goal Attribution by infants

One way for investigative strategies to test whether infants interpret action as goal-directed, is by measuring their responses to means-end sequences that either do or do not seem sensible. Studies which do so, make use of the visual habituation paradigm. This technique takes advantage of a natural tendency for infants to spend more time looking at stimuli that are perceived as novel (and resembles the VOE paradigm, in this way). In one study, Gergely and colleagues (1995) used visual habituation to demonstrate that infants treat actions as goal-directed if the ‘agent’ object reaches the goal in the most rational or direct route available. Infants watched a small ball move toward a larger ball by jumping over a large rectangle. When the rectangle was then stationed behind the small ball (and was therefore no longer an obstacle), infants watched the ball either roll directly to the large ball or perform the same jump as before despite having no apparent reason for doing so. Infants habituated less when watching the ball’s irrational jumping approach, than the new direct rolling approach. Gergely and colleagues (1995) interpreted their results as indicative that young infants attribute to actions a quality of ‘goal-directedness’, viz. that one should obtain a goal using the most rational or efficient means available.

Woodward (1998) also used the visual habituation paradigm to investigate infants’ perception of even simpler actions. Her results revealed that 9-month-old infants expected a human hand to approach one of two objects consistently (another hallmark of having attributed a goal-directedness to an object’s movement), but they do not form the same expectations when watching a mechanical arm, or a wooden rod perform the same actions. In a follow-up study, Woodward (1999) showed infants videos where a human hand either grabbed one of two objects, or merely grazed an object consistently with the back of her hand. She found that when the actor had grabbed one object but grabbed a different object in the test trial, infants looked longer compared to trials where the actor grabs the same object that is placed in a new location. By contrast, when the actor merely grazed an object with the back of her hand, and then grazed a different object in the test trial, infants looked for about as long as they did when the actor grazed the same object in a new location. The direct matching hypothesis would claim that these results line up with their predictions. Infants were able to recruit a matching motor representation of a human actor's hand motion, but not a mechanical claw or a wooden rod, which allowed them to predict the outcome of the hand action and respond with surprise when it reached for something it had previously ignored. Further, their motor representation of the grasping motion and the hand grazing motion allowed them to selectively encode 'grasping' as a goal-directed action, whereas the hand grazing motion was not encoded as goal-directed, by way of comparing it to the infants' matching motor representation of the action.

3.5.2 Infant's predictive looking

Much like AL for testing FB understanding discussed in Chapter 1, infants direct their eyes in anticipation of future actions or to a predictable outcome of an ongoing action. Falk-Ytter, Gredebäck and von Hoften (2006) tested the predictive gaze shifts of 6-month-olds, 12-month-olds and adults when watching two predictable and straightforward videos. The videos consisted of either a human actor moving three objects into a box or the objects flying on their own into the box while the human actor watched. Adults and 12-month-old infants shifted their gaze to the box before the arrival of the objects, but only when the objects were moved by a human hand. 6-month-old infants did not make predictive looks even when the actor moved the blocks. Instead their eyes arrived at the goal roughly 200ms after the blocks were placed there.

Why do older infants and adults anticipate the outcome of manually displaced blocks, and not self-propelled blocks? According to the direct matching hypothesis, the absence of human motion in the self-propelled video prevents the observer from being able to engage in action mirroring, thereby hindering participants from predicting the destination of the blocks. However, there are plausible alternative interpretations. It may be the case that participants did not make the same predictions because the motion of self-propelled objects are not subject to the biomechanical constraints of a human arm. An observer may have reasoned that flying blocks can, by virtue of already disobeying the laws of physics, go wherever they please. Still, it is unusual that their eyes did not fixate toward the only likely target available, since the videos used in the study present no competing destinations for the blocks to enter (the scene displays a table with a single box on one side, the blocks on the other side, and a human seated in the background).

3.5.2.1 Anticipating an action relies on the observer having experience performing that action

Another hypothesis generated by the direct matching view, is that one's past experience performing an observed action facilitates their ability to anticipate the outcome of that action. Previously, it was mentioned that Falck-Ytter and colleagues (2006) did not find evidence that 6-month-old infants anticipated the outcome of the action even when it was performed by a human actor. Accordingly, the direct matching hypothesis would explain that 6-month-olds failed to anticipate the actor's movements because infants at this age lack the requisite dexterity and experience displacing blocks into containers. To lack the proper motor experience is to lack the motor representation needed for action mirroring. Consequently, infants' track, rather than anticipate, the action's outcome.

The idea that one's ability to accurately predict actions is related to one's motor skills is supported by a variety of studies on adults with specific motor expertise. For example, professional basketball players are better at predicting the success of free throw shots, compared to others with comparable visual experience (e.g., coaches and sports journalists; Aglioti et al., 2008). Neuroimaging research showing increased motor cortex activation in expert dancers when watching others dance, compared to a control group of non-dancers (Calvo-Merino et al., 2005; Cross, Hamilton & Grafton, 2006), and similar increases in motor activation were found when trained pianists observe piano playing, compared to non-musicians (Haslinger et al., 2005). Consistent with this view, it has also been shown across a variety of contexts that learning is better facilitated when people play an active role in performing a task, rather than passively observing others perform the task (Yu & Smith, 2012).

How do the limits of one's own motor repertoire contribute to action understanding? A group of researchers have investigated this question directly, by training young infants to perform a certain action, and then measuring their visual habituation to that same action performed by an actor. Sommerville, Woodward and Needham (2005) fitted a group of 3-month-olds with mittens covered by Velcro fabric. Two objects (which would adhere to the mittens upon contact) were placed in front of the infants so that infants could discover for themselves that the mittens adhere to the objects. When these infants watched a visual habituation phase (where a hand, wearing the same mitten, extended and grabbed one of two objects), they dishabituated when the hand contacted a new object from the earlier trials (indicative of having attributed goal directedness). However, another group of infants who were not given the experience using the mittens for themselves, looked equally when the hand approached either object. Furthermore, infants' habituation response was correlated with the extent with which they actively engaged with the toy during the training phase (i.e., the amount of time infants spent looking toward the toy while making contact with it).

In a follow-up study, Sommerville, Hildenbrande and Crane (2008) provided 10-month-olds with an opportunity to learn a novel action themselves (i.e., using a cane to retrieve an out-of-reach toy) or by observing someone else perform the action. Infants took part in an adapted visual habituation paradigm using a cane instead of a mitten and found that infants who received active training with the cane looked significantly longer when the actor used a cane to retrieve a new goal object, whilst infants who merely observed the action, looked equally long following all trials. These two studies are consistent with the direct matching hypothesis. Infants build motor representations from experience producing an action, not merely watching others. Once learned, that motor representation can be immediately repurposed to understand similar actions performed by others.

Ambrosini et al. (2013) used a cross sectional design (6-, 8-, and 10-month-old infants) to test the prediction that individual differences in motor abilities influence how efficiently and accurately they make use of fine postural motor cues to anticipate the target of a particular reach. Participants' motor abilities were assessed by a grasping task: infants were presented with multiple objects of varying size and shape, and the number of fingers used when manipulating each object was recorded (see Butterworth, Verweij, & Hopkins, 1997). Infants then watched the same videos from Ambrosini, Constantini and Sinigaglia (2011), and eye gaze was assessed on accuracy and arrival times to the target. Infants' eye gaze revealed a significant shape (pre-shaped hand vs. no-shaped hand) by target (small or large) by age interaction, which was unpacked to show that 6-month-olds *could* predict the target of a pre-shaped whole hand prehension but could not predict the target of a pre-shaped precision grip, while 8- and 10-month-olds took advantage of both types of pre-shaped hands to predict the appropriate target of the reach. Data from a control group with adult participants confirmed that adults fixated the target of the actor's reach significantly earlier than infants but showed the same effect of target and shape. Concerning individual differences in motor ability and the relationship to gaze arrival time, it was found that infants' grasping ability predicted the degree with which they took advantage of the actor's pre-shaped hand to determine the goal of the reach. And this correlation was observed when age was partialled out. Precision grasping ability was a stronger predictor than whole hand grasping ability, suggesting that possessing finer grained motor skills allows one to more effectively take advantage of kinematic motor cues.

Similar findings have also been reported using 12-month-old infants who watched the videos from Falck-Ytter et al. (2006) after having been granted the opportunity to perform 'containment' actions themselves (i.e., to place a series of balls into a container). Like the results from Ambrosini et al. (2013), a strong correlation was found between the number of balls they placed into the container and the speed of their gaze arrival times when observing others perform the same action (see also Gredebäck & Melinder, 2010; Gredebäck & Kochukhova, 2010).

3.5.3 Direct evidence of direct matching

One problem with the studies mentioned above is that they do not constitute *direct* evidence that motor representations for action anticipation and action production are the same representation. Cannon and Woodward (2008) designed a more specific test to clarify the degree to which representations driving action anticipation match the representations driving action production. Three groups of adults watched the same stimuli from Falk-Ytter et al. (2006). One group simply watched the video, while the other two groups performed concurrent tasks of either working memory or finger tapping. The researchers reasoned that if action anticipation and action production employ the same motor representations, then only the concurrent finger tapping task (i.e., a simultaneous action driven by a *different* motor representation implicating the same effector) should interfere with gaze arrival times to the target. However, working memory tasks are not motor tasks, so gaze arrival times should not be impacted by the concurrent working memory task.

Results supported their prediction: participants who performed a concurrent working memory task fixated the target as quickly as those who simply watched the videos, whereas those who performed the finger tapping task were dramatically slower to fixate the target of the actor's movement. Since participants demonstrated high accuracy on either of the concurrent tasks (each averaging over 95% accuracy), it is unlikely that participants' gaze was impacted by disproportionately challenging concurrent tasks. The authors concluded that the delay of gaze arrival times during the finger tapping task suggests that online motor movements interfered with action prediction. One reason for this interference is that action production and action anticipation recruit the same motor representations.

Taken together, these data clearly indicate a role for the motor system in predicting the goals of actions. Given the previous operationalization of action understanding as, essentially, goal ascription, these data argue that the motor system is crucial for action understanding. The clearest explanation for the connection between these two processes (i.e., action production and action prediction) is that observers' recruit motor representations that are a direct match of that which is being executed in the brain of the actor. Crucial though this process may be, direct matching is not the only means by which actions can be interpreted, as it has also been shown that infants expect morphologically non-human shapes to act in a goal-directed manner on the basis of certain principles of rationality. However, direct matching should allow for an observer to interpret motor cues that other means of action understanding could not (e.g., which object a reach is directed toward on the basis of its pre-shaped hand posture). However, there is much evidence that challenges the direct matching view, and many theoretical alternatives which claim to better explain the data without appealing to extraordinary automatic transferences between motor systems. In the following section, evidence is presented which challenges the direct matching account. After, a selection of noteworthy theoretical alternatives is unpacked.

3.6 Empirical Challenges to the Direct Matching Hypothesis

Next, the focus turns to evidence which has cast doubt on the direct matching hypothesis. In summary, these data are born of two families of research: (1) evidence demonstrating accurate anticipation by participants who do not possess the effector(s) necessary to produce the observed actions themselves, or (2) evidence from participants whose motor cortex is activated when watching object-direct actions which could not be accommodated by the observer's own motor representation. The second section will be divided into evidence from infants and evidence from adults.

3.6.1 Action understanding without action experience: Evidence from upper limb aplasia

If motor representations for action observation are reliant upon the availability of the observer's motor cortex and corresponding effectors, then it would follow that a person born without arms would be delayed in producing looks to the goal of other's reaching actions. Gazzola and colleagues (2007) tested two participants with upper limb aplasia, who watched a series of manual actions, while their blood-oxygen-level dependent (BOLD) signals were recorded using fMRI. These participants' also manipulated objects using their lips or their feet. A control group of typically developing adults was tested and were also asked to manipulate objects with their lips, feet and hands. The aplastics' motor cortex was activated while watching hand motions. Specifically, the fronto-parietal-temporal areas (part of the putative mirror neuron system) were activated during action observation and action execution in both aplastic participants (and typically developed adults). However, most intriguing was the finding that when observing hand actions, aplastics recruited areas of the motor system dedicated to the effector *they* would use to manipulate the object: regions associated with movement of the foot and mouth.

These findings support an alternative explanation of action understanding. While the direct matching view places an emphasis on the 'means by which an action is performed (i.e., the kinematics of the action involved in an object's displacement), as necessary for determining the goals of the action, it is possible that this process operates in the reverse direction. That is, the mirror neuron system in humans may infer the most likely goal of an action, and the 'matching' that occurs in the motor system of the passive observer may be better understood as 'goal matching'. This would fit with evidence from macaques, showing that the majority of mirror neurons are not 'effector specific' but rather 'goal specific' (Gallese et al., 1996). Broadly congruent mirror neurons respond when watching an actor accomplish the same goal regardless of which effector is used. A broadly congruent 'grasping' mirror neuron would respond to an actor picking up a pencil with either his hands, feet, or mouth, but would not respond to the actor merely placing his hand on the pencil.

In section 3.4.4., the speed of adults' predictive gaze was stalled by relatively ordinary constraints. When an observer's wrists are bound, or when simply tapping their fingers on a desk, their action prediction abilities are compromised. If the recruitment of the motor system for action understanding is interrupted by such quotidian disturbances, then surely those who were born without hands would be at a severe disadvantage when predicting the outcome of reaching motions. Vannuscorps and Caramazza (2016) showed the same videos from Ambrosini et al. (2011) to an adult participant with bilateral upper limb dysplasia. This participant (hereafter referred to as 'D.C.') was born without an upper left limb and a shortened upper right limb. D.C. cannot use his right hand to grab or manipulate objects, so he would not be able to produce a precision grip or a whole hand prehension demonstrated in the videos. The direct matching hypothesis would predict that D.C. cannot make use of motor representations to anticipate the outcome of the reaching actions. Therefore, D.C.'s gaze arrival time should be both slower to anticipate the target of the actor's reach relative to a control group, and arrival times should not improve when viewing a pre-shaped hand. However, Vannuscorps and Caramazza (2016) found the opposite pattern: D.C.'s anticipations were more accurate (82% accuracy) and quicker (eyes reached the target 291 ms before the hand) than the control group (average accuracy = 48%; average anticipation latency = 179 ms before the hand). Moreover, the difference between D.C.'s anticipation of shaped and no-shaped hands was greater than typically developed adults, suggesting that D.C. more effectively capitalized off of the kinematics of the hand posture in determining the outcome of the action.

Taken together, these studies constitute strong evidence against the direct matching hypothesis. However, in refuting the direct matching view, Vannuscorps and Caramazza's (2016) data is curious. If D.C. did *as well* as control participants, this would suggest that the direct matching view was wrong, and that the same mechanisms are in play with or without the observer's corresponding effector. The fact that D.C. significantly outperforms control participants on every measure of action prediction suggests that D.C. may not be using the same neural strategies to anticipate the actions of other's arm movements.

However, this begs the question: why should D.C.'s alternative anticipatory strategy surpass controls? This may not be so surprising. A plausible analogy to D.C.'s condition, are those who suffer blindness early in life. Much evidence shows that blind participants outperform control participants on tactile measures (e.g., determining the orientation of a grate using one's fingers; Wong et al., 2011; Sterr et al., 1998), auditory discrimination (e.g., pitch discrimination, Wan et al., 2010; vowel-consonant discrimination, Hugdahl et al., 2004; echolocation, Teng et al., 2012), olfaction (e.g., odor identification, Cuevas et al., 2009; Beaulieu-Lefebvre et al., 2011) and taste detection (Gagnon et al., 2013). Therefore, it may be the case that D.C.'s brain underwent significant neural re-organization in order to make up for his loss of upper limbs. In developing supplemental coordination skills in their feet, it is possible that connections were strengthened through a more effortful skill attribution process, and that once alternative strategies were integrated, superior action prediction abilities emerged. Further investigation is required in order to determine whether aplastic subjects' action prediction skills provide insight into typically developed action prediction abilities.

3.6.2 Replicating Ambrosini, Constantini and Sinigaglia (2011): Internal and external

What is more troubling to the direct matching hypothesis than the surprising competence from aplastics, is how poorly control participants anticipated the target of the hand reaching compared to the original findings. When these videos were first used by Ambrosini, Constantini and Sinigaglia (2011), adults anticipated the correct target 73.2% of the time overall (82% correct when the hand was shaped; 65% when the hand was not shaped), and fixations to the target were made earlier than a shaped hand (-143 ms) but arrived at the same time as a non-shaped hand (+10 ms). By contrast, Vannuscorps and Caramazza (2016) found that their control participants fixated the correct target only 48% of the time overall (51.75% when that hand was shaped, and 44.19% when it was not) and arrived earlier than both a pre-shaped or non-shaped hand (-186 ms and -171 ms, respectively).

However, Vannuscorps and Caramazza (2016) were not the only investigations which tested adults using the videos from Ambrosini, Constantini and Sinigaglia (2011). These videos have also been used as the baseline condition for follow-up studies by the authors of the original study (e.g., Ambrosini, Sinigaglia & Constantini, 2012; Constantini, Ambrosini & Sinigaglia, 2012; Ambrosini et al., 2013; Constantini, Ambrosini & Sinigaglia, 2013; Ambrosini, Pezzulo & Constantini, 2015). Have these internal replication conditions reflected the original study more, or the control participants of Vannuscorps and Caramazza (2016)? The 'unconstrained' participants from Ambrosini, Sinigaglia and Constantini (2012), showed high accuracy (pre-shaped hand = 87.8%; No-shaped hand = 62.2%) and on average, gaze arrived 123 ms earlier than the actor's reaching hand to the correct target. Ambrosini et al. (2013) only analyzed looks to the correct target and, thus only reported the latency of gaze to leave the hand and arrive at the target. They found that adults' overall gaze arrival time targeted the correct object earlier than had been previously reported (590 ms before the actor's hand reached the target). Even 10-month-old infants showed surprisingly early fixations (246 ms before the actor's hand reached the target). Finally, Constantini, et al. (2013) tested adults in three "baseline" conditions using Ambrosini, Constantini and Sinigaglia's (2011) videos. They also did not report accuracy scores, since they analyzed only those trials where participants displayed target-directed gaze behavior. However, they did indicate that 15.6% of trials were excluded on the basis of the participant failing to produce a saccade to the correct target before the end of the movement phase. In light of this information, one could say that adults' accuracy was 84.4%, and average gaze arrival times preceded the actor's hand by 125 ms.

The main takeaway from this collection of data is that there is a demand for additional replication attempts of Ambrosini and colleagues' (2011) findings, from independent labs. How should the present state of this research be interpreted? The short answer is that one cannot know whether humans efficiently take advantage of dynamic hand posture to predict likely action outcomes. On the one hand, the Ambrosini group have found generally the same pattern of results in their own follow-up studies (i.e., earlier gaze arrival times on pre-shaped trials relative to no-shape trials). On the other hand, the only known external replication attempt has shown that adults do not significantly profit from the pre-shaping of a reaching hand in anticipating its target, and when watching hand actions, adults' odds of anticipating correctly are at chance level. Furthermore, even the internal replications are problematic as the speed of gaze is becoming increasingly pro-active in each subsequent analysis. Moreover, in a recent investigation, Low and colleagues have also failed to replicate Ambrosini, Constantini and Sinigaglia's (2011) findings (Low, personal communication, 2018). External replication attempts are generally accepted as stronger evidence than internal replications. There is good reason for this (for a famous example of an effect which could not be externally reproduced despite the original paper providing 8 internal replications, see Bem, 2011 and then Ritchie, Wiseman & French, 2012 for the failures to replicate externally). Overall, the jury is still out on this issue until additional replication attempts emerge or innovative new measures, which tap into this phenomenon through alternative means, are tested.

3.6.3 Evidence from the electrophysiological activation patterns of infant and adult observers

Analyzing the saccadic movements of the eyes is not the only available methodology for investigating action understanding in human infants. Southgate et al (2009) reasoned that, if observers use their own motor system to anticipate the goals of others' actions, then the activation of their motor cortex should be detected by using EEG. 9-month-old infants were fitted with an EEG cap, and data were collected during both a reaching phase (where infants were repeatedly permitted to touch toys that were presented by a mechanical arm) and an observation phase (where infants watched as a human arm emerged from behind a curtain, grabbed an object, and removed it from the scene). Results revealed that infants' sensorimotor alpha rhythm was similarly attenuated in both phases, when either the infant's hand or the actor's hand began to reach. Attenuation of the sensorimotor alpha wave in action observation and action execution has also been found using adult participants (Kilner et al., 2004; Cross et al., 2013). Interestingly, Southgate and colleagues (2009) also found a significant effect of time during the observation phase, as the first three observed trials were associated with an increase in alpha wave activity and the following three observed trials were associated with a decrease in alpha wave activity. This suggests that infants were only able to anticipate the outcome of the reaching action after observing multiple initial trials.

This evidence does not speak in favor of a direct matching hypothesis. Despite the fact that direct matching would predict that motor activation should be found during both action observation and action execution, it would not predict that infants would be required to observe multiple trials before ascertaining the function of the reach. Instead, the visual kinematics of the moving hand should suffice, and matching should not be improved by multiple iterations.

More recent evidence has extended these findings to ask whether the motor activation of observers is somatotopically organized. de Klerk, Johnson and Southgate (2015) measured sensorimotor alpha suppression of both infants and adults during action execution and observation of reaching movements and kicking movements. The sensorimotor cortex is organized somatotopically, such that electrodes placed at a particular point along the cortex will selectively respond to movements of the hand (i.e., C3 and C4), while electrodes placed at other points respond to movements of the leg (see Pfurtscheller et al., 1997). While their sensorimotor alpha suppression was recorded, 12-month-olds infants were offered a toy that was just within their reach, prompting them to attempt to grab it (Southgate et al., 2009), and later had one of their ankles tied by a ribbon connected to a mobile thereby allowing them to move the mobile by kicking their leg. Results revealed that adults and infants showed highly somatotopic differential activation when reaching or kicking. When observing these two actions, adults' sensorimotor cortex again showed highly somatotopic activation. However, when infants watched these actions, activation patterns did not significantly differ. This is surprising since an earlier investigation (where the action was performed live in front of the infant instead of on screen) found that 14-month-old infants' sensorimotor cortex activation was somatotopic when watching hand or foot actions (Saby et al., 2013).

In another study, Southgate and Begus (2013) tested the action anticipations of 9-month-old infants by measuring the activation of their sensorimotor cortex. Infants were shown either a human hand reaching toward one of two target objects, or a mechanical arm, or a self-propelled object. In either case, successive trials displayed a preference, as the hand consistently approached one object. The authors reasoned that, if motor representations are only generated in circumstances where they can be mapped onto the observer's own motor repertoire, then infants should show motor cortex activation only when watching the human hand. However, results revealed changes in the sensorimotor alpha amplitude across conditions.

Southgate and Begus (2013) explain this data as evidence that the motor system's role in action prediction is not driven by corresponding motor representations, but by a more general and context dependent process for identifying goals. This alternative account is supported by evidence showing that adults' motor system is activated when watching actions that they could not produce (Cross et al., 2011; Schubotz, 2007) and by reports that adults' motor activation is significantly less modulated when they watch an actor pantomime actions (i.e., produce actions that are not directed toward an object; Muthukumaraswamy, Johnson & McNair, 2004).

Southgate and colleagues (2014) asked whether motor representations are *also* involved in goal representation. The researchers measured the suppression of the cortical regions associated with goal representation (i.e., the anterior parietal cortex; Ramsey & Hamilton, 2010). The researchers employed a repetition suppression (RS) design: cortical activation has been shown to attenuate as the same stimuli are repeated, and this attenuation is released when novel stimuli is shown (Grill-Spector et al., 2006). 9-month-old infants were shown an animation where a triangle maneuvers toward one of two objects. In the first two trials the triangle makes a consistent choice to approach one object, but on the third trial the triangle instead approaches the novel object. The researchers observed a release of suppression in response to the triangle traveling toward the new goal object, compared to when the triangle moved toward the same object as before. Note that the side where the new or old goal object was positioned was counterbalanced, so this effect cannot be attributed to lower-level differences in the stimuli. This evidence joins a wide collection of research showing that infants can identify the goals of non-human objects (e.g., Biro & Leslie, 2007; Csibra, 2008; Gergely et al., 1995; Hernik & Southgate, 2012; Luo and Baillargeon, 2005; Southgate & Csibra, 2009). Southgate and colleagues (2014) concluded that such evidence challenges the view that goal identification is afforded by directly matched motor representations. Instead, the motor system appears to be involved in identifying the goal of both human (i.e., directly match-able) *and* non-human goal-directed actions.

The evidence described above, showing that infants understand the goals of non-human objects, stands in stark contrast to evidence showing that infants' action prediction corresponds with their motor ability (Cannon et al., 2012; Falck-Ytter et al., 2006; Sommerville et al., 2005; Sommerville et al., 2008; Woodward, 1998; Ambrosini, et al., 2013). How could a direct matching view explain this discrepancy? One possibility is that there are multimodal means by which observers determine the goals of an action. Actions which *can* map onto motor representations are processed by a pathway exclusively designated for human motor simulation, while the actions of motorically un-mappable entities (or of human actions which the observer cannot themselves produce) are processed by virtue of non-motor routes to action understanding (Rizzolatti & Sinigaglia, 2010). According to this view, it would be incorrect to discount the possibility that directly matched motor representations remain an optimal means of predicting and understanding *specifically* human action. However, this view fails to explain much of the data from infant studies. For example, Hernik and Southgate (2012) show that infants (9-month-olds) consider an action to be goal directed if the action is selective (i.e., the agent is presented with a choice between two objects) and efficient. That infants can detect whether the actions of non-human agents are goal directed on the basis of concepts such as efficiency and selectivity is problematic for a direct matching view, as it implies that goal ascription can be better achieved without direct matching. This evidence is in conflict with the direct matching view's explanation that motor routes offer a “richer” understanding of action than non-motor routes (Gallese et al., 2004; Rizzolatti & Sinigaglia, 2010). In the following section, some of the leading theoretical alternative accounts to the direct matching hypothesis are presented.

3.7 Theoretical perspectives which argue against direct matching and motor resonance

The idea that motor representations play some role in action understanding has been accepted by many action theorists. However, the specific reason that the motor system is implicated in action prediction, how they translate visual input to motoric information, and when humans begin to make use of them during ontogeny, divides the field into numerous camps. The direct matching hypothesis is one of those camps, and it has earned an equal share of support and criticism. A comprehensive review of every plausible account is beyond the scope of this thesis. However, in the following section I will attempt to provide a sketch of some noteworthy theoretical challenges to the direct matching hypothesis.

Southgate (2013) reviews much of the literature described in this chapter and argues that the direct matching view is by no means unequivocally supported by infant data. In place of direct matching, Southgate posits that the data favor an alternative account wherein actions are interpreted as one of many available cues (not all of which are ‘motoric’) which allow the observer to attribute a goal to the actor. One of the available cues that can be used is action familiarity, and it is for this reason that there is noteworthy similarity between the motor activation of an observer of an action and someone performing that action. However, this similarity can be better explained by appealing to the familiarity alone, rather than to a systematic matching mechanism. Once the goal is determined, Southgate suggests that the motor system may be recruited for predicting the means by which that goal will be attained. It is this directionality that separates Southgate’s view, from the direct matching hypothesis: direct matching suggests that means are matched, so that the ends can be predicted, whereas Southgate proposes that ends are predicted and then means are motorically simulated. This account bears a resemblance to Biro and Leslie (2007) who propose what they call a ‘cue-based bootstrapping model’. Here, infants begin with an innate concept of ‘goal-directedness’ which is associated with a limited range of ‘cues’ such as self-propelledness and the efficiency of the means by which the goal is attained. As infants develop, the range of potential cues broadens, and more specific types of cues (like those related to the posture of a reaching hand) are integrated. Thus, Biro and Leslie (2007) propose an *early modular gradual expansion.*

Other theorists have homed in on the ‘directness’ of the direct matching view, as its primary shortcoming. Csibra (2007) challenges the notion that motor representations for action production and motor representations for action observation could possibly be a direct match of each other. In his view, differences in the origin of each representation means that observed information may, at best, be transformed into a convincing homologue of a corresponding motor representations; but this will always be an interpretation, never a clone, of the motor representation produced by the actor’s motor cortex. In place of action mirroring, Csibra postulates ‘action reconstruction’. Action reconstruction does not suggest that what is reconstructed is imitative, but rather *emulative* of the observed action. The main difference is that when one emulates another person’s behavior, they start by acknowledging the goal of their action and choose how to best go about accomplishing that goal themselves. Akin to Southgate’s view, Csibra suggests that motor representation does not provide the means for goal ascription but follows *after* goal ascription. The identification of a goal, in this model, is interpreted outside of the motor system.

What should one make of these two classes of debate? First, it should be noted that there is at least one aspect of Csibra’s logic which is irrefutable: on a precise enough scale, all imitation is actually emulation to the degree that one can never co-ordinate their body movements in exactly the same way as someone else. Likewise, any matching visual and motor representation are at best, relatively congruent rather than perfect copies. However, adjusting a direct matching hypothesis into a ‘convincingly matching hypothesis’ does not necessarily imply that goals would be attributed outside of the motor system. It is at least as plausible that parallel processing may be at work here. For example, an observer watching someone use a precision grip to reach out toward a peanut or an apple, can make two predictions roughly simultaneously. First, one may direct one’s gaze toward the peanut rather than the apple, in anticipation of an immediate goal based on an interpretation of the perceptually available kinematics of the precision grip. However, one may also predict that this person has the intention of eating a peanut. Of these two perceivable outcomes, only the latter is necessary for emulation, because only the outcome needs to be achieved; the kinematics are permitted to vary. However, possessing knowledge of the peanut-grabber’s intentions would not lead one to predict that the coordination between an observer's gaze and the peanut-grabber's hand movement, would match the coordination between the participant's gaze and their own hand movements when they themselves reach for a peanut (Flanagan & Johannson, 2003).

Jacob and Jeannerod (2005) provide another challenge homing in on the limitations of a direct matching account. In a nutshell, Jacob and Jeannerod point out that the direct matching hypothesis is insufficient to translate another's intentions to an observer. The authors first acknowledge that 'intention' is an umbrella term, which deserves unpacking. Recall that earlier in the chapter, 'intentions' were aligned with mental states, whereas 'goals' are merely outcomes to which actions are directed (see section 3.2.1). To use Jacob and Jeannerod's own terminology, their central position is that a direct matching account could allow an observer to understand another person's *motor intention* (i.e., their goal) but direct matching could not facilitate the observer's appreciation of one's prior intention, social intention, or communicative intention (all of which are mental states, and more remote relative to motor intentions). For internal consistency, I will continue to use 'goal' in place of motor intention, and 'intention' in place of prior, social or communicative intentions, as dividing up these levels of 'intention' is unnecessary to the present argument.

To illustrate Jacob and Jeannerod's argument, let's start by adapting an analogy of theirs. Consider the action of flipping a switch. No matter what the switch causes (it could operate a light in a room, or it could operate a life-support system in a hospital), the kinematic configurations involved in operating both switches are the same. When the kinematics are identical, direct matching alone could not differentiate between the actions which subserve an intention to turn the lights off, or an intention to euthanize a patient. These intentions can only be translated to an observer by attending to broader contextual information. This problem is commonly referred to as the "ambiguity problem" (Monroy et al., 2017). Thus, direct matching seems a poor mechanism for interpreting actions in context dependent scenarios, or social scenarios. Of course, the direct matching view could counter with a familiar appeal to additional non-motor routes to action understanding. If the majority of scenarios (those where intention or communication are involved) must be referred to non-motor processes, then it seems hardly appropriate to argue that direct matching affords the observer with an optimal (or 'richer'; Gallese et al., 2004) means of decoding action.

Even the proposed evolution of a mirror neuron system which underpins the direct matching hypothesis, is problematic. To quote Jacob (2013), the problem is that “synchronous interpersonal neural similarity across two distinct brains at a single time presupposes asynchronous interpersonal neural similarity at different times in two different tasks” (p. 1134). In other words, for the mirror neuron system of both an observer and an actor to directly match, these separate mirror neuron systems must already operate similarly when performing and observing actions. For such a view to be intelligible, it must commit to the claim that pairs of individuals are selected by evolution. However, as Jacob (2013) notes, natural selection operates on the level of the individual, not pairs of individuals.

In an earlier paper, Hurley (2008) argues that perhaps interpersonal neural similarity is merely a corollary of a more central process. Perhaps mirror neurons are instead intrapersonally ‘re-used’ by an observer in order to interpret another’s behavior. ‘Re-use’ would then be selected by evolution, and interpersonal neural similarity would gradually emerge, but as a byproduct. Jacob (2013) again challenges this view, noting that Hurley (2008) is assuming that mindreading is the fundamental ability that motor simulation is facilitating. Jacob (2013) proposes an alternative: that the function of mirror neurons is to identify an action concept (e.g., grasping). This concept may then become instrumental in such operations as mindreading, but this (and other) larger-scale operations should not be conflated with mirror neuron functionality.

These alternative theoretical accounts are presented here in summary. The intended effect is not to unpack each account in its entirety, but to demonstrate that the direct matching account faces considerable challenges from several fronts. Strictly speaking, it is immaterial to the present thesis if any of these alternatives mentioned above are true, and the direct matching account is false. The core proposal of this body of work is that an efficient mindreading system interacts with the motor system, but as of yet the nuances of such an interaction remain the subject of speculation. However, this is not to say that all possible alternative accounts are compatible with that core proposal, and in the next section one such alternative account is addressed, the implications of which conflict with the premises of the present investigation.

3.7.1 Statistical Learning

Each of the theories described above are themselves subject to lower level domain-general alternative accounts. Much like the behavior rules account which can readily explain infants' success on nonverbal FB tasks, a statistical learning process has been proposed for action understanding. According to such an account, humans are born equipped with impressive statistical learning abilities (Ruffman et al., 2012). Some have proposed that deductions about each action's most likely goal is accomplished using probabilistic modeling and Bayesian inferences (Kilner, Friston & Frith, 2007a, 2007b; Gilet, Diard & Bessière, 2011). There is evidence that infants and adults can discriminate between regularities in behavior (Baldwin et al., 2008; Baldwin et al., 2003), and that sensitivity to probability is detected in corresponding neural responses (Ahlheim, Stadler & Schubotz, 2014). However, statistical learning still faces a core problem: how do perceptual observations become encoded into the motor system? This problem, known as the 'interface problem' is a central issue for motor and statistical learning theories alike.

Monroy et al. (2017) investigated this 'interface problem' by asking whether the observation of either a statistically predictable or random pairing of actions was sufficient to inform infants about likely upcoming actions, but also that their anticipations could be detected in the sensorimotor suppression of their motor cortex. First, 18-month-old infants watched a series of videos showing six objects which could be interacted with in distinct ways. The presentation of the actions was such that the infants were equally exposed to all possible actions the same number of times, however, two of the actions always occurred in a pre-determined sequence (e.g., the action 'slide' was always followed by the action 'twist'), while the others were randomly ordered. During the test phase, infants were shown a deterministic action (e.g., 'slide') and random action (e.g., 'push') and their activations were measured. Results revealed that motor cortex activation was significantly lower in the deterministic condition relative to baseline, while motor activation during the random action did not differ from baseline. These findings speak in favor of the notion that statistical learning forms a foundation for building expectations based on observed visual information. Most intriguing is that infants can apparently 'feed' visual information into their motor system. This would mean that the motor system is not constrained by the observer's motor repertoire, or by any constraints (be they momentary, like tying up one's hands, or permanent, like being born with upper limb dysplasia). This alternative account is therefore troubling for motor theories and future experiments are needed to uncover how this visual information is being fed into the motor system.

3.8 Bridging the gap between vision and action: an attempt to solve the interface problem

'Action understanding' has now been diced up into a number of concepts which serve presumably non-overlapping purposes. For example, ascribing intentions to an agent is different to ascribing goals to an agent. What is less clear is where motor representations fit into either of these categories. Of course, this concern tacitly assumes that representations of goals and representations of intentions can be verified as making differential contributions to action perception.

Evidence from functional brain imagining supports such a distinction. de Lange et al. (2008) showed adults three actions, one that demonstrated an ordinary means to an ordinary goal (an actor lifting a mug to her mouth), one which demonstrated an unusual intention (lifting a mug to the actor's ear), and one which demonstrated an unusual means of bringing about an ordinary intention (lifting the mug to the actor's lips using a strange grip). In addition, participants were instructed to judge whether the intention of the actor was unusual or not, or if the means of accomplishing the intention were unusual or not. This design allowed the researchers to investigate how the activation pattern of various brain regions were modulated by either the perceived typicality of the observed action, and/or the aspect of the stimulus that participants were attending to. The researchers found that the activation of the inferior frontal gyrus (part of the human mirror neuron system) was not modulated by what participants were instructed to attend to. However, the inferior frontal gyrus *was* modulated by observing actions with unusual intentions (compared to ordinary actions) but not when observing actions with unusual means for ordinary intentions. Meanwhile, the medial prefrontal cortex, posterior cingulate cortex and right posterior temporal sulcus (all of which are part of the mindreading or 'mentalizing' network) showed greater activation when participants judged the intentions behind actions, instead of the means of those actions, but was insensitive to the visual differences in the observed actions. These results show that there is a meaningful division between neural processes designated to understanding intentions and processes for interpreting actions as directed toward certain outcomes or goals.

Nanay (2013) raises the point that a comprehensive theory of the purposiveness of action must integrate and distinguish intentions and motor representations. In taking a step toward a comprehensive account of purposive action, Butterfill and Sinigaglia (2014) confront this challenge. First, they argue that some motor representations represent action outcomes, while other motor representations are not outcomes *per se*, but instead guide actions toward certain outcomes. Where motor representations represent action outcomes, it becomes challenging to place them firmly in the camp of either goals or intentions, as either would specify an outcome. But which action outcomes do motor representations specify: the most proximal outcome of an action (e.g., to grasp a mug), or the more broadly encompassing outcome of a chain of discrete actions (e.g., to drink tea, which involves several separate actions)?

The group of researchers who discovered mirror neurons (e.g., Gallese et al., 2004; Rizzolatti et al., 2001) argue that mirror neurons decode others' intentions. However, their hypothesis appears unprepared to withstand the interface problem, as it is far from clear how visual information is first mapped onto a motor representation, and how that motor representation carries propositional content (Pacherie, 2000). To unpack the problem further, intentions are necessarily object directed. The object of a belief, for instance, is the proposition that a circumstance is true. Therefore, if motor representations are intentions, they must be able to represent the object of the propositional attitudes motivating action. While a motor system should be able to represent the dynamic and kinematic features of action, it does not follow that it could represent the object toward which the action is directed.

Butterfill and Sinigaglia (2014) propose that this apparent incompatibility can be undermined by supposing that intentions and motor representations 'interlock'. First, an analogy from Butterfill and Sinigaglia (2014) is borrowed in order to illustrate how information from differing formats can be not only compatible (to the degree that they serve complementary functions) but also non-overlapping (to the degree that neither needs to recruit properties peculiar to the other in order to interlock). Suppose that you are approached by a hurried traveler whom is asking for directions to the train station. One could go about assisting her in a number of ways, some of which differ in format. One could draw her a sketch depicting an aerial view of the surrounding streets and trace a line from one location to another. This representation provides the requested information in a cartographic format. Alternatively, one could draw the train station itself, which would provide her with information relevant to her request, but in a pictographic format. Even if the pictographic illustration of the train station was photorealistic though, it provides little help given the traveler's current distance from the station. The traveler would likely refuse the option that represents her intended outcome and instead choose to take the aerial map, which represents (albeit more crudely) information better suited to guiding her to her goal. Still, the *optimal* choice would be to take both: the aerial map guides her from the current position to the final intersection, and the illustration informs her which of the available buildings at the intersection is the train station (and thus, what actions should follow her reaching the end point of the map).

The process by which motor representations contribute to the prediction of an action outcome are metaphorically akin to the aerial map. First, motor representations are generated (when possible), and allow the observer to predict the outcome of an ongoing action. Motor representations are more crude and minimal, relative to intentions; however, for the purposes of making efficient inferences about others' goals (or coordinating one's own actions toward one's own goal), they are desirable. However, without intentions, motor representations are a high-speed vehicle lacking a pilot. How, then, are motor representations directed to outcomes, without using intentions? One possibility is that motor representations interact with minimal mindreading such that outcomes and kinematics of action can be predicted in context of what others have registered.

3.9 Motor Representations and False Belief Understanding

So far, some possible roles that motor representations could play in the understanding and production of action have been considered. As this chapter has hopefully illustrated, action understanding is a topic of lasting debate. A central focus of the present thesis is to investigate the possibility that motor representations interface with efficient mindreading. Since action understanding is essentially goal ascription, and goals are constituents of beliefs, a comprehensive account of mindreading is incomplete without incorporating action systems. Of course, the establishment of a comprehensive account of human mindreading is by no means the objective of this body of work. Our preoccupation is with efficient processes for mindreading, of which the two-systems account offers the most promising game in town. Therefore, in fostering a possible connection between action understanding and mindreading, it is asked whether and how motor representations influence a minimal mindreading system (or vice versa). As in previous chapters, mindreading will continue to be observed through the contextual lens of the FB task. First, one must ask what precedence has been set for using FB tasks to investigate motor processes.

Southgate and Vernetti (2014) tested 6-month-old infants and adults using a standard unexpected change of location FB scenario. They compared observers' motor activation (as well as their AL) after viewing one of two scenarios. In scenario one, an agent watches a ball jump inside a box in the center of a table. A curtain is lowered to occlude the agent's view, and the ball jumps out of the box and rolls off screen. Now, the agent has a FB that there is a ball in the box, but the observer knows there is not. In scenario two, the agent watches the ball jump out of the box and roll off screen. After the curtain is lowered, the ball rolls back on screen and jumps back in the box. Now, the agent has a FB that the ball is *not* in the box when the observer knows that it is. Note that in either case, the agent has a FB. However, only in scenario one does the agent's FB provide her with an impetus for producing a goal-directed action.

As predicted, the amplitude of both adults' and infants' motor system was significantly suppressed only when the agent has a FB that the object is in the box. This finding was qualified by evidence showing that adults made significantly more anticipatory first looks (as well as significantly longer fixations) toward the agent's hand when her FB would motivate an upcoming action. Taken together, this shows that the motor cortex is activated in anticipation of actions that reflect not only *that* an agent has a FB, but *what* that FB specifically entails.

These results suggest that either infants' motor representations are sophisticated and can efficiently determine what actions will be performed in light of another's FBs, or infants' motor representation suppressed activation following the efficient tracking of registrations (as a proxy for beliefs), which facilitated accurate action predictions. However, if motor representations interact with the efficient mindreading system, and if the efficient mindreading system is subject to signature limits, then motor representations should demonstrate accurate action predictions in object-location but not object-identity scenarios. A recent study by Edwards and Low (2017) tested this prediction.

The FB task developed by Edwards and Low (2017) differed from Southgate and Vernetti (2014) in two critical ways. First, they measured motor activation behaviorally rather than neurophysiologically, by analyzing response times. Activation of the motor cortex has been shown to prime the facilitation of motor responses (Confais, et al., 2012), so those whose motor cortex has been activated by the automatic anticipation of an agent's future action should show more speeded responses. Second, the researchers designed novel location and identity FB conditions. Participants first watched familiarization trials wherein an agent observes a red or blue object move inside of a center box. The agent then demonstrates their color preference by saying "Yay!" and reaching into the box to collect it, or by saying "Yuck" and remaining motionless when the object enters the box. Note that the agent's color preference was counterbalanced, but for ease of understanding, the following procedures are described in context of the agent having a preference for blue objects and an aversion to red objects.

Following familiarization trials, participants were shown FB-location and FB-identity trials. The critical difference between tasks was the object used. In the FB-location trials, the agent watched as either a red ball or a blue ball enters the box, a curtain was lowered, and the ball leaves the box; or the agent watched as either a red ball or blue ball first leaves the box, a curtain was lowered, and the ball enters the box again. In either case, the agent ends up with a FB about an object's presence or absence. In the FB-identity condition, the dual-aspect dog-robot toy from Low and Watts (2013) was used in place of the colored balls. The dog-robot enters (or exits) the box with its red or blue aspect facing the agent. The curtain was lowered and the dog-robot leaves (or returns to) the box. Finally, participants were instructed to press one of two buttons corresponding to images of the agent either grabbing the box or remaining still.

If motor representations are related to the workings of an efficient mindreading system (which can only track registrations rather than full-blown beliefs), one would expect that when the agent has a FB that a desirable object is in the box, responses should be speeded due to motor activations, when that object is a ball, but not when the object is a dog-robot toy. Since registrations encode only the relational attitude between an agent, an object and a location (but not that object's subjective appearance), motor representations linking with registrations should not be able to form the prediction that, from the agent's perspective, they see the dog-robot as blue/desirable and will therefore reach for it. In addition, when the agent has a FB that an undesirable toy is in the box, the minimal mindreading account would predict that responses will be slower when the object is a red ball and speeded when the object is a dog-robot toy. Since the participant's efficient mindreading system should register the dog-robot only on the basis of its appearance from the participant's perspective, their motor system should activate in response to viewing the blue/desirable side, priming by an incorrect action prediction. Finally, the signature limits on registrations should be relatively encapsulated from conscious processing, so a two-systems account would further predict that no corresponding pattern should be found in the error rates of participants' choices. As an explicit measure of belief reasoning such choices would be managed by distinct and flexible processes.

Results indicated that when the agent had a FB that a desired object was in the box, responses were slowest when the agent's FB resulted from a misrepresentation about the object's identity. Furthermore, when the agent had a FB that an undesirable toy was in the box, response times revealed the opposite pattern: now participants made significantly faster choices about the dog-robot than the ball. Importantly, there was no difference in the error rates across conditions. Taken together, results align with a 2-systems view: signature limits should apply to efficient processes but not overt conscious choices. By contrast, a one-system account would not find these results easy to explain. One might argue that the difference in timing has to do with a lower-level perceptual bias, as the participant has to inhibit what they can see of the dog-robot, before they can appreciate that the agent sees it differently. However, this explanation is not supported by response times, as responses were faster in the FB-identity condition when the agent's FB is that an undesirable toy is in the box.

Alternatively, a one-system account may counter this argument by suggesting that responses should be quicker when the perceptual bias leads participants to erroneously conclude that the dog-robot is, say, red because that's how it looks from where they are. However, this would also imply that such patterns from lower-level biases specific to the FB-identity condition should be reflected in error rates as well as the response time for those choices. While it was reported that the proportion of errors was significantly higher in FB-identity conditions than FB-locations conditions, there was no significant difference in error rates across FB-identity conditions. This would suggest that although the identity tasks were more challenging than the location tasks, error rates did not follow the pattern observed in response times: more speeded responses were not more likely to be correct than slower responses, and vice versa. In sum, these results do not map onto a one-system account of mindreading, while they map comfortably onto the predictions from a two-systems account.

3.9.1 Efficient Mindreading Mechanics: A new frontier

In a recent review of the evidence for distinct mindreading systems, Low and colleagues (2016) noted that "research is needed to map the terrain of the efficient … mindreading system" (pp. 13). The topology of such a terrain could be elucidated by converging evidence showing that signature limits are robust across different FB tasks and measures. Chapter 2 discussed evidence that signature limits are found in FB identity tasks measured by AL (Low & Watts, 2013; Low et al., 2014; Wang, Hadi & Low, 2015), response times and error rates (Edwards & Low, 2017; Surtees, Samson & Apperly, 2016), and young children's spontaneous helping (Fizke et al., 2017). Edwards & Low (2017) provide results that support the conjecture that there may be interactions between motor representations and the efficient mindreading system, as their reported pattern of response time facilitation suggests that participants' motor cortex was activated by the same erroneous anticipation of an agent's action that would be expected from a system operating on registrations rather than full-blown beliefs.

3.10 Conclusions

In the following experimental chapters, a new measure of implicit belief reasoning is considered in an interactive FB. In chapter 4 (Experiment 1), it is shown that watching an agent approach a location under a false belief, will trigger looks revelatory of the observers' own planned actions. However, this prediction is only valid when participants in a position to provide assistance to an agent. To this end, an active helping FB task (adapted from Buttelmann et al., 2009) with matching location-FB and identity-FB conditions, as well as matched true belief conditions is presented.

CHAPTER 4

Experiment 1: Signature Limit on False Belief Tracking in An Active Helping Task

4.1 Introduction

How does tracking someone's belief enable humans to cooperate more effectively with one another? Answering this question requires acknowledging and measuring the contributions from different kinds of belief-tracking processes. One theory that is gaining traction delimits two systems for mindreading: one which efficiently tracks mental states, and one which represents mental states as such (Low et al., 2016). Efficient mindreading (also referred to in the present chapter as 'minimal' mindreading) incurs processing limits, whereas flexible mindreading is slow but equipped to represent conceptually challenging mental states.

Here, an investigation is conducted into the signature limits on efficient belief reasoning. A naturalistic-seeming scenario was designed in which adults are asked to help an agent who comes to hold a FB about either the location of their red backpack ('he doesn't know that his backpack was moved') or the identity of their backpack ('he thinks that the red backpack he is looking at is *his* red backpack; but it's really an imposter'). First, a new protocol for comparing patterns of looking during an active helping task is considered. It is then asked whether signature limits (previously evidenced by measuring anticipatory eye gaze) are detected when measuring predictive looking. Next, the focus shifts to recent evidence suggesting that efficient belief reasoning is linked to motor processes (Edwards & Low, 2017), by asking whether differences in looking correlate with the directions participants move, across distinct stages of their helping behavior.

4.1.1. Addressing the gap in the literature: A Compression identity FB task

Do humans suffer signature limits when attempting to efficiently process other's FBs about an object's identity? Some studies measuring infant looking behavior while observing a complex social scenario suggest that they do not (Scott & Baillargeon, 2009; Scott et al., 2015), while other studies have shown that even adults' form incorrect anticipations in identity FB scenarios (Low & Watts, 2013; Low et al., 2014; Wang, Hadi & Low, 2015; Mozuraitis, Chambers & Daneman, 2015). One problem with the available evidence is that the vast majority of identity FB tasks are 'expansion' identity tasks. In an expansion identity FB scenario, an agent develops a FB that there are two distinct objects, when really there is one object, with two distinct aspects. It remains an open question whether or not the signature limits which have been observed in ‘expansion’ identity FB tasks (Low & Watts, 2013; Fizke et al., 2017) also are found in the reverse scenario: the ‘compression’ identity FB task.

A person develops a ‘compression’ identity FB when they believe that only one object is present, when actually there are two identical objects which have swapped places. The common ground between the expansion and compression identity tasks is that they both involve numerical discrepancies between the agent's belief and reality. Only a couple studies have attempted to design compression identity FB tasks (Scott & Baillargeon, 2009; Scott, Richman & Baillargeon, 2015). These studies reported that infants succeed at predicting how an agent with an identity FB, will behave. However, the agent in either study sees that there are two identical objects, at some point during the procedure. Therefore, these studies may not have presented a compression identity FB scenario, in the strict sense. In the absence of the numerical aspect, it is possible that understanding the agent’s FB is about as challenging as understanding a FB about an object's location (see section 2.2.2; Low et al., 2016). This leaves the compression identity FB task untested in the empirical mindreading literature. The present thesis attempts to address this gap in the literature by testing a novel compression identity FB task, where the agent is mistaken about the number of objects in the scene.

4.1.2. The ecological validity of FB tasks

One quality that is rarely prioritized in a FB task is its ecological validity. The ecological validity of a study refers to the degree to which its results represent the actual behaviors of interest as they occur in natural settings (Barkley, 1991). It is often taken for granted, or dismissed as unimportant, whether participants’ VOE, AL or helping responses during a FB task are generalizable to their anticipations and behaviors when they reason about FBs in their day-to-day lives.

To be sure, there are a multitude of other issues which may rightly take priority over concerns about low ecological validity. For example, one might argue that future research has more pressing matters to attend to, like the present replication crisis or the lack of convergent validity from studies using implicit measures. However, it may be wrong to dismiss ecological validity as irrelevant to these discussions. For instance, one reason why reported effects from superficially dissimilar experiments that measure the same phenomenon are nonetheless strongly correlated (as is the case with explicit FB tasks), is because these studies tap a naturalistic behavior. If the available implicit FB tasks do not replicate and their findings do not correlate, one of the possible reasons for this is that the available studies have low ecological validity.

A few studies have designed tasks where the agent's FB or ignorance is treated as authentic. In the 'director task' (Keysar, Linn & Barr, 2000; 2003), the researchers successfully deceived participants into believing that a confederate participant's FB was genuine. However, participants still knew that they were presently taking part in an experiment. The present study aimed to take a further step toward maximizing ecological validity, by presenting an agent whose FB seems to be both authentic and also entirely separate from the experiment in which they happen to be taking part.

One FB task which *could* achieve high ecological validity, is the active helping task (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Knudsen & Liszkowski, 2012). Here, it is tacitly assumed that the infant in the study believes that the agent does not know that the target object has been moved during the agent's absence. However, a wealth of recent conceptual replications of the Buttelmann, Carpenter and Tomasello (2009) FB helping task have shown that it is not safe to assume what concepts infants are operating on when they respond in these tasks (Allen, 2015; Priewasser et al., 2017; Crivello & Poulin-Dubois, 2017)

Nevertheless, it is explored whether the active helping task could be adapted in such a way so that the participant not only believes that an agent's FB is real, but that the conditions which led to them obtaining that FB were not caused by an artefact of the experiment. In doing so, the aim was to achieve higher ecologically validity relative to previous FB tasks. It was reasoned that accomplishing these goals would be possible through the use of an active helping task adapted for use with adult participants.

4.1.3 Efficient mindreading in an active helping task

The findings presented in this thesis are collected from adult participants. The reasons for this are two-fold. First, evidence of signature limits using implicit measures, are more meaningful when they are recorded from an adult sample, as it is relatively safe to assume that adults possess fully fledged mindreading abilities, and their executive functioning abilities are also sufficiently mature. If adults show evidence of signature limits in their performance, it is unlikely that these limits are caused by extraneous task difficulties (as is often argued to be a plausible alternative explanation for errors in young children's responses; Scott & Roby, 2015; Roby & Scott, 2016). Second, unlike several other FB tasks involving non-verbal measures, there have been no active helping FB task designed to tests adults. Traditionally, active helping FB tasks recruit infants or young children (Buttelmann, Carpenter & Tomasello, 2009; Buttelmann, Suhrke & Buttelmann, 2015; Fizke et al., 2017). However, inferring the underlying rationale behind infants' and young children's helping behaviors has become a source of controversy (Allen, 2015; Priewasser et al., 2017; Crivello & Poulin-Dubois, 2017). Theoretical motivations aside, there are methodological benefits to testing younger populations in helping tasks. Infants are highly gullible, and easily convinced about the authenticity of an agent's FB. The same cannot be said for adults, who are often suspicious about the true purpose of an experiment (sometimes leading to fallacious effects; see Durgin et al., 2009). To test how adults provide help to an agent *whom they genuinely believe has come to hold a FB*, an assortment of task features must be introduced and rehearsed in order to achieve deception. In the present study, this challenge was undertaken by engaging in a carefully rehearsed sequence of events which collectively revolved around a clear FB narrative.

4.1.4 Departing from the anticipatory looking framework: How eye movements can and cannot be measured in an active helping FB task

Studies measuring helping and studies measuring eye gaze are the predominant means for studying implicit mindreading, and the only means for doing so with infants. However, there are no studies which measure one's eye gaze during the active helping task. Helping tasks focus solely on overt choices (e.g., Buttelmann, Carpenter & Tomasello, 2009; Buttelmann, et al., 2015). The nearest that a study has come to measuring predictive eye gaze is Knudsen and Liszkowski (2012), who measured spontaneous pointing and found that 12-month-old infants spontaneously point to an object's *actual* location before an agent acts erroneously, on the basis of his FB.

The omission of looking measurements in FB helping tasks is likely the result of challenging differences in methodology, when compared with established looking paradigms. For example, the principal measure used in Southgate, Senju and Csibra's (2007) AL FB task was the direction of the infants' first saccade within a certain time window. Specifically, after a light and sound cue (signaling that an agent was about to make a choice) the authors expected that infants would look *first* to the 'empty box', in anticipation of the agent's next move, rather than toward the 'full box', where the object actually was. One reason that this hypothesis is considered a valid measure of implicit belief reasoning, is that looks to the 'empty box' are homologous to what would be the correct verbal response to the standard FB question: 'where will the agent go to find his toy?' However, in active helping FB tasks, the correct way to help is by opening the *full* box, where the object is presently located. On first glance, this appears to imply a contradiction. In the active helping task, successful belief reasoning is demonstrated by the behavior which would correspond to the incorrect and egocentric response in an AL task. However, the reasons why these two correct responses actually do not imply a contradiction, can be explained by taking into consideration each procedure's underlying social context.

In the active helping task, the correct dependent response is only correct given that the participant has motivation to act. This motivation is quelled under certain social contexts. For instance, a child in a standard verbal FB task does not try to help the character in the story because storytelling has its own underlying social context[1] (i.e., a listener cannot affect the actions and outcomes of the characters in the story). Social contexts can also differ when stimuli are viewed on a video screen instead of in a live event (e.g., Libertus & Needham, 2010). Much like storytelling, humans do not engage with people that are seen on TV in the same way as people in one's shared environment. If an observer sees that someone nearby has developed a FB, and the observer watches that agent begin to act in error on the basis of that FB, this real-world social context supplies the observer with an appropriate context in which to act. Unlike the social contexts of an AL FB task, here an observer *can* effectively change the outcome of that agent's upcoming actions. The logical thing to do in this case, is to approach or point out the location that is currently housing the object. Doing so accomplishes the goal of bringing the agent's belief in line with reality and doing so (in this social context) carries with it its own pro-social value (Warneken, 2013).

Does this difference in social contexts mean that looking measures are a fruitless measure for testing FB reasoning in helping scenarios? Not necessarily. Low et al. (2014), used measures of AL and verbal predictions on a cooperative bridge-moving task. In the task, a table top map

[1] However, this may differ according to culture. Wang, Hadi, and Low (2015) suggest that Semai children may mis-interpret verbal FB questions as asking where the agent *should* go to get the object.

and actions figures were arranged so that two agents were stationed on separate islands, each with a bridge connected to a center island. The target object (a rabbit) was stationed on a third unconnected island. Children (3- and 4-year-olds) listened as a story where it was revealed that one character (Bart) either wanted the rabbit only for himself or wanted the other character (John) to help him get the rabbit together. After Bart crosses to the center island, an experimenter says, “I wonder which bridge Bart will move in order to get the rabbit?” Moving Bart’s bridge would allow both characters to cross over to the island with the rabbit (cooperative), whereas moving John’s bridge would leave John stranded so that only Bart could reach the rabbit (competitive). Results showed that after the ‘I wonder’ prompt, children in the cooperative condition tended to look first to the correct bridge (Bart’s) and also looked there for longer, during the critical time window (1,750 ms). By contrast, in the competitive condition, children in both age groups tended to look first toward the wrong bridge (Bart’s) and looked longer in this direction, whereas verbal predictions increased as a function of age, from below chance in 3-year-olds to above chance in 4-year-olds. It is noteworthy that children’s eye movements behaved as predicted, because Low and colleagues (2014) did not preface the main task with familiarization trials, which are characteristic of some earlier AL tasks (e.g., Southgate, Senju & Csibra, 2007).

If even young children make predictable saccades toward cooperative solutions, it would seem plausible to suppose that adults would make saccades in a helping task. However, determining which direction of looking and when the critical time window looking should be set, as well as what future outcomes eye movements should be interpreted as predicting, is difficult for testing FB reasoning in a naturalistic task. First, a predictive looking hypothesis during an active helping FB task cannot straightforwardly rely on a comparison of looks made toward one location (the empty box) or the other (the full box). The reasons why this cannot be done are largely pragmatic.

Next, one must consider *when* during the helping task event it would be appropriate to capture participants’ eye gaze. In an ideal scenario, participants could be conditioned or prompted to expect that the agent is about to try to retrieve an object after an agent stands still and equidistant from either of the two possible locations housing an object. However, it is difficult to achieve this vignette in a naturalistic scenario. What nonverbal event could provide a signal to the participant that the agent is about to reach for an object they last stored, without cueing participants in on the true purpose of the study? The most natural cue would be the agent's unfolding actions themselves. The problem with this is that looks to the empty box can hardly be considered "anticipatory" if the agent is already directing their actions to one of locations when the critical time window begins. If one collects data on participants' directional eye gaze during this unfolding action, then perceptual side biases cannot be controlled. This confounds any standard comparison of a participant's AL toward the empty vs. full box, as looks to the empty box are just as likely be anticipations of the outcome of the agent's actions based solely on the agent's ongoing actions. Moreover, as the agent approaches the empty box, it becomes increasingly challenging for an eye gaze coder to determine whether a participant's eyes are tracking the agent or targeting the empty box. The bottom line is that any comparison of looking patterns during an active helping FB task cannot simply recycle the logic used to rationalize AL.

4.1.5 A predictive gaze measure for active helping tasks

What of the other location, the space housing the object that the agent truly wants? In existing helping tasks, this is where infants and children tend to direct their actions in order to help the agent. In a FB helping task, when the agent approaches the empty box, the intention behind this action is to obtain the object that is in the other box. For the purposes of differentiation, this location will be hereafter referred to simply as "compartment *Y*" whereas the location which the agent is trying to reach will be referred to as "compartment *X*". In the active helping task, one *can* measure the frequency and duration of an observer's looks to compartment *Y* without incurring the same confounds that delegitimize looks to compartment *X* (where the agent is already directing attention to). In terms of attracting a participant's attention, the compartment *Y* is constantly competing with the events taking place in the opposing direction. Thus, what is problematic for measuring looks to compartment *X* is advantageous for measuring looks to compartment *Y*, as all of the most salient and distracting stimuli in the task are allocated to one side, leaving compartment *Y* isolated on the other side. In summary, looks directed to compartment *X* during the agent's approach, are not experimentally sound candidates for a comparative eye gaze analysis, but looks to compartment *Y* might be.

However, what benefit could an analysis of looks to compartment *Y* offer for furthering the understanding of FB reasoning? The present study suggests that a difference in looks to compartment *Y* can tap into signature limits on efficient mindreading. To flesh out this idea requires briefly revisiting the two-systems account's minimal mindreading framework, homing in on its signature limits, with the view of explaining why the minimal mindreading system would wrongly interpret an agent who has a compression identity FB, as having a TB. There is currently no compression identity FB task that has been tested which fits the conditions proposed by Butterfill and Apperly (2013). The present thesis aims to address this gap in the literature.

4.1.6 How minimal mindreading contradicts itself: "the object that was moved to the right is about to be retrieved on the left".

Minimal mindreading operates on certain principles which acknowledge a registration's correctness. One of these principles states that "correct registration is a condition of successful action. More precisely, in order [for an agent] to successfully perform a goal-directed action with a goal that specifies a particular object, the agent must correctly register that object" (Butterfill & Apperly, 2013, p. 617). The present task intends to investigate the following question: In a compression identity FB helping scenario, how would a minimal mindreading system determine whether the agent's last registration is a condition of successful goal-directed action? This task may be most clearly described by presenting the event sequence of the FB helping scenario to come. First, an agent leaves an object (let's say the object is a red backpack) in a certain compartment *X*, and then leaves the room. Before the agent returns, a stranger moves the agent's red backpack to a different compartment *Y*, then stores his own backpack in compartment *X*, and leaves. Next, the agent returns and starts to walk toward compartment *X*. A 'location' FB version of this event sequence requires that the stranger's backpack and the agent's backpack look different. By contrast, an 'identity' FB version requires that the agent carry a backpack that is identical to the stranger's and that both the agent and the stranger store their respective backpacks in the same way inside compartment *X* (i.e., such that the orientation and positioning of the stranger's bag is also identical to the bag that the agent left in compartment *X*).

Beginning with the FB location scenario, what would a two-systems account predict about an observer's eye movements having processed the agent's registration? Because the observer has seen the red backpack moved to compartment *Y*, and the observer has attributed to the agent a previous registration that specifies his or her red backpack in compartment *X*, the output of the minimal mindreading system should guide the observer's eyes toward compartment *X*.

This predictions is suitable for an AL framework. However, as discussed earlier, a naturalistic helping task does not provide opportunities to condition an observer to expect that the agent is about to act. Moreover, these looks are defined as anticipatory: they *antedate* action. How then should a minimal mindreading system guide an observer's eyes *during* the agent's approach of compartment *X*? Before forging an answer to this question, I would like to pause and openly acknowledge that a full answer to this question cannot be proposed by referring exclusively to the hypotheses stated *explicitly* by the two-systems account. How the eyes should behave in a FB helping scenarios has not been considered by Apperly, Butterfill, Low or other two-systems advocates. Therefore, the following explanation keeps as close to the proposed workings of the two-systems model when possible, and announces any departures into uncharted territories, when necessary.

What aspects of the environment might an observer's eyes be drawn to as an agent begins to approach compartment *X*? In this FB location helping scenario, there seem to be at most, three possible targets. The first is the agent: an observer may simply track the agent's movements. The second is compartment *X*: the observer may continuously target this anticipated outcome as the agent approaches it; however, this could just as likely be informed by a non-mindreading prediction of the outcome of an observed ongoing action. The third possible target is compartment *Y*: the observer may produce looks to the compartment housing the red backpack that the agent falsely believes he or she is about to obtain. Looks tracking the agent are not helpful for measuring FB tracking and looks to compartment *X* are confounded by the direction the agent is moving toward. If those targets are invalid, what function would be served by producing looks to compartment *Y*? Furthermore, if the minimal mindreading system guides eye movements that serve some predictive or planning purpose, what is it that looks to compartment *Y* are foreshadowing? The short answer that this chapter advances is that, in a FB scenario, looks to compartment *Y* are made in preparation of the observer's own goal to help the agent whom they can detect is about to act in error. Doing so requires that the minimal mindreading system also efficiently ascribes goals (or at the very least, that minimal mindreading *informs* efficient goal ascription processes).

Butterfill and Apperly (2016) have considered how the minimal mindreading system would manage goal ascription in an efficient manner. According to their model, belief tracking allows an observer to generate predictions that are based on both facts about the environment, and facts about the agent's registration. In a FB circumstance, belief tracking implies understanding a distinction between the outcome of an agent's action (i.e., he will open compartment *X*) and that which is the agent's goal according to his or her last registration (i.e., to retrieve the red backpack which was last registered in compartment *X*). Butterfill and Apperly (2016) suggest that cognitively efficient belief tracking informs efficient goal ascription, as the agent's last registration contains the information about the object which is the agent's goal. By employing efficient belief tracking and efficient goal tracking, an observer should be able to efficiently calculate that an agent's registration is not a condition of a successful action directed outcome (his or her backpack is in compartment *Y*). Therefore, in the FB location helping scenario, the observer may show looks to compartment *Y*.

However, hypothesizing that an observer should show one type of look does not specify how one should score those looks to most effectively capture the saccade(s) that reflect, in the present case, the output from an efficient mindreading system. In AL paradigms, two classes of looking measures are typically used: first look and/or preferential looking duration. However, recent evidence has shown that these measures can lead to different passing rates on the same task. Burnside et al. (2017) found that children (3- to 7-year-olds) performed significantly below chance in a FB location AL task when scoring for first look or preferential looking duration; whereas adults performed at chance when scoring their first looks and were significantly above chance (64% passing) when scoring their preferential looking duration. Based on their findings, Burnside and colleagues advise that, when testing adults, preferential looking duration is the more robust measure, particularly when the task biases participants to look in one direction during the critical time window. In addition, Burnside and colleagues stress the importance of reporting multiple measures of predictive looking. Although, the present study makes several departures from the AL framework (e.g., focusing on looking toward a single location, excluding prompts, and measuring gaze *while* the agent acts on his or her FB), the following experiments report the percentage of participants who show a look to compartment *Y* during the critical time window, as well as the average duration of looking to compartment *Y* per condition (which is the measure that will be used to test the present hypotheses, per the above mentioned suggestion).

With the looking duration measure specified, the predictions from a location or identity FB helping task can be unpacked. According to a two-systems view, a minimal mindreading system would incorrectly interpret the registration of the agent as being a condition of successful action, when the agent has a compression identity FB. Therefore, it is predicted that the average duration of looking to compartment *Y* will be significantly shorter in the identity FB scenario, than in the location FB scenario. To qualify this prediction, it helps to review the identity FB event chronologically and consider how the event would be processed by a minimal mindreading system. The compression identity FB scenario differs from the location FB scenario in two respects: (1) the stranger replaces the agent's red backpack with an identical red backpack, and (2) when these backpacks are stored in compartment *X* they are at least partially visible. Thus, when the agent returns, he will see a red backpack in compartment *X*, and attempt to retrieve it.

As the stranger moves the agent's backpack from compartment *X* to compartment *Y*, the minimal mindreading system can already detect that the agent's last registration (i.e., that the red backpack is in compartment *X*) is not a condition for successful action. Next, the stranger places his identical red backpack in compartment *X*. Upon seeing the room this way (with the stranger's red backpack in compartment *X*) the critical question is: can minimal mindreading system keep separate that the red backpack and that is now in compartment *X* is *not* the same red backpack that was just placed in compartment *Y and* that the red backpack in compartment *Y* is the one that the agent has the goal of retrieving? It is argued that a two-systems account would say that a minimal mindreading system could not. According to Butterfill and Apperly (2013), a minimal system needs to operate largely independent of working memory, attention and inhibition. However, it is difficult to imagine a cognitive strategy that could successfully keep the two identical backpacks associated with their owner and allocated to the correct places without drawing from either working memory or inhibition. In addition, one main difference between minimal mindreading and flexible mindreading, is that only the flexible system uses (and can internally manipulate) visual representations. Manipulating visual representations allows the flexible system to resolve complex differences between perspectives. However, the minimal system can neither access nor manipulate visual representations and visual experiences. Accommodating representational processing costs would stymie its efficiency. A flexible system could choose between a host of ways to differentiate two identical objects from each other. For example, it could assign separate labels to either of the backpacks based on which backpack belongs to whom (e.g., ‘it is not just a red backpack that is in compartment *X*, it is *Edward’s* backpack’). The minimal mindreading system has none of these tricks up its sleeve to overcome the identical appearance issue. Therefore, when the agent returns to the room and begins to approach the red backpack in compartment *X*, a minimal mindreading would *itself* struggle (if not outright fail from the get-go) to hold these two objects separate.

This brings us to the critical time window in the identity FB helping scenario. The agent has re-entered the room and is beginning to direct his actions toward compartment *X*. What does the minimal mindreading system think is happening as the agent does so? A review of the agent’s last registration will inform the observer’s minimal system that the agent’s last registration was of a red backpack in compartment *X*. Is the agent’s last registration a condition of goal-directed action? Since the observer can see that there *is* a red backpack in compartment *X*, the minimal mindreading system should come to the *incorrect* conclusion that the agent’s registration *is* a condition of goal-directed action. In other words, unlike in the FB location scenario, an observer’s minimal system should output the same information as it would if the agent was acting on a TB. How would this modulate eye gaze compared to the FB location task? Well, in either case, one would suspect looking to compartment *X*. However, the incorrect output from the minimal mindreading system, in the FB identity condition, should result in a reduction in the duration of looking toward compartment *Y*, where the agent’s actual backpack is. The minimal mindreading system incorrectly judges the agent’s last registration to be a condition of successful goal-directed action and thus, will not associate the outcome to which the action is directed with the object that the agent actually intends to obtain. Such a reduction in looking to location *Y* in the FB identity condition would fit with the notion that a minimal mindreading system is operating, and reliably errs on identity FB tasks.

Importantly, the predicted difference in looking toward compartment *Y*, between location and identity scenarios should not be found to differ in TB versions of these two scenarios. If the agent is present and attending to all of the events in the scene, the observer’s minimal mindreading system should expect that the agent’s last registration is a condition of successful goal directed action no matter what is placed in compartment *X*.

4.1.7 A summary of the hypothesized patterns of looking

With these four distinct conditions in mind, it is predicted that the results from

Experiment 1 will reveal a shorter duration of looking to compartment *Y* in the FB identity scenario compared to the FB location scenario. By contrast, it is predicted that no significant difference will be found in the average duration of looking to compartment *Y*, when comparing the TB identity and TB location scenarios.

4.1.8 A motor hypothesis revisited: Does the minimal mindreading hold sway over body movements?

The second research question that Experiment 1 was designed to address was efficient mindreading processes interface with motor processes to facilitate an appropriate action (see Chapter 3). Advances in theories of human mindreading are incomplete without taking into consideration how mindreading connects with theories of action. In the present study, a critical aim was to analyze full-body movements that all participants naturally produce when they provided help. To do this, a lab room was spatially arranged in such a way as to constrain participants' movements toward either compartment *X*, or compartment *Y*, across four distinct stages-of-actions following the agent's request for help.

The first stage of action was *swerving*: after participants stand up, the side of the desk that they choose to swerve around was coded. The second stage of action was *approaching*: as participants walk forward from the desk, the compartment they are standing in front of when they stopped walking, was coded. The third stage of action was *reaching*: the compartment that the participant first reached up toward (operationalized by their hand extending above their shoulder after they have finished approaching) was coded. The fourth stage of action was *retrieving*: the backpack that the participant ultimately handed to the agent was coded.

In this way, forced-choice directional full-body movements that took place from the moment the agent requested for help and their final backpack selection could be coded. In so doing, it was possible to test for relationships between full-body movements and whether participants had looked to compartment *Y*. Why should any relationship be warranted? Recall that Edwards and Low (2017) presented evidence that signature limits can be found in adults' motor cortex activation by way of tracking their response time. They found that the speed of adults' responses suggested that their motor cortex was incorrectly pre-activated: it anticipated an agent would act on the basis of how an object (with two aspects) appeared *to the observer*, instead of the agent.

If signature limits are revealed by analyzing certain eye movements, as well as by measuring sequelae of motor cortex activation, perhaps they can also be measured by analyzing full-body movements. If so, it is predicted that looking to compartment *Y* will be predictive of earlier (but not later) stages-of-action. Since the present work is the first exploration into traces of the minimal mindreading system's influence on full-body movements, it cannot be said that one has reason to predict that there would *only* be a correlation between looking to compartment *Y* and, say, swerving actions, or to both swerving and reaching but not approaching or retrieving. However, based on a central attribute of the efficient system (it's speedy and quick-fading, relative to the slow but lasting flexible mindreading) it is predicted that there would be significant correlation between looking to compartment *Y* and earlier stages of action, and that this correlation will disappear with later stages of action, as more flexible mindreading operations come online and take control.

4.2. Methods

4.2.1 Eye gaze - Data collection protocols

It is important to note that one of the primary goals (to test adults in an ecologically valid FB task) constrains some methodological practices that have become increasingly standard in studies measuring eye gaze in the FB literature. For example, several AL studies make use of sensitive high-fidelity eye tracking equipment. This exploration necessarily restricts the use of such equipment as it would inform participants that their eye movements are being measured (and thus, heighten suspicion about the true purpose of the study). In addition, using this equipment typically requires participants to keep their head steady in a chin rest to stabilize viewing distance, which makes further sacrifices to ecological validity and the task's believability. Instead, this experiment used a high resolution (1920x1080) frame-by-frame (30 frames per second) video recording analysis of participants' eye movements. One risk to scoring eye gaze using human coders is that even a high-resolution frame-by-frame analysis can present ambiguous eye movements which can be challenging to resolve. In order to minimize these ambiguities, it was ensured that the storage compartments for each object was sufficiently far enough from each other to distinguish looks targeting either compartment (space between compartments = 94 cm).

4.2.2 Participants

A per-condition sample size of 24 participants per condition was determined. This sample size allowed for enough participants to fully counterbalance for side biases. The present study was also informed by Buttelmann et al. (2009), who used a sample of 12 infants per condition. Unlike Buttelmann et al.'s study, the agent in the present study was played by either a male or female actor. Thus, the sample size was doubled to fully counterbalance for side biases across both actors. Although this sample size is relatively small compared to some samples that have been recruited in recent helping task replication studies (e.g., Priewasser, et al., 2017; Crivello & Poulin-Dubois, 2017) the present study did not look *exclusively* at each participants' overt choices. Instead, a frame-by-frame analysis of participants' eye gaze and a graded set of their body movements were analyzed, meaning that the number of data points per subject was relatively large. To clarify, in Crivello and Poulin-Dubois' (2017) second experiment, a large sample of infants was tested in the FB task (n = 97). Since their primary measure was infants' overt choices, their number of resulting data points is also 97. For the present study's FB location condition, 24 adults were tested. However, the sum of the total number of frames where their eye gaze was captured plus the addition of four body movement data points per participant yielded a total of 2,279 data points. Moreover, the strategy of using a relatively small number of participants with a large number of data points per participant is consistent with contemporary research strategies investigating sensory and motor systems (Yu & Smith, 2012; Blake, Tadin, Sobel, Raissian, & Chong, 2006; Jovancevic-Misic & Hayhoe, 2009). Therefore, across 4 conditions, a total of 96 undergraduate adults (64 females) between 17 and 45 years (M = 19.16 years) participated in exchange for course credit. All participants provided informed consent prior to the start of the experiment.

4.2.3 Design

All procedures reported in this thesis were reviewed and approved by the Psychology Sub-Committee of the Human Ethics Committee of Victoria University of Wellington. Participants were randomly allocated into one of four conditions: False Belief Identity (FBID), False Belief Location (FBLoc), True Belief Identity (TBID) and True Belief Location (TBLoc).

4.2.4 Materials & Procedure

An assistant (*A*) led the participant (*P*) into the testing room (room length: 4.93m, width: 5.23m). *A* wore a sling on their left arm and a red backpack was slung over their right shoulder. Once in the room, *A* asked *P* to take a seat at a center table (length: 40cm, width: 60cm, height: 71cm). Standing in the far corner of the room (side counterbalanced) was a researcher (*R*). *R* stood facing away from *P* and appeared busy doing work at a standing desk. *R* wore a white lab coat (to convey an impression of authority) and a large pair of headphones, which were intended to deter *P* from trying to communicate with *R* at any point during the procedure. A bookshelf spanned the far wall facing *P*. The two bottom rows of the shelf were filled with books and folders, and these items were arranged in a symmetrical (but seemingly-random) manner so that neither side of the shelf or side of the room was more visually alluring than the other. The top rows consisted of two empty storage compartments (compartment *X* and compartment *Y*) with flaps that opened outwards and up (top shelf was 2.25m above the floor; see Figure 1). A 1.3m step ladder stood underneath either compartment-x or compartment-y (side counterbalanced). In this experiment, two plain-clothes Asian actors (1 female, 1 male) took turns playing the role of *A*, and a European male actor played the role of *R*.

A instructed *P* that the study would consist of two parts: completing a timed block puzzle and filling in a questionnaire. *A* pointed to the plastic blocks on the table and instructed *P* to assemble the blocks to match a pattern as quickly as possible[2]. *A* pressed a stopwatch at the start and end of the puzzle's completion; however, as this test was not actually part of the assessed task, their scores were not used in the analysis. The events following this task were different depending on the condition.

[2] The puzzle was an item from the block design sub-test of the Wechsler Intelligence Scale, 2003.

Figure 1 - The testing room from *P*'s perspective. Red boxes indicate the two compartments (*X* and *Y*). Which compartment was *X*, and which was *Y* was counterbalanced.

4.2.4.1 False Belief - Identity Condition

Once *P* completed the puzzle, *A* removed the blocks from the table and said, “I’ll just go out and get your questionnaire from the printer” (see Figure 2.1). *A* then feigned hesitation and, looking at the backpack, said, “Oh! I’ll just put this away first.” *A* climbed the ladder and stored the backpack inside compartment *X*. Importantly, *A* stored the backpack in an upright orientation, so that the top can be seen when stored (Figure 2.2). Finally, *A* walked out of the room and closed the door. After a two second pause, *R* stirred for the first time. *R* turned around (avoiding eye contact with *P*) and looked up at *A*’s backpack in compartment *X*. *R* then shook his head and exclaimed “tsk-tsk”, expressing frustration (Figure 2.3). *R* then reached up to remove *A*’s backpack from compartment *X* and promptly moved it into compartment *Y* (Figure 2.4). Note that *R* stored the backpack sideways in compartment *Y*, so that when the flap was down, the backpack could not be seen. *R* then quickly returned to his standing desk, reached down and withdrew an identical red backpack (Figure 2.5 - 2.6). *R* stored his red backpack in compartment

X such that it is positioned in an upright orientation (Figure 2.7). Thus, the scene looked identical to how *A* had last seen it, with a red backpack partially visible in compartment *X*. Therefore, upon entry, *A* held a FB that the backpack in compartment *X* was the backpack that was originally stored there. Finally, *R* picked up the ladder, closed it, and walked out of the room with it (Figure 2.8).

The sound of *R* closing the ladder cued *A* to re-enter the room. *R* passed *A* and closed the door as he exited. Meanwhile, *A* was pretending to take an unexpected phone call. The phone call was meant to deter *P* from attempting to alert *A* about the backpack switch too early. *A* walked first to a calendar on the wall near compartment *X* and said, “Okay, yes, we can fit you in on the 22nd. Okay thanks. Bye!” *A* then pocketed the phone, turned toward compartment *X*, approached the compartment and immediately jumped up reaching for the backpack (Figure 2.9). *A* failed to reach the backpack and made a pained inhalation of breath as s/he nurses their injured arm[3]. *A* immediately repeated this action but again failed to reach it. *A* then turned to *P* and said, “Can you help?” (Figure 2.10).

4.2.4.2 False Belief - Location Condition

The False Belief Location condition (FBLoc) was similar to the FBID condition with the exception of the following details. First, after *R* moved *A*’s red backpack from compartment *X* to compartment *Y*, he withdrew from his standing desk a black backpack instead of an identical red backpack. Second, when *A* originally stored the red backpack in compartment *X*, s/he did not store it upright but sideways so that it was invisible inside the compartment. The reason for this was that, if *A* stored the backpack upright in compartment *X*, *A* will return and see a different looking backpack visible in compartment *X*. If s/he jumped to reach it nonetheless, *A*’s intentions would be unclear to *P* (‘does s/he want a black backpack now?’).

[3] Actors were intentionally chosen for the role of *A* who were too short to reach the compartment without the ladder (1.58m and 1.68m tall, respectively) while the actor playing the role of *R* was taller (1.98m) and could easily reach the compartments without the ladder.

Figure 2 – Sequence of events (**1 - 10**) in the FB identity condition where *A* is led to believe that there is just 1 red backpack when there are 2 red backpacks.

4.2.4.3 True Belief - Identity Condition

In the True Belief Identity condition (TBID), after *A* climbed down from the ladder having placed the red backpack upright in compartment *X*, *A* moved to a table along the left side of the room (side counterbalanced; the table was always on the same side of the room as compartment *X*) where there is a stack of papers that *A* flicked through. In this way, *P* was led to believe that *A* was looking for *P*'s questionnaire in this pile of papers, instead of having to leave the room. After 2-seconds, *R* turned around, moved his head side-to-side while staring at *A*'s red backpack, and exclaimed "Tsk-tsk!" When *R* made this utterance, *A* looked up from flicking through the papers and (with a neutral face) watched while *R* removed *A*'s red backpack from compartment *X*, walked to the other side of the room, stored it in compartment *Y*, and then returned to store his (*R*'s) red backpack in compartment *X*. In this way, the TB sequence was identical to the FBID condition, except that *A* tracked what *R* was doing throughout. At no time did *A* and *R* ever say anything to one another, and *R* made sure never to look at *A*.

Suddenly, *A* realized that s/he was receiving an incoming phone call. *A* exclaimed, "Oh!" and then turned away from the scene to take the call. *A* recites the same phone conversation here, as s/he does in the FB conditions. While *A* was taking the phone call, *R* picked up the step ladder and walked out of the room carrying it. The rest of the procedure followed the FB conditions exactly.

4.2.4.4 True Belief - Location Condition

The True Belief Location condition (TBLoc) followed the same procedure as the TBID condition with the exception of the following differences. First, *R* replaced *A*'s red backpack with a black backpack instead. Second, when *A* stored the red backpack in compartment *X*, s/he stored it sideways so that it could not be seen when the compartment flap was closed. As in the FBLoc condition, when *R* turned around, he first lifted the flap so that the backpack can be seen before shaking his head side-to-side and uttered, "Tsk-tsk!" The remainder of the procedure was the same as the others.

4.2.4.5 Participant Responses to A's request for help

Figure 3 depicted the two possible final outcomes of *A*'s request for help. However, prior to this point, there were many possible reactions that *P* could have produced to the helping prompt, and all actors who played the role of *A* were trained to respond appropriately. If *P* did not respond, or said “what?” or something to a similar effect, *A* was instructed to repeat the question: “Can you help me?” If *P* replied that they were too short to reach, then *A* would suggest that they use *P*’s chair to reach it (which was purposefully the only chair in the lab room). If *P* went to where *A*’s actual backpack was (compartment *Y;* Figure 3B), *A* would wait until *P* had opened compartment *Y* and, upon seeing the backpack, would say “Oh!” in surprise *A* would then thank *P* for their help. In the FBLoc condition if *P* went to compartment *X* (Figure 3A) *A* would respond with surprise upon seeing a different looking backpack (“Oh, that’s not my backpack!”) but in the FBID condition, *A* would accept the backpack in compartment *X* as if it was their own.

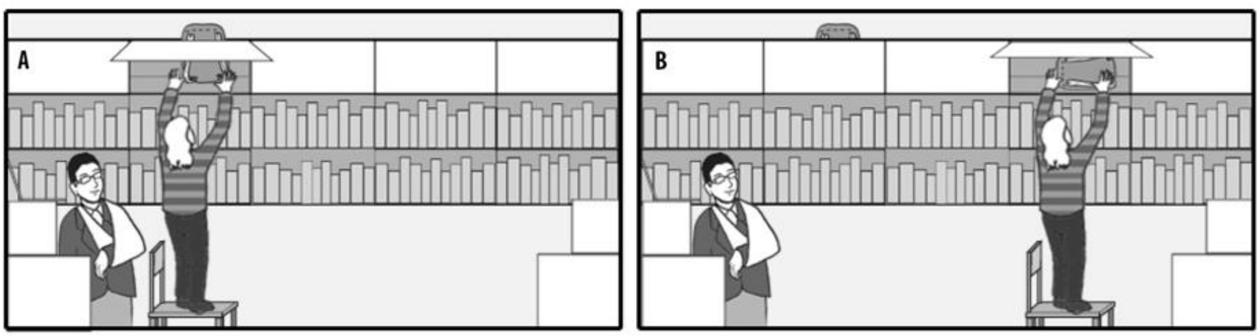


Figure 3 – Correct or incorrect helping actions, during the FBID condition. Since *A* always requests help from beneath compartment *X*, (**A**) depicts *P* incorrectly helping (retrieving *R*’s backpack) and (**B**) depicts *P* correctly helping.

After helping to retrieve a backpack for *A*, *P* was given a questionnaire on emotion recognition (Reading the Mind in the Eyes Task, abbreviated hereafter as RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). The RMET provided an independent check for whether participants across groups were equivalent in their affective empathy towards others. *A* said that s/he would wait outside the room until *P* finished. All participants were fully debriefed via E-mail once all experiments were completed. Delayed debriefing was necessary to avoid contaminating the spontaneous helping responses of potential participants from the university subject pool.

4.2.5 Data Analysis

Videos were recorded from two hidden cameras (Panasonic HC-V770M video cameras). One was positioned in the center of a bookshelf (1.93 meters from the front of the desk) and was used to capture eye movements. The other was positioned on a table behind *P* (1.3 meters behind the desk) which was used to code their body movements, and to synchronize the events corresponding to *P*'s eye gaze. The direction of *P*'s eye movements was assessed by a coder (the primary investigator) on a frame-by-frame basis (30 frames per second) using Adobe® Premier Pro® software. The videos were synchronized using Apple iMovie®. The videos were edited so that the back-camera's footage played on the bottom half of a split screen, and the footage facing *P*'s eyes played on the top half of the screen. In this way, the coder could use the back-camera footage to determine when the critical time windows should begin and end. Once the time windows were established, the coder zoomed in on the upper half of the screen for eye gaze coding (Premier Pro® zoom level = 400%). The critical time windows were individually set by the primary coder prior to scoring, using the following criteria. The time windows began from the first frame wherein *A*'s body was oriented facing compartment *X* (following *A*'s phone call) and ended when *A* completed the first jump. The end of the jump was operationalized as the first frame where *A*'s reaching hand fell below the second shelf. To assess inter-observer agreement, eye movements were scored by a second coder who was trained to distinguish looks to compartment *Y*. A look to compartment *Y* was coded when *P*'s eyes were directed upward and to the side which held *A*'s actual backpack. Figure 4 shows an example of a look to compartment *Y as A* is jumping to reach the bag in compartment *X*. The side where *A*'s actual backpack was located is not confounded by other perceptually stimulating events. The second coder watched the films projected onto a screen from an overhead projector. The second coder also actively raised questions about ambiguous frames. All frames for all 96 participants (total frames = 8,306) were coded by the second coder, and inter-observer agreement was high at 97.82%. All disagreements were resolved through discussion with a third coder.

Figure 4 – An example of the synchronized footage showing *P* looking to compartment *Y* while *A* jumps to reach compartment *X*.

4.3 Results

4.3.1 A Forecast of the order of presentation for Experimental Results

First, a preliminary analysis section presented the results from the RMET, reported any relevant differences found with respect to gender, age, and ethnicity, and compared the critical time window within which gaze behavior was coded and analyzed. The next section presented an analysis of the average duration of looking to compartment *Y* during the critical time window. One recent investigation into typical strategies used to study implicit FB reasoning, suggests that using looking time may be a more reliable measure for testing adults, especially when the events in a task bias a participant to look in one direction over the other (Burnside et al., 2017). Notwithstanding the fact that the present study differs from the AL paradigm in several respects, Burnside and colleagues' (2017) advice was nonetheless taken onboard and duration of looking to compartment *Y* was chosen to measure and compare adults' belief tracking. In the next section, the critical time window is broken down into four sub-windows and average durations of looking to compartment *Y* are compared, per sub-window. This was done in order to investigate where during the events in the critical time window, looking may have differed. The next section provided an analysis of participants' movements across four distinct stages of their helping action following *A*'s request for help. Finally, correlational analyses were run in order to determine whether or not looking to compartment *Y* connected with any of the four stages of action.

4.3.2 Preliminary analyses

Between conditions, participants were equivalent on the RMET (p = 0.821). All participants scored within 3 standard deviations of the mean (FBID: M = 26.92, sd = 4.09; FBLoc: M = 26, sd = 4.12; TB-ID: M = 26.54, sd = 4.3; TBLoc: M = 26.54, sd = 4.11). Thus, it could be assumed that participants in some conditions were not more empathically minded than those in other conditions.

Participants' eye movements were coded on a frame by-frame-analysis during the critical time window. Results revealed no effects of gender, age, or ethnicity on the duration of looking to compartment *Y*. Because the task was performed live, the length of the time window varied (although the scene was regularly rehearsed prior to data collection in order to reduce variation as much as possible). The time window started after *A* had finished with the phone call and as soon as *A* had oriented his/her body to face compartment *X*. The time window ended when *A*'s reaching hand fell past the middle shelf as *A* lands his/her first jump. Overall, the average time window lasted 2,659.03 ms (std. deviation = 574.81).

Time window durations were submitted to a 2 (scenario: location or identity) by 2 (belief: FB or TB) by 2 (Assistant: male or female) ANOVA. Analyses revealed a significant main effect of Assistant [$F(1, 88) = 37.71$, $p < .001$, $\eta^2 = .3$], and a significant interaction between scenario and belief [$F(1, 88) = 13.73$, $p < .001$, $\eta^2 = .135$]. In regard to the effect of Assistant, the male assistant character performed the actions faster (M = 2631.94 ms, SE = 58.06) than the female assistant character (M = 3136.11 ms, SE = 58.06). Post-hoc independent samples *t*-tests confirmed that time windows were significantly shorter during FBID (M = 2606.94, SE = 85.90) than the FBLoc (M = 3031.94, SE = 108.56) condition, $t(46) = -3.07$, $p = .004$, but the time windows in the TBID (M = 3040. 28, SE = 81.94) and TBLoc (M = 2856.94, SE = 108.28) conditions did not significantly differ, $t(46) = 1.35$, $p = .184$.

To determine whether the length of time windows was a predictor of the duration of looking toward compartment *Y*, Kendall's tau correlation coefficient was calculated. Results showed that length of time window was not a significant predictor of the duration of looking toward compartment *Y*, $r_\tau = .14$, $p = .062$. In addition, when separate analyses were run on the data from either assistant, it was found that length of time windows did not predict the duration of looks toward compartment *Y* (male Assistant: $r_\tau = .19$, $p = .076$; female assistant: $r_\tau = .09$, $p = .416$). Moreover, the length of the time window was not correlated with *whether* participants showed a look to compartment *Y* ($r_\tau = .14$, $p = .099$), meaning that shorter time windows did not restrict opportunities for participants to show such a look. Consequently, data were collapsed across male and female actors playing the role of *A*.

4.3.3 Looking to compartment Y

The data on duration of looking to compartment *Y* violated the assumptions of normality (shapiro-wilks = .724, $p < .001$) because a high frequency of zero values was found. Therefore, non-parametric tests were used. A Kruskal-Wallis test was used in place of an ANOVA, with four groups (FBID, FBLoc, TBID, TBLoc). In place of an independent samples t-test, post-hoc comparisons were tested using the Mann Whitney *U* test, which is commonly used to test the likelihood that non-normal data from two independent samples come from the same population and has been used in recent investigations into helping behaviors (Buttelmann et al., 2015; Plötner et al., 2015), as well as studies measuring fixation durations (Wellman et al., 2016; Johannson et al., 2001). These tests will be used in this way to compare non-normal duration data for all experimental chapters in the present thesis.

The Kruskal-Wallis test found that the average total duration of looking to compartment *Y* was not the same across conditions [$\chi^2(3) = 8.134$, $p = .043$] (*Mean ranks:* FBID = 40.96; FBLoc = 58.15; TBID = 41.79; TBLoc = 53.10).

Separate post-hoc Mann Whitney *U* tests revealed that significantly longer looking to compartment *Y* was observed in the FBLoc condition (*Mean rank* = 28.85) compared to the FBID condition (*Mean rank* = 21.71) ($U_{(n1=n2=24)} = 392.5$, exact 1-tailed $p = .018$, $r = .34$). The duration of looks to compartment *Y* did not significantly differ between the TBLoc (*Mean rank* = 27.29) and TBID (*Mean rank* = 21.71) conditions ($U_{(n1=n2=24)} = 221$, $p = .062$, $r = .23$) Average duration of looking to compartment *Y* between conditions is shown in Figure 5. See table 1 for the means, standard deviations, and 95% confidence intervals for each condition.

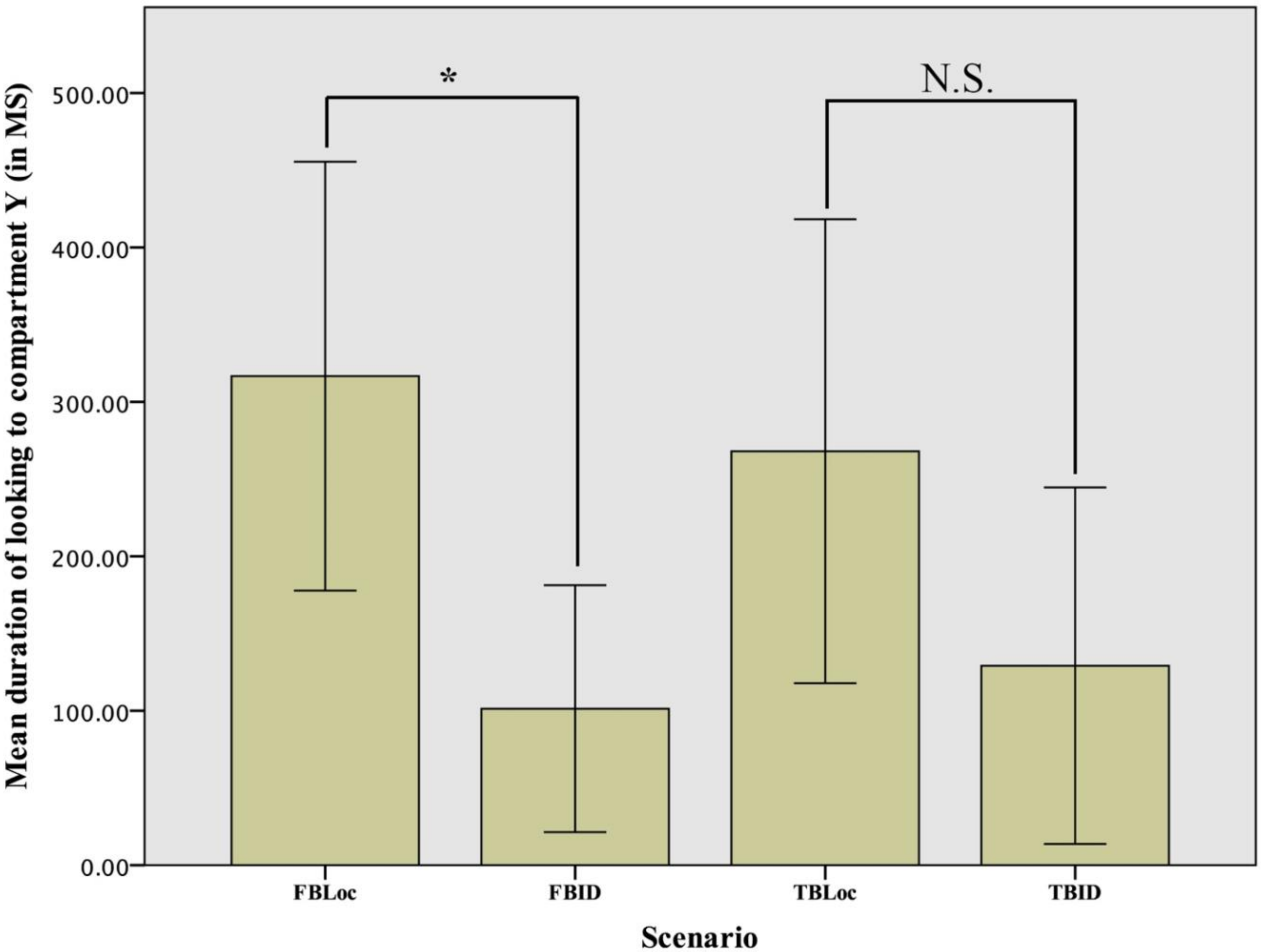


Figure 5 - Mean duration of looking to compartment *Y* (in milliseconds) during the critical time window. * indicates $p < .05$. N.S. indicates no significant difference or $p > .05$.

4.3.4 Durations of looking during four sub-windows

In order to investigate how occurrences of looking to compartment *Y* might have differed over the course of the critical time window, the window was divided into four 'sub-windows' and average durations of looking to compartment *Y* during each sub-window were analyzed. In doing so, it was possible to detect when durations of looking to compartment *Y* differed as *A*'s actions made his or her goal-directed outcome (to retrieve the backpack s/he originally left in compartment *X*) increasingly clear. Until the point where *A* turned to face compartment *X*, it may not be necessarily clear to *P* that *A* was about to attempt to retrieve the backpack. However, the goal may be clearest when s/he walked to the compartment, prepared to jump, leaped into the air, and then finished the jump.

The first sub-time window started when *A* oriented his or her body away from the calendar and toward compartment *X*. This first sub-window ended when *A* finishes approaching the compartment (M = 1827.08 ms, sd = 484.53). The second sub-window started on the next frame and then ended when *A*'s knees were bent to jump (M = 394.44 ms, sd = 166.68). The third sub-window started on the next frame and ended as *A* reached the top of the jump (M = 295.83 ms, sd = 33.25). Finally, the fourth sub-window began on the next frame and ended when *A* finished the jump and his or her reaching arm dropped below the middle shelf (M = 370.14 ms, sd = 105.91).

The average durations of looking to compartment *Y*, in the four groups (FBLoc, FBID, TBLoc, and TBID) were compared using an independent samples Kruskal-Wallis analysis of variance, for each sub-window. Results revealed that looking durations did not significantly differ during sub-window 1 [$\chi^2(3) = 6.507$, $p = .089$] (*Mean ranks*: FBID = 41.40, FBLoc = 56.13, TBID = 43.10, TBLoc = 53.38) or during sub-window 2 [$\chi^2(3) = 6.516$, $p = .089$] (*Mean ranks*: FBID = 47.73, FBLoc = 54.23, TBID = 44, TBLoc = 48.04) or during sub-window 4 [$\chi^2(3) = 1.086$, $p = .780$] (*Mean ranks*: FBID = 47.63, FBLoc = 51.21, TBID = 47.79, TBLoc = 47.63). Durations of looking to compartment *Y* were significantly different during the third sub-window [$\chi^2(3) = 9.128$, $p = .028$] (*Mean ranks*: FBID = 46, FBLoc = 54.06, TBID = 46, TBLoc = 47.94). Follow-up Mann-Whitney tests were conducted to explore the difference in looking to compartment Y during sub-window 3. Results revealed that durations of looking to compartment *Y* were significantly greater in the FBLoc condition (M = 31.944 ms, sd = 88.18) compared to the FBID condition (M = 0 ms, sd = 0; $U_{(n1=n2=24)} = 240$, $p = .039$, $r = .30$). Duration of looking to compartment *Y* did not significantly differ between TBLoc (M = 2.778 ms, sd = 13.608) and TBID (M = 0 ms, sd = 0) conditions during this sub-window ($U_{(n1=n2=24)} = 276$, $p = .317$, $r = .15$). See Table 1 for the means, standard deviation and 95% confidence intervals for the above-mentioned analyses. See Figure 6 for the average durations during these sub-windows and per conditions.

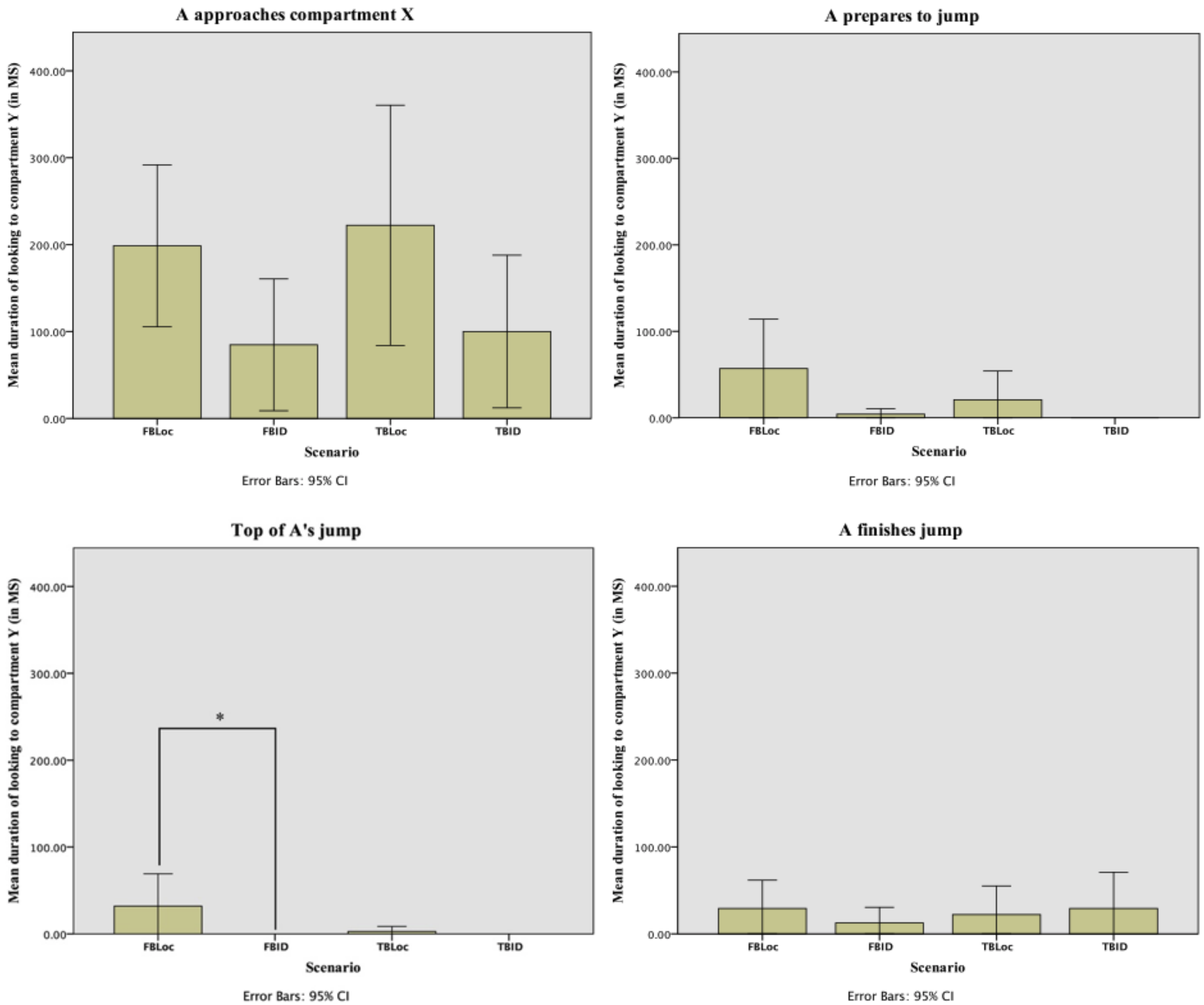


Figure 6 - Mean duration of looking to compartment *Y* (in milliseconds) across four distinct critical sub-windows. In the first sub-window, *A* approaches compartment *X*. In the second sub-window, A bends his or her knees in preparation to jump. The third sub-window ends when *A* reaches the top of his or her jump. The fourth sub-window ends when A has landed his or her jump.

* denotes p <.05.

Table 1 - Mean, standard deviation and 95% Confidence Intervals for duration of looking to compartment *X* during each of four sub-windows.

Time Window	Scenario	Mean	Std. dev.	Lower Bound	Upper Bound
Overall Critical Time Window	FBLoc	316.67	328.81	177.82	455.51
	FBID	101.39	189.45	21.39	181.39
	TBLoc	268.06	355.70	117.86	418.25
	TBID	129.17	273.17	13.82	244.51
Sub-Window 1: *A* approaches *X*	FBLoc	198.61	220.78	105.39	291.84
	FBID	84.72	179.63	8.87	160.58
	TBLoc	222.22	327.29	84.02	360.42
	TBID	100.00	208.05	12.15	187.85
Sub-Window 2: *A* prepares to jump	FBLoc	56.94	134.95	0	113.93
	FBID	4.17	14.95	0	10.48
	TBLoc	20.83	78.52	0	53.99
	TBID	0	0	0	0
Sub-Window 3: Top of *A*'s jump	FBLoc	31.94	88.18	0	69.18
	FBID	0	0	0	0
	TBLoc	2.78	13.61	0	8.52
	TBID	0	0	0	0
Sub-Window 4: *A* finishes jump	FBLoc	29.17	76.97	0	61.67
	FBID	12.5	42.63	0	30.5
	TBLoc	22.22	77.81	0	55.08
	TBID	29.17	98.94	0	70.94

4.3.5 Motor responses - Coding Criteria

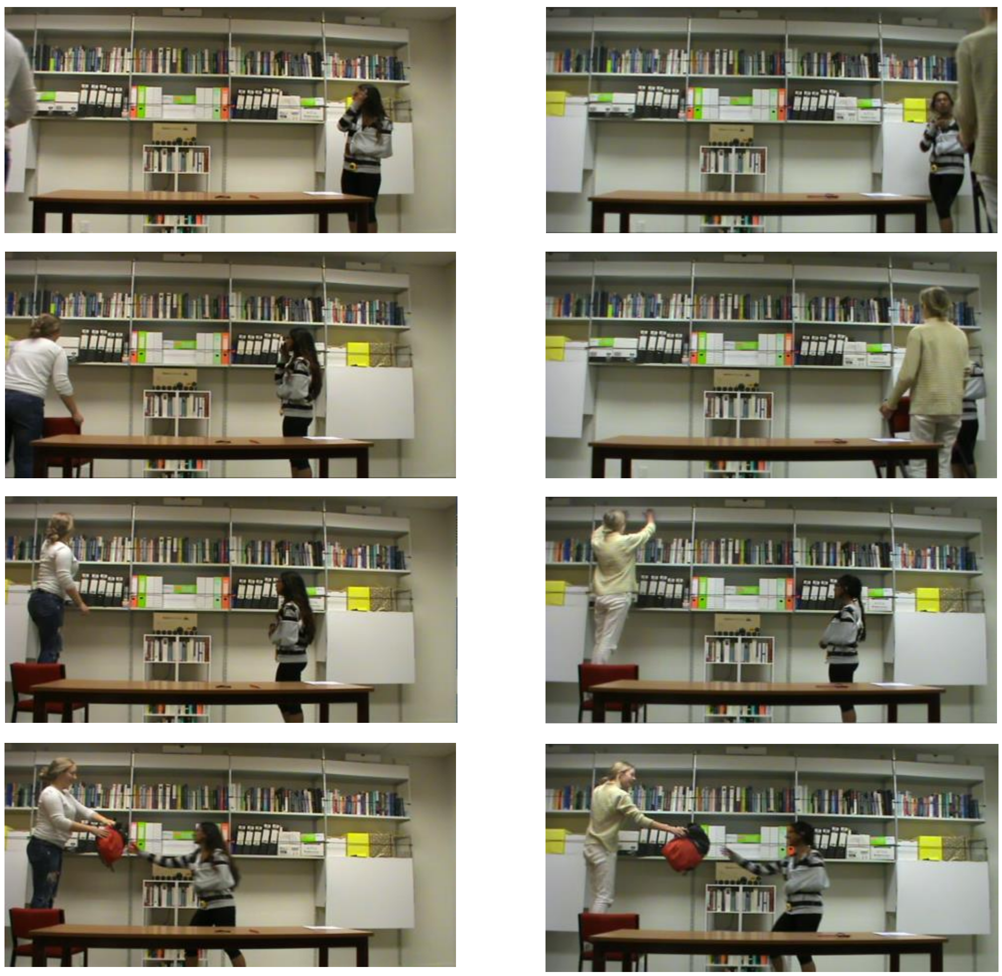
In response to *A*'s request for help, the course of individual participants' motor output in each of the different conditions was divided into four stages of action: swerving (stage-1), approaching (stage-2), reaching (stage-3), and retrieving (stage-4) the correct or incorrect backpack. In each stage, *P* was scored based on whether his or her actions were directed to the side where *A* was requesting help (compartment *X* side) or to the side where *R* had stored *A*'s backpack (Compartment *Y* side). Each stage was coded once *P* passed each of the four thresholds for the first time.

For 'swerving', the corner of the desk *P*'s body passed first was coded. If *P*'s body only partially passed one corner of the desk before changing side and moving to the other corner, the coding was based on the corner the participant's body passed *completely*. For example, if *P* stood up, turned toward the compartment *X* side to pick up their chair, and then carried it around the compartment *Y* side, this was coded as a swerve to the compartment *Y* side. However, if *P* got up and walked to the compartment *X* side, such that their body completely passed the corner of the table, and then he or she returned to get the chair before walking around the table on the compartment *Y* side, this was coded as a swerve to the compartment *X* side.

For stage-2, 'approaching' was coded based on the compartment that *P* first walked toward following their swerve. Specifically, stage-2 was judged as the point when *P* stopped walking from desk to a compartment. It was then determined which compartment *P* was standing in front of. If *P* began to approach compartment *X* but changed direction and approached compartment *Y* and stopped, this was coded as an approach to compartment *Y*.

For stage-3, 'reaching' was coded as which compartment *P* touched first. A 'reach' was coded based on *P* having touched (or, if they couldn't reach, fully extending their arm) one of the compartments. If *P*'s arm began to rise but did not fully extend, this was not coded as a reach.

Finally, stage-4 was coded as being which backpack *P* ultimately retrieved for *A*. In the identity conditions, there are two red backpacks that *P* could choose between, so in this condition, it is possible to give *A* the wrong backpack and *A* was therefore instructed to accept it as if it was their own. However, in the location conditions, there are two different colored backpacks. If *P* tries to hand *A* the black backpack instead of a red backpack, the scenario would logically result in *A* not taking the backpack and instead saying: "Oh! That's not my backpack!" This fourth stage of action is included in order to show whether participants attempted to give *A* the wrong backpack when possible. However, in all other conditions this is not possible.



A **B**

Figure 7 - Two participants' stages of action. A value of "0" is scored for a stage of action toward compartment *X* and a value of "1" for a stage of action toward compartment *Y*. Participant **A** (left) would be scored a 1 for all action thresholds (1, 1, 1, 1) and participant **B** (right) would be scored as making incorrect swerving and approaching actions and would be scored as (0, 0, 1, 1).

4.3.6. Motor Responses - Comparisons

Separate hierarchical log-linear analyses with backwards elimination revealed that scenario (ID or Loc) as well as belief (FB or TB) had significant effects on stage-1 action (p = 0.018 and p < 0.001) and stage-4 action (p = 0.019 and p < 0.001), but only belief was relevant to stage-2 and stage-3 actions (ps < 0.001). To recap, the prediction was that participants' earlier body movements would differ between the FB conditions, but not between TB conditions. In the FB conditions participants were predicted to move toward compartment *Y* in the later stages of action. However, during the early stages of action (i.e., stage-1), participants' in the FBID condition should show more swerving to compartment *X*, compared to those in the FBLoc condition. For a schematic visualization of all participants' movement patterns across the four stages-of-action, see Figure 8.

For stage-1 action, in the FBID condition, 63% of participants swerved toward compartment *X* (one-tailed binomial p = .154). In the FBLoc condition, 75% of participants swerved toward compartment *Y* (one-tailed binomial p = .012) [$\chi^2(1) = 6.86$, p = 0.009]. In the TBID and TBLoc conditions, participants were more likely to swerve toward compartment *X* (83% and 88%, respectively) [$\chi^2(1) = .167$, p = 1]. For stage-2 action (i.e., approaching), 63% of participants in the FBLoc condition and 54% of participants in the FBID condition approached compartment *Y* [$\chi^2(1) = 1.422$, p = .371]. For stage-3 action (i.e., reaching), more participants reached toward compartment *Y* in both the FBLoc (67%) and FBID condition (63%; one-tailed binomial ps > .05) [$\chi^2(1) = .375$, p = .380]. Regarding stage-4, the majority of participants in FBID (75%; one-tailed binomial p = 0.011), and all participants in FBLoc (100%) correctly gave *A* the backpack from compartment *Y*. Although performance was significantly above chance, a small but significant number of participants (25%) handed the incorrect backpack to *A* in the FBID condition [$\chi^2(1, 48) = 6.86$, p = 0.009].

For stages 2, 3, and 4 actions, 100% of participants in the TBID and TBLoc conditions approached, reached and retrieved the backpack from compartment *X*.

Figure 8 - Schematic representation of *P*'s (*N* = 96) stages of action between conditions: (**A**) FBID, (**B**) FBLoc, (**C**), TBID, and (**D**) TBLoc. Helping action was divided into four stages: (1) swerving around the center desk, (2) approaching compartment *X* or *Y*, (3) reaching toward compartment *X* or *Y*, and (4) retrieving the bag in either compartment *X* or *Y* for *A*. Dotted lines represent thresholds for each stage of action.

4.3.7 Relationships between gaze and Motor Responses

The final analysis investigates the possibility that signature limits on efficient belief tracking can be detected in early full-body movements when helping is broken down into distinct stages. More specifically, it is predicted that participants who show looks to compartment *Y* during the critical time window will be more likely to produce early directional movements toward that compartment. To begin, it is useful to ask how many participants in each condition showed at least one look to compartment *Y* during the critical time window.

In the FBID condition, looks to compartment *Y* were in 8 participants (33.33%) during the critical time window. In the FBLoc condition, looks to compartment *Y* were detected for 14 participants (58.33%). In the TBID condition, looks to compartment *Y* were detected for 8 participants (33.33%). And, in the TBLoc condition 12 (50%) participants showed a look to compartment *Y*. Chi-squared analyses revealed that participants were equally likely to show a look to compartment *Y* in the TBID condition compared to the TBLoc condition [$\chi^2(1) = 1.371$, exact 1-sided $p = .190$] as well as when comparing the FBID to the FBLoc condition [$\chi^2(1) = 3.021$, exact 1-sided $p = .073$]. Although most participants did not show any looking toward compartment *Y*, the following analysis asks whether those who *did* were more likely to direct their early actions (e.g., swerving around their desk) toward compartment *Y* than those who did not. A two-systems account would suspect that a significant correlation will be found between looking and early stages of action when the efficient mindreading system is able to detect that the agent has a FB (the FBLoc condition). The clearest prediction concerns how an efficient mindreading system should direct early action in a FBLoc condition; it is less clear whether a two-systems account would predict that any correlation would be found in the FBID condition and the TB conditions. According to the two-systems account, the minimal mindreading system should incorrectly treat *A* as having a TB in the FBID scenario. On the one hand, perhaps the minimal mindreading system does not output to action during TB scenarios in the same way as it does when it detects a FB. If so, there may not be any correlation between looking to compartment *Y* and actions, in the FBID or TB scenarios. Conversely, if minimal mindreading does output to motor processes in both TB and FB scenarios, then there may be an even stronger correlation found in the TB and FBID scenarios than in the FBLoc scenario, as looking and actions are directed away from compartment *Y*. Notwithstanding this uncertainty, there is one clear prediction: *all* three scenarios (TBID, TBLoc and FBID) should either reveal strong positive correlations or no correlations.

Results were in line with these predictions. It was found that participants who produced a look to compartment *Y* also tended to swerve (stage-1 action) toward compartment *Y* [$\chi^2(1) = 11.2$, $p < .001$] but this correlation was only found in the FBLoc condition. For FBID and TB conditions, looking toward compartment *Y* was not associated with any stages of action (*p*s > .05). Furthermore, in the FBLoc condition, looking to compartment *Y* was not associated with stages 2 (approaching), 3 (reaching) and 4 (retrieving) of action (*p*s >.05).

4.3.8 Secondary analysis of Experiment 1 with participants excluded for giving the wrong backpack

As noted in the previous section, some of the participants in the FBID condition handed the wrong backpack to *A*. Since this option is not available to participants in the FBLoc condition (as *A* was instructed to comment on the different looking backpack) or the TB conditions (as *A* was present while the backpack was switched), it is unclear what these participants are doing in the FBID condition. Further checks of the videos confirmed that all participants watched the backpacks switch. It seems unlikely that these participants' failure was caused by a failure to attend to critical information during the task. In any case, secondary analyses were conducted with those participants who handed *A* the wrong backpack ($n = 5$) excluded.

4.3.8.1 Looking to Compartment Y

An independent samples Kruskal-Wallis test (with FBID, FBLoc, TBID and TBLoc as between-subjects factors) revealed that total durations of looks to compartment *Y* were significantly different even when the 5 participants were excluded ($\chi^2(3) = 9.186$, $p = .027$). A Mann Whitney U test revealed that, for FBLoc (*Mean rank* = 28.85) and FBID (*Mean rank* = 20.15), total durations of looking toward compartment *Y* were significantly different ($U(\text{n1=n2=n4}) = 183.5$, exact 1-tailed $p = .009$, $r = .34$).

4.3.8.2 Body Movements

In the original ($n = 24$) FBID sample, 63% of participants swerved toward compartment

X, a percentage that did not differ from chance (one-tailed binomial p = .154). After removing the 5 participants who provided *A* with the wrong backpack, swerving remained at chance level (52% toward compartment *X*; one-tailed binomial p = .5). Furthermore, when n = 19, a chi-squared analysis still showed that looking to compartment *Y* was not predictive of swerving direction [$\chi^2(1) = 3.316$, $p = .091$]. This was also found with respect to later stages of action (all ps $\geq .373$).

4.4 Discussion

Several hypotheses were proposed in Experiment 1. First, adults should show significantly more looking toward compartment *Y* in the FBLoc condition compared to the FBID condition. Second, adults should be more likely to make initial directional body movements (swerving) toward compartment *Y* in the FBLoc condition compared to the FBID condition. These differences should not be found in matched TB conditions. Finally, if the output of an efficient mindreading system influences not only eye movements but motor representations, then early full-body movements should be predicted by when providing help.

The experiment uncovered three main findings. First, adults showed significantly longer looking to compartment *Y* in the FBLoc condition than in the FBID condition. This finding is consolidated by the TB conditions, where adults' looking to compartment *Y* did not significantly differ between identity and location conditions. Second, when adults helped an agent who had a FB about the identity of an object, their early stages of action (swerving) were directed toward compartment *X*; whereas, when adults helped an agent with a FB about the location of an object, their swerves were directed toward compartment *Y*. Third, a significant correlation between looking to compartment *Y* and swerving. Finally, as predicted, this correlation was only detected in the FBLoc condition, and this correlation diminished below significance level during later stages of action.

4.4.1 Tapping into signature limits: A two-systems explanation

The effect of scenario (Location vs. Identity) on the duration of adults' looking to compartment *Y* fits with reports of differential AL in children and adults (Low & Watts, 2013; Low et al., 2014, Wang, Hadi & Low, 2015) and adds to the FB literature by showing that, in an active helping task, looking measures can tap into signature limits of the efficient mindreading system. The results bear on a key challenge for the two-systems theory of mindreading. Until now, the two-systems theory lacked an account specifying which actions (if any) are influenced by efficient mindreading. The present findings suggest that early stages of action performed in response to another's FB may be influenced by efficient belief-tracking processes, whereas later stages of the seemingly continuous action stream seem to be coordinated with the help of less automatic more flexible belief-tracking processes. This opens up a new avenue through which to test the output from efficient belief tracking: testing for correlations between eye movements and early full-body movements in active helping tasks. Furthermore, these results suggest that in circumstances where a minimal mindreading system is able to detect that an agent is about to act on a FB, the motor system directs the helper toward the location housing the object that the agent last registered in the location the agent is gesticulating toward. By contrast, when a minimal mindreading system is unable to detect that the agent is acting on a FB (because, say, that agent's FB is about the identity of an object), the motor system directs the helper toward the location where the agent is gesticulating, as it would when the agent has a TB.

Here, it is prudent to adopt a cautionary tone: the present results are the first of their kind in many regards and should be treated with the same skepticism that all exploratory research should be prepared to endure. The present study is the first active helping FB task to measure gaze behavior, or to divide a helping behavior into distinct stages of action, or to use adult participants. Moreover, given the recent replication crisis from a wide range of psychological research (Open Science Collaboration, 2015) and specific non-replications in the mindreading literature specifically with respects to looking paradigms (Kulke et al., 2018; Burnside et al., 2017) and helping tasks (Crivello & Poulin-Dubois, 2017; Priewasser et al., 2017) it is important to stress the value of externally replicating these findings. It is also relevant to note that the sample sizes for Experiment 1 are relatively small at $n = 24$. Since the believability of the study was a high priority, each participant could only take part in a single trial. It is therefore even more imperative that these results be confirmed in a sample size that achieves higher statistical power. While the present findings provide preliminary evidence that can be explained by a 'two-systems' account, these findings do not represent a decisive test of the two-systems account over competing accounts. These accounts are expanded in the following section.

4.4.2 Alternative Interpretations: representations, rules, and reasons

The two-systems account offers something of a ‘middle of the road’ between domain-specific modular accounts and domain-general statistical learning models. As such, the two-systems account is challenged by theories of both higher and lower level reasoning. This section represents some of the leading alternative accounts to the two-systems account and considers whether and how such theories could explain the pattern of results found in Experiment 1.

The main opponent to the two-systems account is the one-system account, or early mindreading account (EMA). According to this account, a representational mechanism facilitates the understanding of other’s goals, knowledge, and FBs, from early infancy. The system responsible for this capacity improves over time, however, its core “skeletal causal structure” is present from birth and is fully operational by around the second year of life (Baillargeon, Scott & He, 2010). The EMA has argued that infants are prepared not only to predict the actions of an agent who has an identity FB (Scott & Baillargeon, 2009) but even to predict how one agent should act in order to implant an identity FB into the mind of another (Scott, Richman & Baillargeon, 2015). If infants have this capacity, then the EMA should expect that adults have no difficulty with either the location or identity version of this task, and thus, looking durations would not significantly differ in the FBLoc or FBID condition. With respect to correlations, stages of actions also should not correlate with eye movements more in the FBLoc scenario compared to the FBID scenario. The results from Experiment 1 do not support these predictions.

One explanation that EMA researchers often raise in response to evidence for signature limits regarding identity FBs, is that the identity task carried with it extraneous performance demands that were not present in the location task. In response to this concern, the present study designed the location and identity scenarios so that they match as closely as possible. The only differences between the location and identity conditions are the color, and position of the backpack when placed in compartment *X*. An argument could conceivably be made that the presence of the red backpack in compartment *X* carried with it a greater inhibition cost to observers, as they would need to ignore the identical backpack and remember that it is not the same one *A* put there. However, given that the present study tests and adult sample and all participants watched the bags switch and that the switch is takes place only a few seconds prior to *A* returning and jumping, it seems safe to expect that adults could manage to hold these two different items in their mind and treat them as such. Therefore, the results from Experiment 1 do not fit with the view that identity FBs and location FBs are equally difficult.

How would the present findings fit with the predictions from a behavior-rule perspective? Such accounts argue that human mindreading is rooted in an early-developing ability to detect statistical regularities in human behavior, and to generalize these behaviors into rules for how people are likely to act in future circumstances. Ruffman (2014) provided one such rule-based explanation for the active helping task by Buttelmann, Carpenter and Tomasello (2009). Infants have learned through experience that people tend to search for objects until the object is retrieved. Since the agent in the FB condition has demonstrated ‘search’ behavior, infants go to the box housing the object. However, when the agent has watched the object move (TB), they recognize that the agent must not be searching for the object, so they help by going to the empty box. A behavior-rule theorist would expect that adults would go toward compartment *Y* in the FB conditions and compartment *X* in the TB conditions.

Other lower-level accounts may offer further challenges to the present results. For example, Heyes (2014) argues that looking time and directional looking studies are subject to alternative explanations based on a stimulus’ perceived novelty. According to this account, any difference in the conditions that may attract an observer’s attention in a particular direction, may be the cause for any differential pattern of looking. In the present FBID condition, a red bag is visible in compartment *X*, but this bag is not seen in the FBLoc condition. For this reason, a perceptual novelty account may predict that participants may have looked less toward compartment *Y* in the FBID condition because they were looking at the red bag in compartment *X*.

It is possible that some hybrid model between behavior-rules and perceptual novelty may account for the present findings. If adults were operating on rules but were also influenced by certain perceptually novel features of the task, like the partially visible backpack, then this could account for the present pattern of results. One further possibility is that an association bias explanation may also explain these results. Perner and Ruffman (2005) demonstrated how an infant could solve Onishi and Baillargeon’s (2005) task by generating lower-level perceptual associations between the agent, an object and its location. However, it is unclear whether these associations are formed by certain object properties such as color and shape. If so, then it may be the case that adults made an association between *A*, the *red* backpack, and compartment *X*. Since *R* places a red backpack in compartment *X* in the FBID condition, an observer could have made an error based on the present state of the world (red backpack in compartment *X*) matching the three-way association that was previously constructed. This could lead to a bias toward compartment *X* in the FBID condition that would not apply to the FBLoc condition. Therefore, the present results do not allow one to rule out lower-level association and perceptual bias accounts for explaining adults' behavior.

Interpreting the meaning behind helping behaviors is not as straightforward as it may appear. A number of conceptual replications of Buttelmann and colleagues' (2009) study have challenged Buttelmann's interpretation of infants' behavior as necessarily driven by belief reasoning (e.g., Allen, 2015; Perner, 2014; Crivello & Poulin-Dubois, 2017; Priewasser, et al., 2017). In one study, Priewasser and colleagues (2017) not only failed to replicate Buttelmann and colleagues' (2009) pattern of results, but in a second experiment, found evidence that infants may have solved the helping task by appealing to the agent's objective reasons for action rather than by making inferences about the agent's subjective mental states. Priewasser and colleagues' findings suggested that infants may not have been thinking about the beliefs of the agent at all but were instead interpreting the agent's actions through a basic form of teleology.

According to the teleological account, infants see events in the world as objective facts that provide reasons for actions to occur and experiences dictate the likelihood that those actions will occur (Perner & Esken, 2015; Perner & Roessler, 2010). Relative to mindreading, teleology is a simpler explanation for infants' competency on FB tasks, because objective reasons are governed by tangible events (e.g., the agent was here while the bag was in compartment *X*), whereas subjective reasons are the propositional contents of mental states (e.g., The agent has a subjective belief that the backpack is in compartment *X,* so the agent is heading to compartment *X* to retrieve it). An observer operating on teleology only needs to possess a subjective view of what is good or bad (i.e., what would lead to a more pleasant state of affairs, or a less pleasant state of affairs). The objective facts about the world lead to an understanding about what would need to happen in order for a more or less pleasant state of affairs to result. For example, in Buttelmann and colleagues (2009) study, an infant could simply have reasoned that the agent wants the toy that she was previously playing with and is therefore searching for that toy when she struggles to open the empty box. To help her, an infant operating on teleology would go to the full box to retrieve the toy, simply by recognizing that the agent wants the toy and therefore has good reasons to open the other box. No belief reasoning was necessary in order to accomplish this.

A teleology account may be able to account for the pattern of looking durations observed in Experiment 1. To the observer, it is an objective fact that the agent stored a red backpack in compartment *X*. It is also an objective fact that the agent was out of the room when this red backpack was moved to compartment *Y*. In the FBLoc condition, the former fact conveys that the agent has good reasons to return to compartment *X*, and the latter fact conveys that *when* the agent returns to compartment *X*, the agent's goal will not be obtained. Thus, a teleological account would expect to see more looking to compartment *Y*, in the FBLoc scenario, compared to the TB cases (where the agent remained in the room during the swap, so wherever the agent goes, s/he must have good reason for going that way). Moreover, a teleological explanation may also explain the behavior observed in the FBID scenario. The color and shape of the bag are also objective facts. In the FBID scenario, a teleologist observer may have reasoned that: 'the agent was present when the red backpack was stored in compartment *X*. And the agent was not present when the backpack was moved. Therefore, the agent has good reason to go to compartment *X* to retrieve his or her red backpack." In the FBLoc scenario, this understanding of events allows the observer to reason that going to compartment *X* will not result in the agent obtaining the goal of their action, because there is no red backpack in compartment *X*. However, in the FBID scenario, an identically shaped, colored and positioned backpack is placed inside compartment *X*. Therefore, it is possible that an observer who is operating on teleology for efficient inferences about others' actions will make a similar error to that which the two-systems account would predict. The teleologist may incorrectly consider the objective facts about the present environment as matching the objective facts about the environment when the agent stored their red backpack. Thus, the looking durations found in Experiment 1 are as explainable by appealing to a teleological framework for interpreting actions in relation to objective reasons, as they are by appealing to the workings of an efficient minimal mindreading system.

In conclusion, the present findings are not able to say conclusively that the observed pattern of behaviors were necessarily the product of signature limits on efficient mindreading.

While the observed pattern of results aligns with the predictions from a two-systems account, it is also equally plausible that a teleological account (Perner, Priewasser & Roessler, 2018) or an even lower-level association biases (Perner & Ruffman, 2005) and perceptual novelty (Heyes, 2014) could explain these data.

4.4.3 Links to Chapter 5

Is it possible to tease apart a two-systems prediction from the predictions made by a teleology or association bias account? One possible way of doing so would be to explore the notion that efficient belief tracking interfaces with motor processes. Of the potential valid explanations for the data from Experiment 1, only the two-systems account has postulated a connection to motor representations. In Experiment 1, suggestive evidence is found for such a connection, as looking to compartment *Y* was correlated with early stages of helping action during the FBLoc condition only. In chapter 5 (Experiment 2), the proposed interface between motor processes and efficient belief tracking is explored. Specifically, Experiment 2 asks whether the connection between gaze behavior and early stages of helping is disrupted when access to motor representations are interrupted.

CHAPTER 5

Experiment 2: Motor Representations for Tracking False Beliefs

5.1 Introduction

5.1.1 Abstract

Recent preliminary evidence suggests that an efficient system for mindreading may be linked to motor representations (Edwards & Low, 2017). This evidence dovetails with recent philosophical conjectures that efficient mindreading defers to motor representations for understanding an agent's goals (Butterfill and Apperly, 2016). In the following experimental Chapter, two Experiments investigated this proposed connection between efficient mindreading and motor representations, by drawing upon earlier research showing that motor representations for action prediction are stalled by restraining the observers' ability to perform the observed actions (Ambrosini, Constantini, & Sinigaglia, 2012). In the present study, adults' hands were tied behind their backs while observing a seemingly naturalistic FB scenario where an agent comes to hold a FB about the location of her or his backpack. It was reasoned that tying adults' hands would produce a corresponding negative effect on gaze behavior (i.e., duration of looking toward compartment *Y*), as well as the direction adults would choose to swerve around their desk immediately after the agent asks for their help. Finally, it was predicted that the previously reported correlation between looking to compartment *Y*, and swerving direction (see Chapter 4), would not be found when participants' hands were tied.

Results were in the predicted direction but did not reach significance (Experiment 2a). To address the possibility that earlier findings may have been undermined by spurious results from the 'hands free' condition of Experiment 1, an internal replication of the 'hands free' FBLoc condition was run with a larger sample size for higher statistical power (n = 48). A new sample of 'hands tied' FBLoc participants were also tested (n = 24) and their performance was compared against the replication 'hands free' participants (Experiment 2b). Results from the new 'hands free' condition successfully replicated the significant correlation (between looking to compartment *Y* and swerve behavior) found in the FBLoc condition from Experiment 1, lending support to the earlier findings. Moreover, Experiment 2b found that those whose hands were free were more likely to show a look to compartment *Y* and tended to look longer toward compartment *Y*, than those whose hands were tied. The results of these experiments provide suggestive evidence that when motor representations are stalled, efficient mindreading no longer

holds sway over early stages of action.

5.1.2 A summary of the findings from Experiment 1

In Experiment 1, the overall duration of adults' looking to compartment *Y* during the critical time window were significantly shorter when observing an agent acting on a FB about an object's identity, as opposed to a FB about an object's location. After the agent requested help, adults' immediate full-body movements (i.e., toward which side of their desk they swerved) tended to be directed toward the compartment *Y*, whereas they swerved at chance when the agent had a FB about the identity of the backpack. In matched TB scenarios, adults were uniform in their tendency to swerve toward compartment *X*, and their looking durations to compartment *Y* did not significantly differ when the agent's TB was about object identity or object location.

These results are in line with the two-systems theory which posits that humans operate on distinct systems for processing how others' actions are informed by their beliefs. The main thrust of empirical research into distinct mindreading processes has focused on testing the core tenet of the theory: that children's and adults' performance on explicit and implicit measures of belief reasoning recommends two distinct systems (e.g., Clements & Perner, 1994; Low & Watts, 2013; Low et al., 2014; Wang, Hadi & Low, 2015; Masangkay et al., 1974). Beyond this, some studies have investigated the automaticity of a minimal mindreading system (Samson et al., 2010; Surtees & Apperly, 2012; Kovács et al., 2014) and others have investigated its degree of encapsulation (Grosse-Weismann et al., 2017; Low, 2010; Schneider et al., 2012). However, little research has been conducted on whether and how a minimal mindreading system informs action. Incidentally, there has been recent speculation that the minimal mindreading system may defer to motor processes in order to make swift predictions of the outcome of an agent's ongoing action (Butterfill & Apperly, 2016). Moreover, recent empirical evidence has emerged that suggests that efficient minimal mindreading informs adults' expectations of others' upcoming action (Edwards & Low, 2017).

In Experiment 1, signature limits on efficient mindreading explains why adults showed less looking to compartment *Y* as agent approaches compartment *X* during the FBID scenario. The notion that a minimal mindreading system influences early actions is further supported by adults' swerving behavior which was correlated with looking to compartment *Y* in the FBLoc scenario and was at chance level in the FBID scenario. Experiment 2 investigates a complementary yet separate claim. If efficient mindreading interfaces with motor processes such that signature limits on efficient belief tracking affect actions, then efficient belief tracking should be affected by stalling an observer's access to their motor representations. This logic is unpacked in the following section.

5.1.3 Connecting motor representations to minimal mindreading

It is generally agreed by most action theorists that motor representations play a fundamental role in action understanding (the details of that role notwithstanding; see section 3.7). In a nutshell, motor representations can represent action outcomes or guide action execution. This is useful for controlling one's own goal-directed actions. But motor representations also facilitate an observer's ability to predict the outcome of others' actions. Butterfill and Apperly (2016) have postulated that motor representations have a role to play in efficient belief tracking. The idea is that efficient belief tracking *implies* efficient goal attribution (see Michael & Christenson, 2016) and that an account of minimal mindreading must therefore be prepared to explain how the goals of others' actions are attributed without relying on sophisticated higher-level cognitive processes. Butterfill and Apperly (2016) consider one means by which goals could be ascribed in a minimal cognitive architecture: by deferring to what motor representations represent.

5.1.4 Exploiting the fragility of motor representations

According to Butterfill and Apperly (2016) efficient belief tracking defers to motor representations for functional goal ascriptions, which enables a minimal mindreading system to predict how an agent will behave. A strong test of this conjecture would be to stall, disrupt or disable what the efficient system is deferring to. Thus, an experimental manipulation which interrupts observers' ability to generate motor representations should forestall goal ascription, and by extension, impair accurate efficient mindreading. Research shows that motor representations can be disrupted in certain ways. They can be stalled using TMS (Constantini et al., 2013) or when participants engage in a different action to that which is being observed (Constantini, Ambrosini & Sinigaglia, 2012). But perhaps the simplest way to stall motor representations, is by restricting an observer's ability to perform the observed actions themselves.

In one study, Ambrosini and colleagues (2012) stalled motor representations by momentarily tying up participants' hands. Adults watched an actor reach toward two potential target objects which required different hand postures to grasp. When adults' hands were unrestrained, the proactivity of their predictive saccades to the target of the reach was benefited by the pre-shaping of the actor's hand on its approach. That is, relative to when the actor's hand was not pre-shaped, in the pre-shaped trials observers were quicker to initiate saccades away from the actor's reaching hand and arrived earlier to the correct target prior to the hand's arrival. Those whose hands were restrained (i.e., tied behind their backs) also were earlier to arrive at the target object when the hand was pre-shaped, however, they were also significantly slower than those whose hands were unrestrained. According to the authors, observers' access to their motor representations for producing hand actions were delayed when they were in a position that restrained them from being able to perform the action themselves. In Experiment 2, a similar experimental restraint is applied to adults while observing the active helping task from Experiment 1.

The aim of Experiment 2 was to stall adults' access to specific motor representations which Butterfill and Apperly (2016) suggest unlocks efficient goal attribution for efficient mindreading. According to Ambrosini, Constantini and Sinigaglia (2012), tying up participants' hands stalls their access to motor representations for hand actions because the recruitment of motor representations is highly somatotopic. Therefore, in order to effectively stall motor representations for a certain action, it is important that the type of restraint is specific to the action. In the present helping task, in order to open the flap on compartment *X* (and access the backpack within), a 'pincer grip' is needed which involved the thumb and index finger. To maximize the potential of stalling motor representations that were specific to opening the flap, participants were asked to hold a pencil in their hands using a 'pincer grip' while their hands were tied behind their back (see Figure 9). In so doing, Experiment 2 tested the conjecture that a disruption to somatotopically specific motor representations will result in less looking to compartment *Y* compared to those whose hands were free. Specifically, it is hypothesized that adults in a 'hands tied' version of the FBLoc scenario would show a shorter duration of looking to compartment *Y* compared to those in the 'hands free' FBLoc scenario. Finally, it was hypothesized that adults' initial movements (i.e., swerve performance) would be significantly different to their swerve performance observed in the FBLoc 'hands free' condition, and that the connection previously observed between looking to compartment *Y* and swerving, will not be found in the ‘hands tied’ scenario.

5.2 Methods

5.2.1 Participants

24 undergraduate adults (mean age: 19.3; range 17-31; 21 female) participated in exchange for course credit. 83.3% of participants were either of NZ European ethnicity or Maori, and 16.7% indicated ‘other’.

5.2.2 Materials & Procedures

All participants provided informed consent prior to the start of the experiment. The same plain-clothes Asian actors took turns playing the role of the assistant (*A*) and the same European actor wearing a white lab coat played the role of the researcher (*R*). As in Experiment 1, *A* (carrying a red backpack over his or her left shoulder and wearing a sling on his or her left arm) led the participant (*P*) into the laboratory and invited *P* to take a seat at a center table. *R* (wearing headphones) was already in the room when *P* arrived.

After *P* was seated, *A* instructed that the study would consist of two parts: completing a timed block puzzle and filling in a questionnaire. On this occasion, the blocks were on top of a stand that was placed behind *P*’s chair. *P* was instructed to solve the puzzle as quickly as they could with an added challenge: their hands were bound behind their backs using an elastic band. A GoPro® HERO3 Black video camera was attached to a tripod and synced to an Apple® iPhone® 6 (running the GoPro® Quik Application) positioned on the table in front of them. By using the video feedback from the iPhone® display, *P* could manipulate the blocks while their hands were tied behind their backs. Once the puzzle was completed, *A* removed the iPhone®, blocks, and GoPro®.

A then instructed *P*: "*For the next task, I need you to keep your hands behind your back as they are, and to also hold this pencil in both hands using pincer grips* (See Figure 9). *I just need to go run out and fetch the finger pressure sensor which we are going to use for the next part of the experiment*." *A* then feigned hesitation and, looking at their backpack, said, "*Oh! I'll just put this away first*." *A* climbed up the ladder and placed the red bag sideways inside compartment *X* (no part of it remained visible). *A* climbed back down the ladder and exited the room saying "*I'll be right back*" to *P*.

Figure 9 - *P* with hands constrained by elastic band and holding the pencil in a 'pincer grip' as *A* requests help.

The 'finger pressure sensor' was mentioned only to provide a rationale for keeping *P*'s hands tied after the block puzzle task was completed. If *P* asked about this, *A* informed *P* that keeping their hands tied in this position was important to "prime them for an upcoming balancing task". Of course, there was no balancing task, and *A* never finds the 'finger pressure sensor'.

The following sequence o*f* actions was identical to the FBLoc condition of Experiment 1: After a 2-second pause, *R* moved the red backpack to compartment *Y*, stored his black backpack inside compartment *X* and left the room with the ladder. While *R* is leaving, *A* entered the room and rushed toward the calendar while pretending to be taking an important call. After finishing the phone conversation (and *R* has exited), *A* turned and walked forward to the location where the ladder used to be and attempted to jump to reach compartment *X*. Failing to reach the compartment, *A* then turned around to *P* (whose hands are still bound behind their back) and asked for help.

A third hidden camera positioned along wall to *P*'s right confirmed that all participants kept their hands tied and held the pencil with both hands using pincer grips, throughout the procedure. In response to *A*'s request for help, *P* typically stood up and only then asked if it was okay to remove the elastic band and let go of the pencil. *A*'s stock response was, "*Oh! That's alright, you can let that go for the moment*." If *P* attempted to give *A* the red bag, *A* would say "*Thanks for helping me get my bag*." *A* would then look through the bag and say, "*Hmm, that's strange! The sensor is not in here and it wasn't outside either. Okay, I think we'll just have to skip that task and move onto the final part*." *P* was then invited to return to the desk and complete the RMET. Participants were debriefed only after all participants had completed the experiment, to ensure that the participant pool was not contaminated.

5.3 Results

The results of Experiment 2 are presented in the same order as Experiment 1. The order of presentation is as follows. First, a preliminary analysis section presents the results from the RMET, reports any relevant differences found with respects to gender, age, and ethnicity, and compares the critical time window within which gaze behavior was coded and analyzed. The next section presents an analysis of the average duration of looking to compartment *Y* across conditions. In the next section, the critical time window is broken down into four sub-windows and average looking durations to compartment *Y* are compared, per sub-window. The final section provides an analysis of participants' movements across four distinct stages of action. Finally, the number of participants who showed a look to compartment *Y* during the critical time window is reported, and a correlational analysis reports whether or not looking to compartment *Y* connected with any stages of action.

5.3.1 Preliminary Analyses

The performance of hands-tied participants was compared against hands-free (FBLoc) participants from Experiment 1. Participants were equivalent in their performance on the RMET task ($t(46) = .845$, $p = .403$). Participants all scored within 3 standard deviations of the mean (FBLoc 'hands free': $M = 26$, $sd = 4.12$; FBLoc 'hands tied': $M = 24.96$, $sd = 4.42$). Since both conditions are FBLoc conditions, these two conditions are referred to hereafter as either 'hands free' or 'hands tied'. As in Experiment 1, eye movements were coded by the primary investigator on a frame-by-frame (30 frames per second) analysis. A second coder trained to distinguish looks toward compartment *Y* in the videos, observed all frames from all 24 participants in the 'hands tied' condition (total frames = 1,747 frames). The second coder raised questions about any frames that were ambiguous, and inter-observer agreement was 95.363%. All disagreements were resolved through discussion with a third coder.

The same criteria used in Experiment 1 for establishing the critical time window, was used in Experiment 2. The time window extended from the first moment when *A*'s body faces compartment *X*, following *A*'s phone call, and ends when *A*'s reaching arm falls below the middle shelf as s/he lands the first jump (mean duration = 2426.39 ms, *sd* = 399.03). The time window in the 'hands free' condition (*M* = 3031.94 ms, *SE* = 108.56) was significantly longer than that of the 'hands tied' condition (*M* = 2426.39 ms, *SE* = 81.45; $t(46) = 4.462$, $p < .001$). For 'hands tied' participants, a one-way ANOVA revealed that time window was not significantly different when the role of assistant was played by the male or female actor [$F(1, 22) = .047$, $p = .830$].

5.3.2 Looking toward compartment Y

Participants' total duration (in milliseconds) of looking to compartment *Y*, during the critical time window, were compared ('hands free': *M* = 316.67, *sd* = 328.81, 95% CI [177.82, 455.51]; 'hands tied': *M* = 195.83, *sd* = 412.29, 95% CI [21.74, 369.93]). A Mann Whitney *U* test revealed that looking durations significantly differed between the 'hands tied' (*Mean rank* = 20.9) and 'hands free' (*Mean rank* = 28.1) conditions ($U_{(n1=n2=24)} = 201.5$, exact 1-tailed $p = .026$, $r = .28$).

5.3.3 Durations of looking during four sub-windows

The average duration of looking to compartment *Y* was calculated during the same four sub-windows specified in Experiment 1.

The first sub-window extended from the point where *A* was oriented toward compartment *X*, to the point where *A* finished approaching the compartment (*M* = 1454.17 ms, *sd* = 354.65). The second sub-window extended from that point until *A*'s knees were bent in preparation to jump (*M* = 325 ms, *sd* = 151.72). The third sub-window extended from that point until the top of *A*'s jump (*M* = 293.06 ms, *sd* = 25.97). The fourth sub-window extended from the top of *A*'s jump until *A* landed and his or her reaching arm dropped below the middle shelf (*M* = 354.17 ms, *sd* = 80.94).

In order to assess whether and when average durations of looking to compartment *Y* differed between conditions, separate Mann Whitney U tests were conducted for each sub-window. Results revealed that the average amount of looking to compartment *Y* during sub-window 1 was significantly greater for 'hands free' participants (*Mean rank* = 27.71) than 'hands tied' participants (*Mean rank* = 21.29; $U_{(n1=n2=24)} = 211$, exact 1-tailed $p = .037$, $r = .26$). However, during the following three sub-windows, average looking durations did not significantly differ (all *p*'s > .05). See Table 2 for the means, standard deviations and 95% confidence intervals per condition, for each sub-window.

Table 2 - Means, standard deviations and 95% Confidence Intervals for the overall critical time window and the four sub-time windows.

Time Window	**Condition**	**Mean**	**Standard deviation**	**95% CI lower bound**	**95% CI upper bound**
Overall Critical	'hands free'	316.67	328.81	177.82	455.51
Time Window	'hands tied'	195.83	412.29	21.74	369.93
Sub-Window 1:	'hands free'	198.61	220.78	105.39	291.84
A approaches *X*	'hands tied'	136.11	282.32	16.90	255.32
Sub-Window 2:	'hands free'	56.94	134.95	0	113.93
A prepares to jump	'hands tied'	26.39	85.11	0	62.33
Sub-Window 3:	'hands free'	31.94	88.18	0	69.18
Top of *A*'s jump	'hands tied'	11.11	54.43	0	34.10
Sub-Window 4:	'hands free'	29.17	76.97	0	61.67
A finishes jump	'hands tied'	22.22	61.91	0	48.36

5.3.4 Body Movements

Participants' body movements following *A*'s request for help were divided into the same four stages as Experiment 1 (swerving, approaching, reaching and retrieving the backpack for *A*). Figure 10 shows a schematic visualization of *P*s' movements across four stages of action. Participants' swerve direction displayed stronger spatial attraction towards compartment *Y* when hands were free (75%; one-tailed binomial $p = .012$) but swerving was at chance when their hands were tied (50%). For stage-2, participants approached compartment *Y* more in either condition (70% of hands free participants; 63% of hands-tied participants). In stage-3, participants tended to reach for compartment *Y* (71% of hands free participants; 75% of hands tied participants). Regarding stage-4, all participants in either condition correctly gave the appropriate bag to *A* (as mandated by *A*'s instructions). Separate hierarchical log-linear analyses with backwards elimination revealed no significant effect of restriction on any of the body movements (all *p*s > .05).

It was predicted that the previously observed significant correlation between looking to compartment *Y* and swerving actions when participants' hands were free [$\chi^2(1) = 11.2$, $p = .002$], would not be found when their hands had been tied. In the FBLoc 'hands free' condition from Experiment 1, looking to compartment *Y* was detected during the critical time window for 14 (58%) participants. In the 'hands tied' condition, looks to compartment *Y* were detected for 8 (33%) participants during the critical time window. A chi-squared analysis revealed that participants were not significantly more likely to show a look to compartment *Y* when their hands were free rather than tied [$\chi^2(1) = 3.021$, exact 1-tailed $p = .073$].

Finally, it was asked whether tying participants' hands resulted in a disconnect between their looking to compartment *Y* and swerving behavior. A chi-squared analysis revealed that hands-tied participants' looking to compartment *Y* was not correlated with their subsequent swerving actions [$\chi^2(1) = .750$, $p = .333$] or with later stages of action (all *p*s > .05). That is, those who showed a look to compartment *Y* in the 'hands-tied' condition, were no more likely to swerve in that direction. However, these correlations must be interpreted with caution as none of the above effects on participants' body movements were found to be significant.

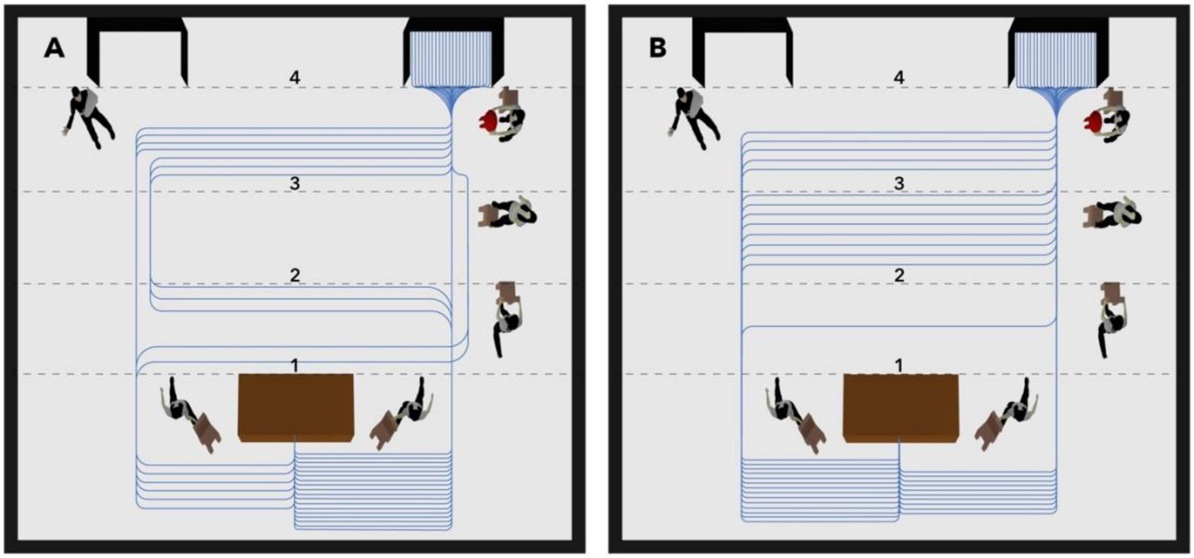


Figure 10 - Schematic representation of *P*'s (*N* = 48) stages of action between conditions: (**A**) 'hands free' and (**B**) 'hands tied'. Note that both conditions are FBLoc scenarios and (**A**) is the same graph as in Figure 8 (**B**).

5.4 Discussion

One limitation of this experiment is that the same data from the FBLoc condition of Experiment 1 was used to compare against the 'hands tied' condition of Experiment 2. Since the 'hands fee' FBLoc condition is predicted to differ from the FBID condition and TB conditions from Experiment 1, as well as from the 'hands tied' condition from Experiment 2, it is possible that these findings all reflected a spurious pattern of behaviors in the FBLoc condition from Experiment 1. Furthermore, a post-hoc power analysis revealed that the present experiment had low statistical power (β = .406). Therefore, the likelihood of making a type-II error is relatively high. To address this possibility, a follow-up experiment was conducted to test the same comparison using new 'hands free' data from a larger sample (*n* = 48) in order to achieve higher statistical power.

5.5 Experiment 2b - 'Hands Tied' and 'Hands Free' Replication

The comparisons made in experiments 1 and 2a involve relatively small samples (*n* = 24). Given that the current state of the mindreading literature is grappling with failures to replicate many earlier and influential findings (see Kulke et al., 2018; Powell et al., 2017; Priewasser et al., 2017; Burnside et al., 2017) the next goal was to replicate the findings from Experiment 2a. One limitation of Experiment 2a is that the comparison bears on the results from the FBLoc 'hands free' condition in Experiment 1. If the results from the original FBLoc condition were the result of a spurious false positive, then this would negate all previous comparisons. Therefore, in Experiment 2b, the FBLoc 'hands free' condition was run again, with a larger sample size (*n* = 48) for higher statistical power. In addition, a new sample (*n* = 24) of adults was recruited in a strict replication of the FBLoc 'hands tied' condition, thereby testing an overall sample of 72 adults.

5.5.1 Participants

72 new participants were randomly assigned to either the FBLoc 'hands free' (*n* = 48) or FBLoc 'hands tied' (*n* = 24) condition. Participants were undergraduate adults (57 females) between 17 and 31 years of age (*M* = 19.04 years). 90.3% of participants were either NZ

European or Maori; 9.7% indicated ‘other’ All participants provided informed consent prior to the start of the experiment.

5.5.2 Materials & Procedure

There were two changes to the materials involved. First, the RMET test was removed because it did not link to any useful data, and it also took a long time for participants to complete. In order to run participants through the study more efficiently, the RMET was substituted for a brief brain-teaser task. Second, Experiment 2b included an exit questionnaire which was meant to control for a possible confound in earlier experiments, namely that the procedure may not have accomplished its goal of presenting a convincingly naturalistic FB interaction. If participants were wise to the fact that they were in the midst of a staged performance, their awareness of the task may have primed them to attend more vigilantly to the actors’ behavior and thus, to engage in deliberate and flexible mindreading early on. An Exit Interview Questionnaire was introduced which was adapted from a questionnaire used by Schneider, Nott and Dux (2014). Originally, this questionnaire was designed to test whether participants had caught onto the purpose of priming and automaticity studies (Bargh, 1994). If participants had sensed that the FB task was the critically relevant task, this should be reflected on the exit interview. Participants who indicated anything to do with the helping scenario (if they mentioned ‘retrieving the bag’, ‘helping the experimenter’, or ‘choosing which bag was the right bag’, or anything to a similar effect) were excluded from the analyses and subsequently replaced. 9 participants in the ‘hands free’ condition (38%), and 2 participants in the ‘hands tied’ condition (8%) indicated something to do with the helping task on their exit questionnaire and were replaced.

5.5.3 Procedure

The procedure followed Experiment 2a exactly. The only task-relevant difference was that the original male and female Asian actors, who played the role of A, were replaced by a different female NZ Pakeha actor.

5.5.4 Results - Preliminary Analyses

Eye movements were coded on a frame-by-frame analysis using the same critical time window as in Experiments 1 and 2 (see section 4.2.5 for details). Eye gaze and body movement data were scored by the primary investigator. A second coder was trained to distinguish looks to compartment *Y* in the videos. The looking data from 20% of participants in both the ‘hands free’ and ‘hands tied’ replication conditions (total n = 72) were randomly chosen to be coded by the second coder (total number of frames scored by second coder = 970). Inter-observer agreement was 93.20%. All disagreements were resolved by discussion with a third coder.

An independent samples t-test revealed that the length of the critical time windows in the ‘hands free’ condition (M = 2,209.03 ms, sd = 436.12) and the ‘hands tied’ condition (M = 2,181 ms, sd = 386.48) did not significantly differ ($t(70) = .573$, $p = .452$).

A new female experimenter played the role of *A* for all new participants (n = 72). A one-way ANOVA with assistant as a between-subjects factor revealed that time windows did not significantly differ when the role of *A* was played by either the original male, original female or new female actor [$F(2, 93) = 2.697$, $p = .073$]. Finally, age, gender or ethnicity did not have an effect on participants’ duration of looking to compartment *Y* in the ‘hands free’ condition. In the ‘hands tied’ condition, ethnicity was correlated with duration of looking to compartment *Y* (p = .007), however, this correlation can be explained by the fact that there were a usually high number of participants in this sample, of ‘NZ European’ ethnicity (22 out of 24 participants).

5.5.5 Looking to compartment Y

Participants’ total duration (in milliseconds) of looking to compartment *Y,* during the critical time window, were compared (‘hands free’: n = 48; M = 304.86, SE = 59.15; ‘hands tied’: n = 24; M = 68.06, SE = 33.30). A Mann Whitney *U* test showed that adults looked significantly longer when their hands were free (*Mean rank* = 41.02), rather than tied (*Mean rank* = 27.46; $U_{(n1=24,\, n2=48)} = 359$, exact 1-tailed $p = .001$, $r = .42$).

5.5.6 Durations of looking during four sub-windows

Average duration of looking to compartment *Y* was analyzed across four distinct sub-windows, in the same manner as Experiment 1. This first sub-window started when A had oriented her body away from the calendar and toward compartment *X* and ended when *A* finished approaching the compartment (M = 1303.70 ms, sd = 396.21). The second sub-window started on the next frame and ended when *A* had bent her knees to jump (M = 306.94 ms, sd = 126.58). The third sub-window started on the next frame and ended as *A* reached the top of her jump (M =

243.98 ms, *sd* = 39.49). Finally, the fourth sub-window began on the next frame and ended when *A* landed her jump and her reaching arm dropped below the middle shelf (*M* = 345.37 ms, *sd* = 99.81).

Table 3 presents the mean, standard deviation, and 95% confidence intervals for the duration of looking to compartment *Y* during either condition and for each sub-window. Separate Mann Whitney U tests were used to assess whether and when participants' durations of looking to compartment *Y* differed during each of the four sub-windows. Results revealed that, for sub-window 1, average looking to compartment *Y* was significantly greater in the 'hands free' condition (*Mean rank* = 40.92) than the 'hands tied' condition (*Mean rank* = 27.67; $U_{(n1=48,\, n2=24)}$ = 364, exact 1-tailed $p = .001$, $r = .36$). No significant difference was found during the following three sub-windows (all *p*s > .05).

Table 3 - Means, standard deviations and 95% Confidence Intervals for the overall critical time window and the four sub-time windows.

Time Window	**Condition**	**Mean**	**Standard deviation**	**95% CI lower bound**	**95% CI upper bound**
Overall Critical Time Window	'hands free'	304.86	409.81	185.87	423.86
	'hands tied'	65.06	163.15	0	136.95
Sub-Window 1: *A* approaches *X*	'hands free'	204.86	319.20	112.18	297.55
	'hands tied'	20.83	72.11	0	51.28
Sub-Window 2: *A* prepares to jump	'hands free'	25.69	79.37	2.65	48.74
	'hands tied'	9.72	47.63	0	29.83
Sub-Window 3: Top of *A*'s jump	'hands free'	25.69	67.10	6.21	45.18
	'hands tied'	9.72	36.09	0	24.96
Sub-Window 4: *A* finishes jump	'hands free'	48.61	112.77	15.87	81.36
	'hands tied'	27.78	76.56	0	60.11

5.5.7 Body Movements

Participants' body movements following *A*'s request for help were divided into the same four stages as in Experiment 1 (swerving, approaching, reaching and handing the bag to *A*). Separate hierarchical log-linear analyses with backwards elimination revealed no significant effect of restriction on any of the body movements (all *p*s > .05). Participants' swerve direction did not significantly differ from chance: when hands were free, 60.4% of adults swerved toward compartment *Y*, and 54.17% of adults swerved toward compartment *Y* when their hands were tied (one-tailed binomial *p*s > .05). For stage-2, participants tended to approach compartment *Y* when their hands were free (72.9%; one-tailed binomial $p = .001$) but did not differ from chance when their hands were tied (63% approached compartment *Y*; one-tailed binomial $p > .05$). For stage-3, participants tended to reach for compartment *Y* (83%; one-tailed binomial $p < .001$) but reaching did not differ from chance (63%) when hands were tied. Regarding stage-4, all participants in either condition correctly gave the appropriate bag to *A*, as mandated by the procedure. Figure 11 shows a schematic visualization of *P*s' movements across four stages of action.

The final prediction was that participants who made a look to compartment *Y* during the critical time window would also swerve toward compartment *Y*. This prediction is informed by the significant positive correlation (between looking to compartment *Y* and swerving) found in the FBLoc condition from Experiment 1. Specifically, it is predicted that a significant positive correlation will be found when participants' hands are free but not when participants' hands are tied. In the 'hands free' replication condition, looking to compartment *Y* was detected during the critical time window, for 26 (54%) out of 48 adults. In the 'hands tied' replication condition, looking to compartment *Y* was detected for 5 (21%) out of 24 adults. A chi-squared analysis revealed that participants were significantly less likely to show any looking to compartment *Y* when their hands were tied rather than free [$\chi^2(1) = 7.251$, $p = .006$]. Lastly, correlational analyses were run to determine whether the significant association between looking to compartment *Y* and swerving was replicated in the 'hands free' condition and was not found in the 'hands tied' condition. Results supported these predictions. For 'hands free' participants, stage-1 of action was associated with whether participants had shown a look to compartment *Y* during the critical time window [$\chi^2(1) = 6.463$, $p = .012$]. Furthermore, as in Experiment 1, looking to compartment *Y* did not correlate with later stages of action (all *p*s ≥ .158). However, for 'hands tied' participants, looking to compartment *Y* was not associated with the first stage of action [$\chi^2(1) = .087$, $p = .585$] or any later stages of action (all *p*s > .05).

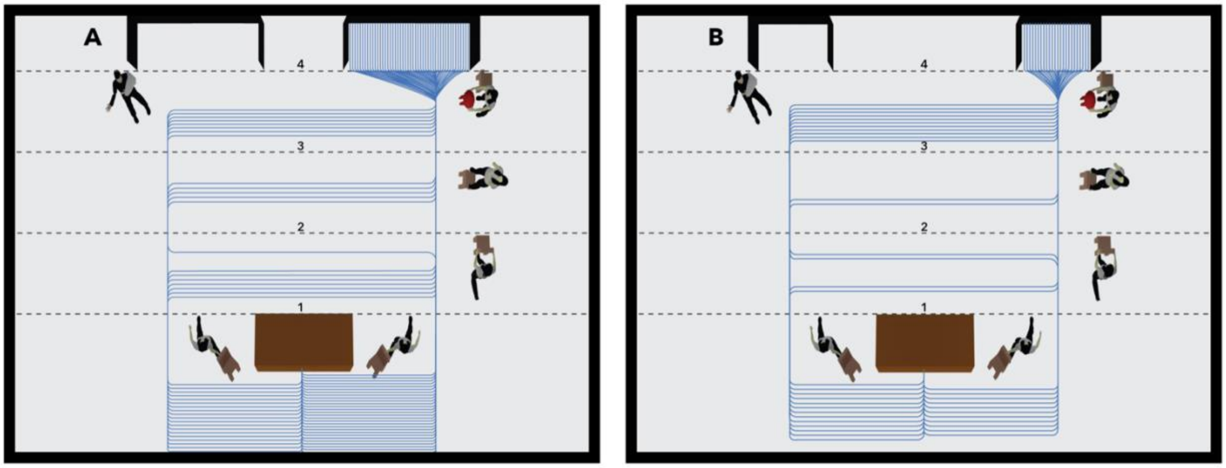


Figure 11 - Schematic representation of *P*'s (*N* = 72) stages of action between conditions (A) ‘hands free’ (*n* = 48), and (B) ‘hands tied’ (*n* = 24).

5.6 Discussion

Experiment 2a and 2b had three main findings. First, participants’ looking to compartment *Y* was negatively affected by having their hands tied behind their back. This finding was supported by three comparisons. Average looking durations were shorter for adults whose hands were tied than those whose hands were free. Also, significantly fewer adults showed *any* looks to compartment *Y* during the critical time window when their hands were tied (Experiment 2b only). Finally, using a more detailed breakdown of the critical time window into four sub-windows, it was found that average durations of looking to compartment *Y* differed during the first sub-window (when *A* was walking from the calendar toward compartment *X*).

Second, the correlation between looking to compartment *Y* and swerving that was found in the FBLoc condition from Experiment 1 was replicated in a higher-powered sample. However, no such correlation was found in the ‘hands tied’ conditions from either Experiment 2 or 2b.

Third, tying up adults’ hands did not produce a corresponding effect on their body movements when their natural helping actions were analyzed across four stages of their helping action. Based on the findings from Experiment 1, it was predicted that stage-1 action would significantly differ when hands were tied. Results of Experiment 2 and 2b did not support this prediction. The implications of these findings will be discussed in turn in the following three sections.

5.6.1 Looking to Compartment Y

Looks to compartment *Y* are interpreted, in the present thesis, as reflecting an efficient goal ascription process sensitive to an agent's registrations. If a minimal form of goal ascription is afforded by motor representation, and if goal ascription is integral to efficient belief tracking, then experimentally stalling motor representations should negatively impact looking to compartment *Y*. The findings support this prediction. Tying up adults' hands resulted in shorter looking durations to compartment *Y* and, when the critical time window was divided into four sub-windows, average looking to compartment *Y* was significantly shorter during the earliest time window (as *A* approaches compartment *X*). Note that in Ambrosini and colleagues' (2012) study, gaze arrival times were delayed when participants' hands were tied. The present study suggests that, when hands are free, motor representations can specify the goal of an agent's movement surprisingly early on in an unfolding action. However, when adults' hands were tied, their motor representations were slower to do so.

5.6.2 Disconnecting looking from swerving

Experiment 2b replicated the finding from Experiment 1: that looking to compartment *Y* and swerving behavior were correlated in the FBLoc condition. Specifically, those who showed a look to compartment *Y*, in the 'hands-free' condition, were more likely to then swerve toward compartment *Y* after the agent requested their help. However, this correlation was no longer found in the 'hands tied' condition from both Experiment 2 and 2b. This suggests that disrupting an observer's access to motor representations diminished belief tracking to the point that their gaze behavior and swerving no longer tracked the agent's belief. One possible explanation for this finding is that the efficient mindreading system has a tenuous influence on overt action production, and by stalling motor representations, the output from a minimal mindreading system is disconnected from motor processes.

5.6.3 Tying up one's hands did not affect their stages of action

Tying adults' hands behind their back was not found to produce a significant effect on any stages of their helping action, compared to those whose hands were free. It was predicted that stalling motor representations would negatively affect earlier stages of action, but not later stages. Although a significant finding in support of this prediction was not found, there was tenuous evidence supporting this prediction in both Experiment 2 and 2b, as adults' stages of action improved (from chance to above chance levels) more quickly when their hands were free rather than tied.

5.6.4 Erring on the side of caution in an age of non-replication

I would be remiss if I did not remind the reader that this investigation is, using a novel FB task, and would greatly benefit from strict external replication and conceptual adaptations to rule out plausible alternative interpretations. Until these findings are confirmed from such follow-up investigations, none of these findings presented herein should be taken as *decisive* evidence. Moreover, there are multiple reasons why these results should be interpreted with caution. Given that there have been concerns about the reproducibility of Ambrosini and colleagues' (2011) findings (the procedure for which the original 'hands tied' condition was compared against; Ambrosini et al., 2012; see Vannuscorps & Caramazza, 2016) the effect of tying one's hands should itself be investigated further before one places too much stock in the present study's findings.

Notwithstanding the above-mentioned reasons for maintaining a critical stance toward the present studies' ability to deliver valid and replicable results, how confident can one be that the observed difference in looking to compartment *Y* was related to a disruption to motor representations? In other words, could this interpretation of the results be confounded by the influence of some extraneous task variable? This seems unlikely. The procedures of the hands-free and hands-tied conditions are identical, the room in which the task was run was set up in the same way for all participants, and both conditions are fully counterbalanced with respect to the side where *R* was standing and the side where *A* was standing.

If a disruption to motor representations *is* the cause for the observed difference in looking to compartment *Y*, it would suggest that motor representations interface with belief tracking and goal ascription even before *A*'s goal is made expressly clear in his or her actions (i.e., while *A* approaches compartment *X*). It is especially surprising that stalling motor representations for hand actions would produce such an early effect on gaze, for two reasons. First, this is not the point where their hands would be reaching using a pincer grip. The rationale for using the pincer grip in the first place, was to stall the motor representations specific to understanding the later stages of the reaching action (e.g., at the top of *A*'s jump). It may be the case that adults whose hands were free used motor representations to make the early forward inference (on the basis of *A* walking toward compartment *X*) predicting that when *A* stops in front of the compartment, s/he will reach for the incorrect bag.

One problematic alternative explanation for the findings from Experiments 2 and 2b, is that by tying participants' hands behind their back and asking them to hold the pencil between their thumb and forefingers, participants were made to incur a cognitive load or a general distraction from the task. If this is the case, then these results do not reflect meaningfully goal ascription and belief tracking interacting with motor processing. To address this concern, the next Experiment sought to formulate an equally loading and distracting task that should not stall adults' specific motor representations for reaching. Therefore, in Chapter 6 data from the 'hands free' and 'hands tied' conditions from Experiment 2b are compared with a new sample of adults (n = 24) whose feet were tied, instead of their hands.

CHAPTER 6

Experiment 3

6.1 Introduction

Experiment 2 found that tying adults' hands behind their back impacted their looking to compartment *Y* (where an agent's backpack was hidden) when an agent tries to reach compartment *X* (where the agent last registered his or her backpack). This finding fits with a recent conjecture from a two-systems account, that an efficient minimal mindreading system interfaces with motor processes. However, one alternative explanation for the findings from Experiment 2 is that tying participants' hands distracted them, and this had an impact on their eye movements. In addition, it is possible that participants' eye movements differed due to a domain-general cognitive load, rather than having to do with stalling specific motor representations. To address this challenge, a new group of adults (*n*= 24) were tested in a condition where their feet were tied instead of their hands. According to the direct matching hypothesis one route to understanding and predicting others' actions is by activating specific regions of the observer's motor system which would be recruited to perform the action themselves. Since there is evidence that restricting access to an observer's hands stalls the arrival of saccades to the target of a reaching action, Experiment 2 aimed to replicate this effect by requiring participants to hold a pencil behind their backs using a pincer grip, as this same grip is required to open the flap of either compartment.

Importantly, motor representations should *not* be impaired by merely restraining an observer, in any way. Only the restriction to areas of the body that would be needed to reach and open the compartment, should affect gaze behavior. The bodily movements which specify the goal of *A*'s action (of trying to reach for the backpack in compartment *X*) are specific to the arms and hands. Therefore, tying up participants' ankles and asking them to be similarly engaged by balancing an object between their feet, as they were when holding the pencil using a pincer grip, should *not* impair adults' motor representations when observing the same procedure.

In Experiment 3, the results from a FBLoc 'feet tied' condition are presented. Participants were asked to keep their ankles bound using a headband and, at the same time, balancing an inflated ball between their feet. It was predicted that participants in the 'feet tied' condition would show looking similar to the participants in the FBLoc 'hands free' condition. Furthermore, it was predicted that participants' swerving direction would be directed toward compartment *Y* significantly more than participants in the 'hands tied' condition. Finally, it is predicted that the same correlation between looking to compartment *Y* and swerving to compartment *Y* will be found when participants feet are tied. Conversely, if tying up the participants' hands only served as a general distraction to participants, then tying up their feet should incur the same distraction and thus, participants in the 'feet tied' condition should display similar looking and swerving behaviors in the 'hands tied' condition.

6.2 Methods

The same European actor from the previous experiments played the role of the Researcher (*R*), and the same female NZ Pakeha actress from Experiment 2b played the role of the Assistant (*A*). The data from the 'feet tied' condition was compared against the 'hands free' replication (n = 48) and 'hands tied' replication (n = 24) data from Experiment 2b.

6.2.1 Participants

Participants were 24 undergraduate adults (19 females) between 18 and 23 years of age (M = 19.04 years). 75% of participants were NZ European and 25% indicated 'other'. All participants provided informed consent prior to the start of the experiment. The same exit questionnaire from Experiment 2b was used to detect any participants who had realized the true purpose of the study. The exit interview revealed that 1 of the participants (4%) was aware of the true purpose of the study and was replaced.

6.2.2 Procedure

The FB procedure for Experiment 3 was identical to the procedure outlined in the FBLoc condition from Experiment 1. Prior to the FB procedure, participants (*P*) were invited to take a seat at a center desk and were told that the study will involve two parts: a block puzzle task and a 'balancing task'. The first part was the same block puzzle task from Experiment 1. After *P* had finished the block puzzle task, *A* removed the blocks and told *P* the following:

> *"Okay, now the second part of the experiment involves a seated balancing task. But before I start you on that, I would like you to stretch this elastic so that it goes around your ankles. Also, you will need to position your feet on top of this ball. I'd like you to maintain this position for about a minute, and I will let you know when you can stop. While you do that, I'm just going to run outside and grab the materials we will need for*

the balancing task. Okay, I'll just put this away and be right back"

A then walked to the ladder and stored her backpack in compartment *X*, and the remainder of the FB procedure was conducted in the same way as the original FBLoc condition. After *P* helped *A* retrieve her backpack from compartment *Y*, *A* opens the backpack and briefly searches through it, and then says:

> *"Shoot! Okay well I can't seem to find the bit that I'm after. There's a device that I'm supposed to use for the balancing task and I thought it was out there or possibly in my backpack, but I can't find it. [pause] Okay, well, what I can have you do instead, is complete this brain teaser exercise."*

The purpose of the 'balancing task' and the lost device was to provide some explanation to *P* as to the reason for keeping their feet tied during the FB procedure. This is similar to the "finger pressure sensor" that is mentioned in the 'hands-tied' conditions. Since the exit questionnaire was completed after *P* helped retrieve *A*'s backpack, these details were included as a preventative measure, with the hopes of mitigating further suspicion of the true purpose of the helping task, during the interim between helping *A* and completing the exit questionnaire. This was important as it aimed to minimize the possibility of participants catching onto the relevance of the helping task *after* the helping task finished.

6.3 Results

In the interest of remaining consistent with the previous chapters, the results section is organized in the following way. First, a preliminary analysis section compared the critical time window within which gaze behavior was coded and analyzed. The next section presented an analysis of the average duration of looking to compartment *Y* across conditions. In the next section, the critical time window was broken down into four sub-windows and average duration of looking to compartment *Y* were compared across conditions, and per sub-window. The final section provided an analysis of participants' movements across four distinct stages of action. The results section concluded by analyzing whether looking to compartment *Y* was correlated with any stages of action. In accordance with previous chapters, when data was found to be non-normal, non-parametric tests were used. Specifically, a Kruskal-Wallis test was used in place of the parametric ANOVA, and Mann Whitney *U* tests were used in place of *t*-tests.

6.3.1 Preliminary Analyses

The performance of ‘feet tied’ participants was compared to the ‘hands tied’ participants (n = 24) and ‘hands free’ participants (n = 48) from Experiment 2b. Eye movements were scored by the primary investigator on a frame-by-frame (30 frames per second) basis. A second coder was also trained to distinguish looks to compartment *Y* from other looks. The looking data from 20% of the data from the ‘feet tied’ condition was randomly chosen for reliability coding (total number of frames scored by the second coder = 341). Inter-observer agreement was 95.8%. All disagreements were resolved through discussion with a third coder. The critical time window was established by applying the same criteria used in Experiments 1 and 2. In the ‘feet tied’ condition, the average time window was 2,393.06 ms (sd = 427.78). Separate independent samples t-tests revealed that the duration of the time windows in both the ‘hands free’ condition (M = 2,209.03 ms, sd = 436.12) and the ‘hands tied’ condition (M = 2,181.94, sd = 386.48) did not differ from the duration of the time windows in the ‘feet tied’ condition (ps > .05).

6.3.2 Looking to compartment Y

The data specifying the duration spent looking to compartment *Y* was found to be non-normal (shapiro-wilks = .639, $p < .001$). A Kruskal-Wallis test with three groups (‘hands free’ (*Mean rank* = 55.55, ‘hands tied’ *Mean rank* = 37.33, and ‘feet tied’ *Mean rank* = 45.56) found that the average time spent looking to compartment *Y* was not the same in all three groups [$\chi^2(2) = 8.986$, $p = .011$]. However, since Experiment 2b already revealed a significant difference between the ‘hands free’ and ‘hands tied’ conditions, separate Mann Whitney U tests were used to assess whether looking durations in the ‘feet tied’ condition differed from either the ‘hands tied’ or ‘hands free’ conditions. Results showed that the average duration of looking to compartment *Y* in the ‘feet tied’ condition (*Mean rank* = 26.63) was not significantly different from the ‘hands-tied’ condition (*Mean rank* = 22.38; $U_{(n1=n2=24)} = 237$, $p = .097$, $r = .19$). In addition, the average duration of looking in the ‘feet tied’ condition was not significantly different from the ‘hands free’ condition ($U_{(n1=48,\ n2=24)} = 454.5$, $p = .059$, $r = .23$; *Mean rank* of ‘hands free’ = 39.03; *Mean rank* of ‘feet tied’ = 31.44). Figure 12 shows the average duration of looking to compartment *Y* during the critical time window for all three conditions. Table 4 shows the means, standard deviation and 95% confidence intervals for the tests described above.

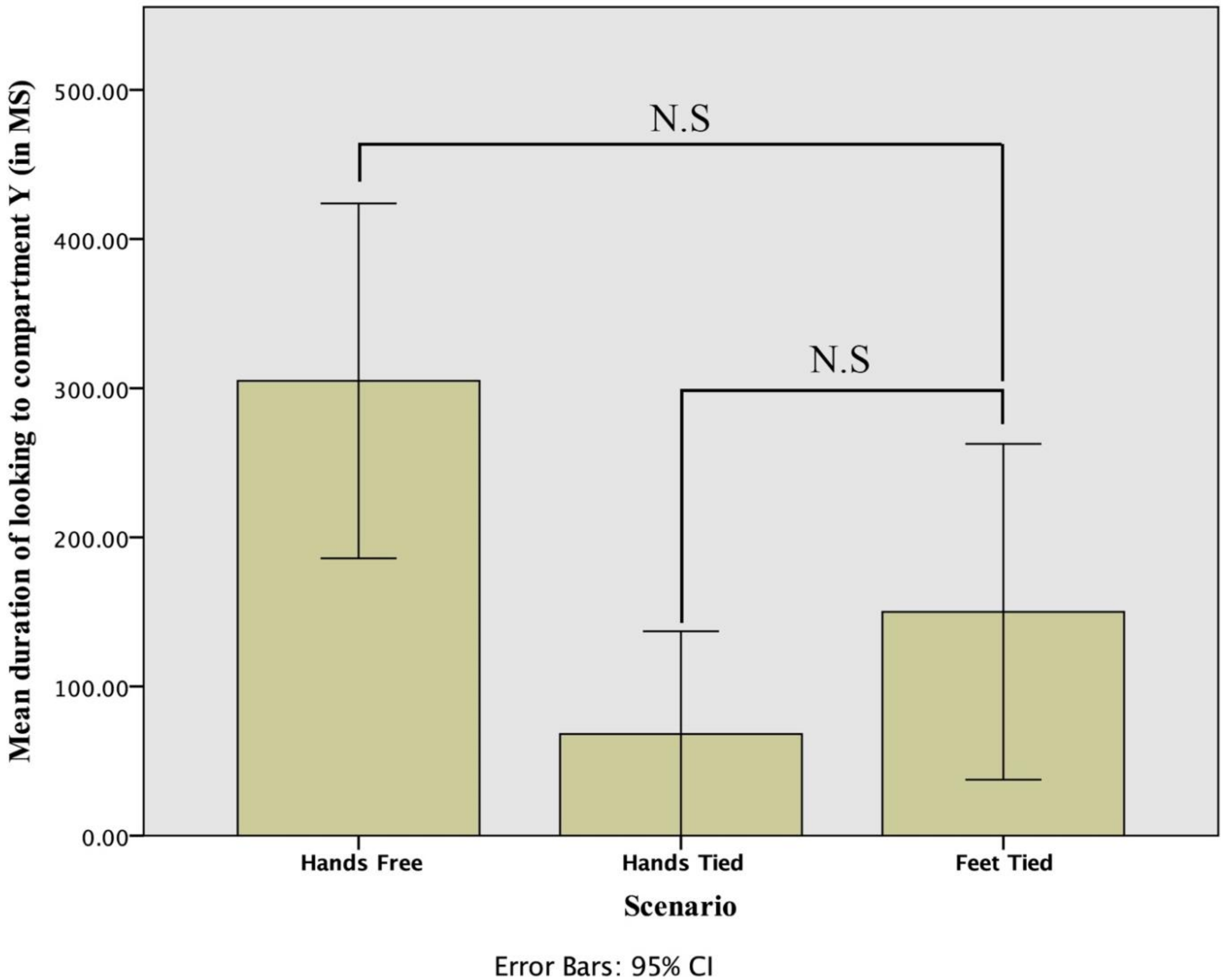


Figure 12 – Mean duration of looking to compartment *Y* (in milliseconds) during the critical time window.

6.3.3 Durations of looking to compartment Y during four sub-windows

The average duration of looking to compartment *Y* was calculated during the same four sub-windows specified in the previous chapters. The first sub-window extended from the point where A is oriented toward compartment *X*, to the point where *A* finished approaching the compartment (M = 1284.85 ms, sd = 314.34). The second sub-window extended from that point until *A*'s knees were bent in preparation to jump (M = 412.12 ms, sd = 168.57). The third sub-window extended from that point until the top of *A*'s jump (M = 256.06 ms, sd = 68.53). The fourth sub-window extended from the top of A's jump until A has landed and his or her reaching arm dropped below the middle shelf (M = 443.94 ms, sd = 159.82).

A Kruskal-Wallis test was used to test whether there were differences between three groups ('hands free', 'hands tied' and 'feet tied') across the four sub-windows. Results revealed that average pro-active looking durations were not the same during sub-window 1 [$\chi^2(2)$ = 10.191, p = .006] (*Mean ranks*: 'hands free' = 55.27, 'hands tied' = 37.60, 'feet tied' = 45.85). Separate Kruskal-Wallis tests showed that average pro-active durations did not significantly differ during the following three sub-windows (all *p*s < .05). To unpack the significant difference found in the first sub-window, separate post-hoc Mann Whitney U tests were used. Note that previous analyses have already shown that the 'hands tied' and 'hands free' were significantly different during this sub-window. When comparing the average pro-active durations made in the 'hands free' (*Mean rank* = 38.85) or 'feet tied' (*Mean rank* = 31.79) conditions, results revealed that looking was not significantly different ($U_{(n1=48,\ n2=24)}$ = 463, exact 1-tailed p = .059, r = .18). In addition, looking to compartment *Y* was not significantly different when comparing 'hands tied' (*Mean rank* = 22.44) and 'feet tied' (*Mean rank* = 26.56) data ($U_{(n1=n2=24)}$ = 238.5, exact 1-tailed p = .059, r = .23). See Figure 13 for separate looking duration graphs for the four sub-windows. Also, see Table 4 for the means, standard deviations and 95% confidence intervals for the data displayed in Figures 12 and 13.

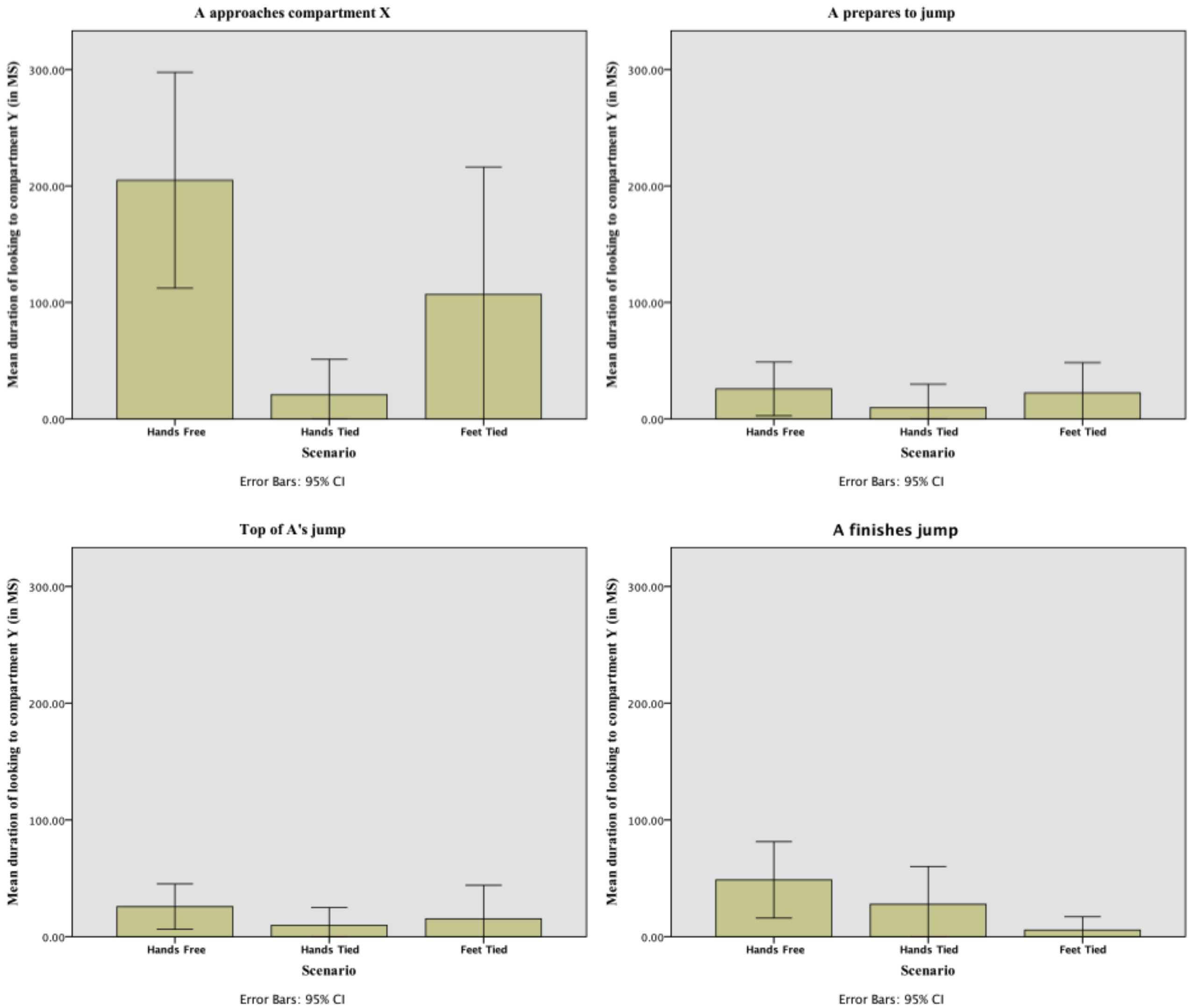


Figure 13 – Mean duration of looking to compartment *Y* (in milliseconds) across four distinct critical sub-windows. In the first sub-window, *A* approaches compartment *X*. In the second sub-window, *A* bends his or her knees in preparation to jump. The third sub-window ends when *A* reaches the top of his or her jump. The fourth sub-window ends when *A* has landed his or her jump.

Table 4 - Means, standard deviations and 95% Confidence Intervals for the overall critical time window and the four sub-time windows.

Time Window	**Condition**	**Mean**	**Standard deviation**	**95% CI lower bound**	**95% CI upper bound**
Overall Critical Time Window	'hands free'	304.86	409.81	185.87	423.86
	'hands tied'	65.06	163.15	0	136.95
	'feet tied'	150.00	266.67	37.40	262.60
Sub-Window 1: *A* approaches *X*	'hands free'	204.86	319.20	112.18	297.55
	'hands tied'	20.83	72.11	0	51.28
	'feet tied'	106.94	258.76	0	216.21
Sub-Window 2: *A* prepares to jump	'hands free'	25.69	79.37	2.65	48.74
	'hands tied'	9.72	47.63	0	29.83
	'feet tied'	22.22	61.91	0	48.36
Sub-Window 3: Top of *A*'s jump	'hands free'	25.69	67.10	6.21	45.18
	'hands tied'	9.72	36.09	0	24.96
	'feet tied'	15.28	68.09	0	44.03
Sub-Window 4: *A* finishes jump	'hands free'	48.61	112.77	15.87	81.36
	'hands tied'	27.78	76.56	0	60.11
	'feet tied'	5.56	27.22	0	17.05

6.3.4 Motor Responses

Figure 14 presented a schematic visualization of all participant's movement patterns across four stages of action. Separate log linear analyses with backwards elimination were conducted to compare participants' performance on the four stages of action in the 'feet tied' and 'hands tied' conditions. Results showed that restriction did not have a significant effect on any of the four coded stages of action (all *p*s > .05). The same analyses were run comparing 'feet tied' and 'hands free' stages of action, which also showed that restrictions had no significant effect on any stages of action (all *p*s > .05).

Figure 14 – Schematic Representation of *P*'s (N = 96) stages of action between conditions: (**A**) 'hands free' (n = 48), (**B**) 'hands tied' (n = 24) and (**C**) 'feet tied' (n = 24). Note that both conditions are FBLoc scenarios and (**A**) and (**B**) are the same graphs presented in Figure 11 (Experiment 2b).

Finally, it was reasoned that that tying up participants' ankles should not disrupt a connection between efficient mindreading and the direction of participants' early helping action. Therefore, it was predicted that looking to compartment *Y* during the critical time window would be correlated with swerving in the 'feet tied' condition. Looks to compartment *Y* during the critical time window were detected for 9 (37.5%) participants in the 'feet tied' condition. By contrast, looks to compartment *Y* were detected in 26 (54.17%) participants' looking data in the 'hands free' condition, and in 5 (20.83%) participants' looking data in the 'hands tied' condition. A chi-squared analysis revealed that participants were not significantly more likely look to compartment *Y* when their feet were tied compared to when their hands were tied [$\chi^2(1) = 1.613$, exact 1-sided $p = .171$]. In addition, participants were not significantly more likely to look to compartment *Y* in the ‘hands free’ than in the ‘feet tied’ condition [$\chi^2(1) = 1.779$, exact 1-sided $p = .139$]. In regard to the connection between swerving and looking to compartment *Y* found in the ‘hands free’ conditions, no significant correlation between looking and swerving was found in the ‘feet tied’ condition [$\chi^2(1) = .548$, $p = .375$]. In addition, no association was found between to compartment *Y* looking and stage-2 action (approaching) [$\chi^2(1) = 2.517$, $p = .122$] in the ‘feet tied’ condition. Surprisingly, a significant correlation was found between looking to compartment *Y* and stage-3 action (reaching) [$\chi^2(1) = 6.993$, $p = .011$] in the ‘feet tied condition’.

6.4. Discussion

Experiment 3 was designed to test for the possibility that tying one’s hands in Experiment 2a and 2b did not have the specific effect on motor representations, and their supposed interaction with efficient belief tracking. For the conclusions drawn from the results of Experiments 2a and 2b to be valid, it is implied that observers made use of their motor system to predict and interpret others’ actions by activating regions of the motor cortex that would be responsible for producing the same actions themselves. It was for this reason that Ambrosini and colleagues (2012) hypothesized that tying adults’ hands would impair their ability to take advantage of the posture of an actor’s reaching hand for predicting which of two objects the actor is about to grasp. When observers were not in a position to perform the action themselves or were engaged in a different motor task whilst observing the reaching action (Constantini, Ambrosini & Sinigaglia, 2012), adults’ predictive looking was selectively stalled (i.e., the speed of pro-active looking to the correct target was not modulated by whether the actor used a pre-shaped grip or no grip, when reaching toward the objects).

In Experiment 2a and 2b, the aim was to maximize the likelihood of stalling an observer’s access to motor representations necessary for determining the outcome of an agent’s reaching motion. Adults’ hands were tied behind their backs and they were also engaged in holding a pencil between their hands using the same kind of ‘pincer’ grip needed to open the flap on either of the compartments. As predicted, significantly less looking to compartment *Y* was observed during the agent’s approaching and reaching action. In addition, the correlation observed between looking to compartment *Y* and swerving toward compartment *Y* was not found when participants' hands were tied, even though both conditions involved the agent having a FB about an object's location. These findings were replicated in a second experiment (Experiment 2b) with higher statistical power. It is important to make sure that these results were not the product of a domain-general distraction or cognitive load incurred by the tying of participants' hands. Therefore, Experiment 3 tied participants' feet, but left their hands free. If the stalling of motor representations specifically responsible for executing and understanding hand actions explained adults' differential performance in Experiment 2 and 2b, then tying their feet should have no effect on their ability to make use of motor representations for interpreting the agent's reaching action. Therefore, it was predicted that participants' looking to compartment *Y* would resemble the 'hands free' condition and differ significantly from the 'hands tied' condition.

The results from Experiment 3 were ambiguous. On the one hand, those whose feet were tied spent as long looking at compartment *Y* as those whose hands were free. However, those whose feet were tied also looked as long to compartment *Y* as those whose hands were tied. In addition, participants were as likely to make a look to compartment *Y* if their feet were tied, as those whose hands were tied, or those whose hands were free. In terms of their stages of action prior to handing the bag to *A* (i.e., swerving, approaching, and reaching), tying participants' feet, hands or leaving them entirely unrestrained, had no significant effect. Thus, the data from the 'feet tied' condition is situated squarely between the 'hands free' and 'hands tied' data. One finding was less ambiguous: looking to compartment *Y* was not correlated with participants' swerving behavior in the 'feet tied' condition. Surprisingly, looking to compartment *Y* predicted later (stage-3: reaching) action rather than earlier action.

The results of Experiment 3 are challenging to decipher. One possibility is that tying up the feet did not successfully supply an equally loading task for observers. In meeting the challenge of designing an equally loading concurrent task to having one's hands tied, a manipulation that involved a similar restriction on observers' feet seemed to be the only appropriate alternative. However, keeping an inflatable ball beneath one's feet may have been less challenging than keeping a pencil held between one's fingers. Another potential problem is that the agent does use their feet to approach and then jump to reach compartment *X*. However, if motor representations for walking and jumping actions are also disrupted by restraining an observers' legs, then the duration of looking to compartment *Y* should be significantly less than those whose hands and feet were unrestricted. Since looking durations did not bear any significant difference to either opposing condition, there is little evidence here to support or refute the role of motor representations in efficient mindreading.

One means by which future research could help shed light on these results, is by restricting adult observers' feet while observing the videos from Ambrosini et al. (2011). In their videos, the actor is seated, and their lower body is offscreen. This should maximize the irrelevancy of tying one's feet, with regards to impairing motor representations for action prediction. If tying participants' feet impacts one's ability to generate matching motor representations of walking agents, then tying up an observers' feet should have no impact on the pro-activity of observers' gaze arrival time when the actor's feet are out of view and unrelated to the observed action.

One result which challenges the predictions from a two-systems account, was the absence of a correlation between looking to compartment *Y* and swerving direction, in the 'feet tied' condition. This finding provides tenuous support for the alternative interpretation (i.e., that restrictions to observers' hands or feet applied a domain-general cognitive load). However, without any clear difference in looking durations, the absence of this correlation constitutes relatively weak evidence. Another surprising finding was that, in the 'feet tied' condition, showing a look to compartment *Y* was correlated with the third stage of action, i.e., reaching to compartment *Y*. Neither the two-systems account nor any of the available alternative theories would have expected a delayed correlation when participants' feet were tied. The most plausible explanation is that this correlation was a spurious false positive and an example of how a sample size of $n = 24$ may be inadvisable for testing predictions using this procedure. By contrast, since a correlation between looking to compartment *Y* and swerving behavior was found in the FBLoc scenario from Experiment 1 *and* in the replication condition from Experiment 2b (using a larger sample size of $n = 48$), this correlation is less likely to have resulted from spurious patterns of behavior.

6.4.1 Conclusions

These results from Experiment 3 were ambiguous, since the data from a ‘feet tied’ condition did not differ from either the ‘hands tied’ or ‘hands free’ conditions. Future research into an overlap between motor representations and efficient belief tracking should take these results into consideration when designing experimental manipulations aimed to interrupt motor representations by restraining specific body parts. Like the findings from previous studies, these conditions should be externally replicated (with larger sample sizes) before the present results are treated as evidence either strongly favoring or opposing the claim that tying one’s hands impacts their action prediction and efficient mindreading abilities.

CHAPTER 7

General Discussion

7.1 Introduction

Three empirical chapters explored the proposal that motor processes interface with efficient mindreading. The final chapter summarizes the main findings from the present work, considers what implications these results have for a two-systems account of human mindreading, and whether alternative theories could offer explanations for these findings. Finally, some of the limitations of the methodology are discussed, as well as some suggestions for future research to address these limitations.

7.1.1 A Summary of Findings

The present thesis tested a novel framework for measuring the eye gaze and body movements of adult participants in a naturalistic-seeming active helping FB task. Chapter 4 (Experiment 1) demonstrated that observers' duration of looking to compartment *Y* (holding the backpack which the agent truly wants), differed according to what the agent believed. When the agent had a FB about the backpack's location, looks to compartment *Y* were greater than when the agent had a FB about the identity of a backpack (i.e., an identical backpack stored in the same place where the agent last stored their backpack). In matched TB conditions where the agent observed the transfer of the backpacks, the same contrast in looking duration was not observed. After the agent asked the participant for help reaching the backpack, participants' stages of helping action were divided into four distinct stages (swerving, approaching, reaching and retrieving the backpack). It was found that the veracity of the agent's belief (FB or TB) and the conceptual complexity of the agent's belief (location or identity) significantly affected the direction that participants swerved around their desk, and which backpack participants ultimately retrieved for the agent. Finally, looking to compartment *Y* predicted participants' subsequent swerve direction but this correlation was only found when the agent had a FB about the location of his or her backpack.

In Chapter 5 (Experiment 2a and 2b), adults' hands were tied behind their backs while observing the same procedure as Experiment 1's FB location scenario. The average duration of looking to compartment *Y* was significantly shorter in the 'hands tied' condition, compared to participants in the 'hands free' FB location condition from Experiment 1. Tying participants' hands did not produce a significant effect on their stages of helping action, however, the correlation between looking to compartment *Y* and swerving direction, was not observed in the hands-tied condition. In a follow-up Experiment (Experiment 2b), the same comparison was tested again using a larger sample for higher statistical power in the ‘hands free’ condition (n = 48). The same pattern of results from Experiment 2a were successfully replicated in Experiment 2b. Specifically, looking durations were significantly longer when participants’ hands were free rather than tied, and looking to compartment *Y* predicted the swerving direction of ‘hands free’ participants but not ‘hands tied’ participants. In addition, the correlation previously observed between looking to compartment *Y* and swerving, was successfully replicated, and no such correlation was found in both ‘hands tied’ conditions.

Finally, in Chapter 6 (Experiment 3), a control condition was tested whereby adults’ feet were tied in order to address the possibility that the effect from hand tying was not the product of a domain-general cognitive distraction. The duration of looking shown by participants whose feet were tied, did not significantly differ those whose hands were free or those whose hands were tied. In addition, tying participants’ feet did not affect their directional movements across their four stages of action. No correlation was found between looking to compartment *Y* and swerving direction in the ‘feet tied’ condition.

7.2 Implications for an efficient minimal system for tracking registrations

An important feature of the two-systems account of human mindreading is that the efficient system (which operates on registrations as proxies for beliefs) shows signature limits. Previous investigations into signature limits have focused on the meaningful distinction between location and identity FBs. While FBs about an object’s location are able to be expressed in terms of registrations (e.g., the agent last registered his or her backpack inside compartment *X*, but the backpack is now in compartment *Y*) FBs about an object whose appearance causes the agent to think that the backpack in compartment *X is* the backpack that the agent last registered in compartment *Y* are beyond what registrations can convey. The results from Experiment 1 are in line with this description, as evidenced by observers’ differential duration of looking to compartment *Y*. Those who observed an agent who developed a FB about the location of their backpack spent longer looking at compartment *Y* in preparation of their own helping action, compared to those in FBID. In the present identity scenario, registrations cannot encode for the fact that an agent would see one bag as the other. In this case, the minimal mindreading system incorrectly attributed a TB to the agent and therefore, does not plan to help the agent retrieve a backpack from compartment *Y*, thus resulting in a reduction in looking to that compartment.

Earlier investigations into the signature limits of the minimal mindreading system have focused on responses such as participants' first saccade recorded during a critical time window, or differential duration of AL to one of two boxes (e.g., Low & Watts, 2013; Low et al., 2014; Wang, Hadi & Low, 2015). Experiment 1 is the first to report evidence of signature limits in gaze behavior in an active helping FB task. In addition, Experiment 1 reports evidence from the first "compression" identity FB task (i.e., the agent has a FB that there is only one object present, when in reality there are two identical objects which never are shown at the same time). Since previous positive evidence for signature limits has been found only in "expansion" identity FB tasks (where the agent has a FB that there are two objects in separate locations, when really there is just one object with multiple aspects) this supports the notion that numerical discrepancy is the defining characteristic of an identity FB (Butterfill & Apperly, 2013; Low et al., 2016).

More recent signature limits research has found tentative evidence of signature limits in belief tracking revealed through the pre-activation of adults' motor system (Edwards & Low, 2017). Specifically, it was shown that adults' motor system was incorrectly primed to expect an agent would act as would be appropriate if the agent's view of the scene was the same as the participants' view. The researchers concluded that the traceable influence of signature limits on motor activation suggests that efficient belief tracking and motor processes may interact. Two-systems theorists have speculated about the possibility that efficient belief tracking could potentially defer to motor representations (when possible) for efficiently attributing goals to an agent (an antecedent of minimal mindreading; see Michael & Christenson, 2016; cf. Butterfill & Apperly, 2016). In the present thesis, Experiment 1 provides further suggestive evidence in support of an interface between efficient belief tracking and motor processes, by showing that a significant correlation between looking to compartment *Y* and swerving action, was found in the location FB condition and not in the identity FB condition.

Taken together, these results fit with the predictions from a two-systems account. However, there are some noteworthy caveats. In the next section, some findings that conflict with the two-systems account are highlighted and discussed.

7.3 Alternative explanations for the present data

The findings from the present body of work could be explained by alternative accounts. The two-systems account faces challenges from two fronts: accounts arguing that a single mindreading system underpins implicit and explicit mental state attribution throughout the lifespan (the EMA), and accounts arguing that competence on implicit tasks may be facilitated by simpler cognitive strategies than even a minimal model of mind. In several replication attempts, it has been revealed that Buttelmann et al.'s (2009) results are subject to alternative explanations which are at least as (if not more) plausible than the authors' mentalistic interpretations (e.g., Allen, 2015; Priewasser et al., 2017). In the following sections, attempts are made to explain the present findings by appealing to alternative accounts, in order to more clearly explicate the challenges for future research continuing this dissertation's line of inquiry.

7.3.1 The EMA

The EMA would expect to find adults well prepared for the type of belief reasoning tapped by this procedure. With respect to looking duration, an EMA theorist would expect that adults would spend the same amount of time looking to compartment *Y* in either the location or identity FB scenarios. The results from Experiment 1 are in conflict with this prediction, as looking duration was significantly shorter in the FBID scenario, relative to the FBLoc scenario. How a unified system for mindreading informs one's later action, has yet to be postulated by the EMA. Presumably, it would be expected that adults will help an agent with a FB by directing their actions toward compartment *Y* in either FB scenarios, and toward compartment *X* in either TB scenarios. The fact that, in the FBID scenario, adults tended to swerve toward compartment *X* and only later adjusted their trajectory toward compartment *Y*, would be at odds with these predictions.

Although it is unclear what the EMA would predict in regard to potential correlations between looking to compartment *Y* and subsequent stages of action, it seems safe to assume that whatever an EMA theorist would predict, the FBLoc and FBID scenarios would both correlate (or not). Since the EMA has not accommodated a role for motor representations for mindreading, it is likely that an EMA theorist would also predict that tying one's hands or tying one's feet would not have any impact on an observer's mindreading ability. The results from Experiments 2a and 2b would challenge these predictions. The results from Experiment 3 are ambiguous with respect to these predictions.

Taken together, the present findings are not easily explained by the EMA, and in several instances, the findings are in direct conflict with the EMA's predictions. Thus, the two-systems account can more clearly and comprehensively explain the present experiments' findings.

7.3.2 Lower-level accounts: Statistical learning and association biases

The present findings could also be explained by certain lower-level alternative strategies. Such accounts argue that nonverbal responses could reflect an understanding of behaviors as informed by experiences, rather than some sophisticated and systematic grasp of others' mental states. Through statistical learning, humans generate a series of 'behavior rules' based on perceived actions and outcomes. Since people tend to search for things where they last placed them, a behavior-reading observer should also predict that the agent in the present study will go to compartment *X* in the FB scenario (and therefore, the observer should look to compartment *Y* where the agent's object resides). However, rule-based explanations alone would not explain why adults would show less looking to compartment *Y* in a FBID scenario.

Another product of statistical learning which could explain differential nonverbal responses in a FBID scenario, are associations. Perner and Ruffman (2005) argued that Onishi and Baillargeon's (2005) results (i.e., that infants look reliably longer when an agent's actions violate their FB) could be the result of the infant having formed a three-way association between the agent, an object and its location. Although the three-way association explanation was used to explain why infants look longer in one test trial over another, perceptual associations may also offer an explanation for the present results. An observer may have formed an association between the agent, the red backpack, and compartment *X*. Note that in the FBID scenario, the scene ends up looking identical to how it did when the agent exited. With this association in mind, an observer's fixations may have been directed toward the backpack in compartment *X*. Because of the matching visual association between agent, object and location, an observer may initially be led to act as if the agent is retrieving the same object that s/he originally stored there, and only later (upon further consideration) does the recognize their error. The same explanation could apply to the FBLoc scenario. Thus, the results from Experiment 1 could be explained by appealing to a hybrid model of lower-level processes that are the product of statistical learning such as behavior-rules and three-way associations, rather than necessarily having resulted from domain-specific belief tracking. Although these lower-level explanations are problematic for the two-systems account, there has yet to emerge any nonverbal FB task the results from which could only be explained by a mentalistic interpretation, and not by any complementary behavior reading explanation (see Lurz, 2011).

Furthermore, there are similar statistical learning explanations for motor activation prior to predictable actions (Monroy et al., 2017). However, such accounts have yet to be reconciled with the evidence showing that the motor system is impaired under certain external constraints (Ambrosini, Constantini & Sinigaglia, 2012; Constantini, Ambrosini and Sinigaglia, 2012; Constantini et al., 2013). Assuming that a statistical learning account for motor activation could accommodate the evidence showing that external constraints impair action understanding, it is therefore plausible that statistical reasoning could even explain the observed results from Experiments 2a and 2b. Taken together, the results from the present study do not permit one to reject the possibility that adults' looking durations and helping actions were informed by domain-general statistical learning.

7.3.3 'Pure' Teleology vs. Minimal Mindreading

The results of the present study could also be explained by teleological accounts. Relative to belief tracking and mindreading, teleological accounts are simpler explanations for infants' and young children's nonverbal behaviors in various FB tasks. The teleologist bases his or her inferences about other people's actions on observable objective facts about the environment, which are then used to derive an agent's the likely reasons for producing one action or another (Perner & Esken, 2015; Perner & Roessler, 2010). By contrast, mental states such as beliefs, are subjective facts that cannot be observed, only inferred. According to Perner, Priewasser and Roessler (2018) teleology does not emerge from a specific cognitive 'system' or 'network'. Instead, teleology is a *practical reasoning schema* constructed and reinforced by experience, for interpreting the actions of others in light of a universal desire to bring about a more attractive state of the world. It is further explained that the appreciation for others' mental states does come into play but later in development. During infancy, humans are 'pure' teleologists, meaning that other's actions are interpreted based on objective facts about the world independent of the agent's perspective on the world.

In the FBLoc scenario of Experiment 1, a teleological observer could reason that because the agent was not present when their backpack was switched (the agent's absence is an objective fact about the environment, as is the backpacks switching), the agent could be seen as having good reasons to go to compartment *X* in order to try to retrieve his or her backpack. An observer using teleology would consequently look toward compartment *Y* in preparation to help the agent find the backpack that he or she is after. However, in the present study, applying teleology relies on the observer being able to keep in mind which backpack is in compartment *X* and compartment *Y*. While doing so is manageable in a FBLoc scenario, the identical appearance of the two red backpacks in the FBID scenario may be as problematic for a teleologist as it is for a minimal mindreading system. If, in the FBID scenario, an observer considered that the agent's goal is to retrieve the red backpack, then an observer may conclude that the agent has good reasons to get the backpack that he or she wants by going to compartment *X*, which the agent can see is presently housing the red backpack. Therefore, much like how a two-systems account predicts that a minimal mindreading system will come to the incorrect conclusion that the agent's registration is a condition for successful goal-directed action, a teleological account may also expect that observers incorrectly inferred that the agent's actions would permit him or her to obtain the goal. Therefore, teleology could just as successfully account for the difference in looking to compartment *Y* observed in the FBLoc and FBID scenarios from Experiment 1.

Furthermore, in both TB scenarios, it is an objective fact that the agent was present when both backpacks were moved into their respective compartments. For this reason, a teleologist should expect the agent to go to compartment *Y*, where they have seen their backpack moved. However, if the agent goes toward compartment *X* a teleologist might also conclude that the agent must have good reasons to do so. Therefore, observers should not look to compartment *Y* as much as those in the FBLoc scenario; the observer should be prepared to help the agent reach the object in whichever compartment they are reaching toward. Moreover, whatever object replaces the red backpack should not affect the duration of observers' looking to compartment *Y*. In Experiment 1, the two TB conditions did not differ from one another. This is in line with the predictions from both a two-systems account and a teleology account.

However, the two-systems account and the teleology account do not make entirely equivalent predictions about observers' looking behaviors in TB scenarios. While a mindreading hypothesis would expect to find highly polarized patterns in TB (compared to FB) scenarios, a teleology account would expect that observers in a TB helping scenario would be unsure about how to help the agent. For example, a recent attempt to replicate Buttelmann and colleagues (2009), reported that although 18-32-month-old children were significantly more likely to approach box B (where the object resided) when the agent had a FB, their selections were at chance when the agent had a TB with 43% of infants selecting box A (Priewasser et al., 2017). By contrast, the original study reported that 81% of infants chose box A (the empty box) in the TB condition. The logic behind Buttelmann and colleagues' hypothesis was that infants who are operating on genuine belief concepts should realize that the agent is looking for the object in box B, in the FB condition, and that the agent must *not* be looking for the object in box B, in the TB condition. However, without a clear and contrasting pattern of responses in the TB condition, one cannot conclude that infants' selections were based on reasoning about the agent's belief or about the objective reasons that the agent has for acting in a certain way.

Priewasser and colleagues (2017) made a simple adjustment to the original procedure by adding a third box (box C). Unlike box A and box B, box C never houses the desired object. Therefore, in a new FB condition where the agent stored the toy in box A, did not witness the toy moved to box B, and then tried to open box C, the mentalistic hypothesis differs from a teleological hypothesis. Buttelmann and colleagues (and the EMA account) would expect that infants and young children will help the agent open box C. The agent has a FB, however, since the agent is not acting on their FB (by going to box A), the agent must have some different reason for going to box C, and thus, that is the box that they want opened. By contrast, when an infant or young child operates on teleology and is not processing the agent's mental states, they should would expect that the agent has the goal of retrieving their toy. Therefore, a teleological account would predict that in the three boxes FB condition, infants would go to box B (where the toy is) even when the agent goes to box C. Priewasser and colleagues' (2017) results were strongly in support of the teleology prediction: when the agent goes to box C in a FB condition, significantly more children helped him or her by opening box B where the toy was hidden.

Future research could make a similar minimal adjustment to the present procedure in order to test whether adults are operating on a teleology or belief tracking when planning their helping behavior. Since the present procedure involves the displacement of two objects to two respective locations, two empty compartments could be added (to counterbalance for perceptual side biases). In a FBLoc scenario, the agent could re-enter the room and, following the phone conversation, turn and approach one of the compartments which have always been empty. A two-systems account (as well as the EMA, in this case) would expect adults to look less to compartment *Y* (still the compartment that houses the backpack) since the agent's FB is that their red backpack is in compartment *X* (not in this unrelated compartment). A teleological account would expect adults to look as much to compartment *Y* as those in the FBLoc scenario. In this way, the present study could be refined to help tease apart a belief reasoning from teleological reasoning explanations for adults' helping actions. However, for the time being the teleological account can offer a plausible alternative explanation for these results and therefore challenges a two-systems explanation.

7.4 Limitations and Future Directions

The current study has several limitations, some of which have already been discussed. Improvements could come from a number of sources including eye gaze data collection methodology, increased sample size, and procedural adjustments that minimize the length and variance in time windows. These limitations open up new avenues for future research into active helping tasks, spontaneous belief reasoning, signature limits, and ecological validity. Most importantly, the present results would greatly benefit from strict external replication before these findings are heralded as clear evidence favoring one theory over others.

7.4.1 Predictions about minimal mindreading in the True Belief scenario

In Experiment 1, a two-systems hypothesis of adults' performance in the TBID and TBLoc scenarios predicted that looking durations would not significantly differ. In addition, stages of action also should not significantly differ and should be largely directed toward compartment *X* in either of the TB scenarios. One prediction which was less clear was whether looking to compartment *Y* should correlate with swerving in the TB scenario. In the FBLoc scenario, a two-systems hypothesis predicted that looking to compartment *Y* would correlate with swerving because the minimal mindreading system may interact with motor processes. Since a minimal mindreading system should be able to recognize that an agent in a FBLoc scenario is acting on a registration that is not a condition for successful goal-directed action, early actions should reflect this early understanding and should therefore show looks to compartment *Y* and subsequently swerve that way. In the TBID and TBLoc scenarios, the minimal mindreading system should come to the opposite conclusion: that the agent's last registration *is* a condition of successful goal-directed action and that the observer should not look nor swerve to compartment *Y*. According to this logic, it could also be predicted that a correlation *should* have been found between looking to compartment *Y* and swerving. Therefore, it is unclear whether the results from the TB scenarios were in line with the predictions from a two-systems account. The TB scenarios further show how the logic underpinning otherwise straightforward comparisons can be clouded in helping tasks. Future research should consider taking further steps to uncover how participants think about TB helping scenarios, as there is also evidence showing that TBs are not processed more automatically than FBs (Back and Apperly, 2010) and that even 4- and 5-year-old children struggle on TB tasks (for a review see Hedger and Fabricius, 2011). One means by which insight into TB reasoning could be uncovered, is by measuring infants' and children's gaze in replications of relatively established FB helping tasks in the developmental literature which include TB control scenarios (e.g., Fizke et al., 2017; Buttelmann et al., 2009).

7.4.2 Eye Tracking Measures: the benefit of new technology

In setting up an ecologically valid active helping FB task, it was clear that high fidelity eye tracking equipment could not be used. Current eye tracking technology requires participants to use a chin-rest for head stabilization, to calibrate their eyes to the tracking system, and in most cases, to observe the stimulus on a computer screen. To maximize the ecological validity of a helping task's results, it was necessary to design a procedure that could convince an adult participant that an agent's FB was the product of a set of spontaneous circumstances. Therefore, eye tracking was scored by human coders using a high-resolution video camera pointed at the adults' face. One problem with this approach is that it is challenging to distinguish looks to one location from another, if those two locations are close to one another. Many participants' eyes tracked the agent during his or her approach, but as the agent began to reach upward toward compartment *X*, it became increasingly challenging to determine whether a participants' eyes were targeting the agent's reaching hand or compartment *X*. Thus, the majority of looks to compartment *X* were ambiguous. Although there were also theoretical reasons to disregard the coding of looks to compartment *X* (i.e., one might look there simply because the agent is gesturing toward it, rather than in anticipation of the agent acting on his or her FB) the ambiguity problem made looks to compartment *X* impossible to include in the present analysis.

One way in which this could be solved is by using wearable eye trackers (e.g., Tobii Pro® Glasses). This new technology records participants' eye movements as well as what participants are looking at. This would greatly improve the precision of all looking data, and would solve the ambiguity problem in looks to compartment *X*. Of course, there are still theoretical concerns about what a two-systems account would predict about looks to compartment *X*. On the one hand, since it is hypothesized that fewer looks to compartment *Y* will be observed in the FBID condition, then one might expect to find more looks to compartment *X*. On the other hand, there are a number of other locations than compartments *X* and *Y* where a participant may look. Just because there are fewer looks to compartment *Y* does not mean that there will necessarily be more looks to compartment *X*. Thus, another potential hypothesis regarding compartment *X* could be that looks will not differ across conditions.

7.4.3 Further establishing the critical time window

Ecological validity is a high priority in the present study. As such, there are some variables which are permitted to vary in order to maintain a convincing sequence of events. One such variable is the length of the critical time window. Although the actors in the present experiment thoroughly rehearsed the event sequence, the critical time window still varied. Future replications of the present study should consider using a professional actor for the role of *A* and requiring an even stricter rehearsal of the critical time window in order to make it as similar as possible across participants. Doing so would further improve the experimental control of the data collection whilst retaining its ecological validity.

7.4.4 Non-Replication of Ambrosini, Constantini and Sinigaglia (2012): Alternative comparisons

Chapter three discussed multiple means by which motor representations may be experimentally disrupted. Tying up one's hands is just one of those ways. The legitimacy of this claim comes from Ambrosini, Constantini and Sinigaglia (2012) who showed that adults' pro-active gaze shifts to the target of an ongoing reaching action were not improved by the actor reaching with a pre-shaped grip. By contrast, observers whose hands were unrestrained were significantly faster to fixate the target object when the actor used a pre-shaped grip instead of reaching with a closed fist. However, a recent external replication attempt did not find that adults made use of the hand shaping cue (Vannuscorps & Caramazza, 2016). This is problematic for the present study (Experiments 2a, 2b and 3), as it threatens the rationale for tying up participants' hands (or feet) in the first place. To address this concern, additional external replication attempts of both Ambrosini, Constantini and Sinigaglia (2011, 2012) are needed.

If future investigations fail to replicate the original findings, the present study could still offer a means by which to test how eye movements and stages of helping actions differ when motor representations are stalled. For instance, there is evidence showing that engaging in a concurrent motor task could also disrupt motor representations (Brass, Bekkering & Prinz, 2001; Kilner, Paulignon & Blakemore, 2003; Constantini, Ambrosini & Sinigaglia, 2012). Therefore, an alternative means for future adaptations of the present study to impair participants' motor representations would be to ask participants to complete a finger tapping task following the block puzzle task. Participants could be told that a tape recorder will be used to score their performance and that the agent will go and retrieve their questionnaire from the printer while the participant completes the finger tapping task. In this way, the FB scenario could unfold while the participant is engaging in the finger tapping task. If engaging in hand actions impairs motor representations, then this procedure should be able to test the same predictions tested in Experiments 2a and 2b. Moreover, a verbal working memory task (where participants say a word out loud instead of tapping) could be used in place of tying their feet, as this should supply the same general distraction to the participant while leaving their hands free.

7.4.5 Looking following a prompt

In many traditional AL FB studies, the critical looking time window is preceded by an 'I wonder…' prompt which primes children's predictive saccades (Clements & Perner, 1994; Low et al., 2014). Since the present experimental procedure does not include any such prompt to indicate to participants that the agent is about to head toward compartment *X*, it is possible that participants did not recognize what the agent was about to until too late in the critical time window. One possible means by which a prompt could be introduced into the procedure is to modify the agent's phone monologue such that the call ends with, "No, I couldn't find it. Oh! Yes, it's possible that it's in my backpack. Okay, I'll bring it up in a second."

This prompt could trigger participants to expect that the agent will now try to retrieve his or her backpack. Future replications should include this prompt in order to demonstrate whether verbal prompting in active helping FB tasks results in earlier or a longer duration of goal-directed predictive eye movements to either compartment.

7.4.6 Testing children in helping tasks

As the present work has shown, in order to adapt the procedure from Buttelmann et al., (2009) to test adult participants, a multitude of adjustments were necessary. However, the present procedure could be minimally adjusted in order to test children or infants. For example, the compartments could be cubbies on the floor, rather than on high shelves Instead of a sling on his or her arm, the agent could instead have a limp and walk with a cane. When the agent stores the backpack, s/he drops it and uses the cane to slide it into the compartment. The backpacks are then switched, and the researcher demonstrates how the compartments can be locked with a pin (much like the procedure from Buttelmann et al., 2009). When the agent returns, s/he tries to open the empty cubby, and then turns and asks the participant for help. In this way, infants' and children's stages of action and gaze behavior can be assessed and compared with adults' responses opening up an avenue for testing several new developmental hypotheses. Furthermore, by testing the gaze and actions of young children in the present procedure and in earlier helping tasks, the convergent validity of such measures could be assessed.

7.4.7 Distance between compartments

Another variable which may have contributed to the present findings is the distance between participants and the two compartments. Crivello and Poulin-Dubois (2017) attempted to replicate Buttelmann, Carpenter and Tomasello's (2009) findings, but administered the task on a table (15 cm from the infant) instead of on the floor (1 meter from the infant). Crivello and Poulin-Dubois found that 37% of infants in the FB condition selected the box containing the toy (a value which did not differ significantly from chance). By contrast, Buttelmann and colleagues (2009) found that 72% of infants selected this box. In a second experiment, Crivello and Poulin-Dubois placed the boxes at the far corners of a table (50 cm from the infant). Infants' selections remained at chance (58%). One explanation for Crivello and Poulin-Dubois' findings is that infants in their task had less time to decide which box to open.

Future research into helping behavior in FB tasks should consider varying the distance between compartments and participant to determine whether this seemingly minor situational feature plays a role in adult responses. In the present study, the compartments are 2.93 meters from the participant. If 1 meter allows infants sufficient time to decide which box to choose in a FBLoc scenario, then it is unlikely that adults in the present study were given insufficient time to make a decision. It would be interesting to see if these results are able to be replicated when the compartments are closer to participants. If the results successfully replicate (looking and swerving correlate and swerve direction significantly differs between FBLoc and FBID conditions), it would suggest that FB reasoning is less fragile in adults compared to infants.

7.5 Conclusions

This dissertation presents findings which support the predictions from a two-systems account of human mindreading and provides evidence against the notion that mindreading is operated by a single and unified mindreading system. In addition, the results lend support to the existence of an interface between efficient belief tracking processes and motor processes. Experiments 1, 2a and 2b support the notion that a minimal mindreading system defers to motor representations for efficient goal ascription, and that impairments to motor representations has a negative effect on belief tracking, even in relatively straightforward FB circumstances. However, not all results presented herein were as a two-systems account would predict. The results from Experiment 3 were ambiguous, and therefore, future research should consider investigating how physical constraints impact one's eye movements along with alternative means by which motor representations may be disrupted.

The two-systems account is not able to claim explanatory sovereignty over these results. A hybrid of lower-level alternative accounts (e.g., behavior-rules and association biases) can also explain these data. The present findings also fit with a teleological account. Therefore, these findings cannot definitively verify that a two-systems account of human mindreading facilitated these adults' behaviors when helping to reunite an agent with his or her hidden backpack. This dissertation makes an initial step toward an approach dovetailing distinct yet complementary methodologies (i.e., helping tasks and looking time measures). Strict and conceptual replication attempts are needed in order to establish this procedure for future investigation. Pending successful replication, the helping FB task presented in this dissertation may provide a viable means for teasing apart competing accounts of human mindreading in infants, children and adults.

The future of implicit mindreading is far from clear. A recent wave of non-replications shows that implicit mindreading is, at least, unstable. More cautious skeptics have questioned whether implicit mindreading is a real phenomenon. However, while implicit mindreading hangs in limbo, the search for reliable signature limits remains one of enduring relevance. Whatever the system is that coordinates saccadic eye movements during FB scenarios, whether stemming from domain-specific belief tracking or domain-general statistical learning, it is useful to define its conceptual boundaries and hurdles. Thus, future investigation of predictive gaze behavior during circumstances where an agent holds a FB about an object's numerical identity (expansion or compression), should be undeterred by the present replication crisis, and by threats to implicit mindreading. If future evidence continues to indicate that certain patterns of looking predict subsequent actions when responding to an agent's FB, then the logical framework underpinning such looking behaviors must play a constitutive role in any comprehensive account of mindreading. So long as predictive eye movements systematically dissociate from verbal responses (or, in the case of helping tasks, later stages of action), a two-systems account still offers a promising explanans of human mindreading.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*(9), 1109–1116. https://doi.org/10.1038/nn.2182

Ahlheim, C., Stadler, W., & Schubotz, R. I. (2014). Dissociating dynamic probability and predictability in observed actions—an fMRI study. *Frontiers in Human Neuroscience*, *8*, 1-13. https://doi.org/10.3389/fnhum.2014.00273

Allen, J. W. P. (2015). How to help: Can more active behavioral measures help transcend the infant false-belief debate? *New Ideas in Psychology*, *39*, 63–72. https://doi.org/10.1016/j.newideapsych.2015.07.008

Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology*, *106*(3), 1437–1442. https://doi.org/10.1152/jn.00118.2011

Ambrosini, E., Pezzulo, G., & Costantini, M. (2015). The eye in hand: predicting others' behavior by integrating multiple sources of information. *Journal of Neurophysiology*, *113*(7), 2271–2279. https://doi.org/10.1152/jn.00464.2014

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking Ahead: Anticipatory Gaze and Motor Ability in Infancy. *PLoS ONE*, *8*(7), e67916. https://doi.org/10.1371/journal.pone.0067916

Ambrosini, E., Sinigaglia, C., & Costantini, M. (2012). Tie my hands, tie my eyes. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 263–266. https://doi.org/10.1037/a0026570

Apperly, I. (2011). *Mindreaders: The cognitive basis of "Theory of Mind."* New York: Psychology Press

Apperly, I. A, & Robinson, E. J. (1998). Children's mental representation of referential relations. *Cognition*, *67*(3), 287–309. https://doi.org/10.1016/S0010-0277(98)00030-4

Apperly, I. A., & Robinson, E. J. (2001). Children's Difficulties Handling Dual Identity. *Journal of Experimental Child Psychology*, *78*(4), 374–397. https://doi.org/10.1006/jecp.2000.2571

Apperly, I.A, & Robinson, E. (2003). When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children's reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, *85*(4), 297–311.

https://doi.org/10.1016/S0022-0965(03)00099-7

Apperly, Ian A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, *106*(3), 1093–1108. https://doi.org/10.1016/j.cognition.2007.05.005

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

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

Apperly, Ian A., Riggs, K. J., Simpson, A., Chiavarino, C., & Samson, D. (2006). Is belief reasoning automatic? *Psychological Science*, *17*(10), 841–844. https://doi.org/10.1111/j.1467-9280.2006.01791.x

Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, *9*(1), 7–31. https://doi.org/10.1111/j.2044-835X.1991.tb00859.x

Avis, J., & Harris, P. L. (1991). Belief-Desire Reasoning among Baka Children: Evidence for a Universal Conception of Mind. *Child Development*, *62*(3), 460. https://doi.org/10.2307/1131123

Back, E., & Apperly, I. A. (2010). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition*, *115*(1), 54–70. https://doi.org/10.1016/j.cognition.2009.11.008

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological Reasoning in Infancy. *Annual Review of Psychology*, *67*(1), 159–186. https://doi.org/10.1146/annurev-psych-010213-115033

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants Parse Dynamic Action. *Child Development*, *72*(3), 708–717. https://doi.org/10.1111/1467-8624.00310

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, *106*(3), 1382–1407.

https://doi.org/10.1016/j.cognition.2007.07.005

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology*, *19*(2), 149–178. https://doi.org/10.1007/BF00909976

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

Bartsch, K., & London, K. (2000). Children’s use of mental state information in selecting persuasive arguments. *Developmental Psychology*, *36*(3), 352–365. https://doi.org/10.1037/0012-1649.36.3.352

Bem, D. J. (2011). Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*(3), 407.

Birch, S. A. J., & Bloom, P. (2004). Understanding children’s and adults’ limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8*(6), 255–260. https://doi.org/10.1016/j.tics.2004.04.011

Birch, S. A. J., & Bloom, P. (2007). The Curse of Knowledge in Reasoning About False Beliefs. *Psychological Science*, *18*(5), 382–386. https://doi.org/10.1111/j.1467-9280.2007.01909.x

Biro, S., & Leslie, A. M. (2007). Infants? perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science*, *10*(3), 379–398. https://doi.org/10.1111/j.1467-7687.2006.00544.x

Blake, R., Tadin, D., Sobel, K. V., Raissian, T. A., & Chong, S. C. (2006). Strength of early visual adaptation depends on visual awareness. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(12), 4783–4788.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*, 3–22.

Burnside, K., Ruel, A., Azar, N., & Poulin-Dubois, D. (2017). Implicit false belief across the lifespan: Non-replication of an anticipatory looking task. *Cognitive Development*.

https://doi.org/10.1016/j.cogdev.2017.08.006

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–342. https://doi.org/10.1016/j.cognition.2009.05.006

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology*, *131*, 94–103. https://doi.org/10.1016/j.jecp.2014.11.009

Butterfill, S. A., & Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind & Language*, *28*(5), 606–637. https://doi.org/10.1111/mila.12036

Butterfill, S. A., & Apperly, I. A. (2016). Is goal ascription possible in minimal mindreading? *Psychological Review*, *123*(2), 228–233. https://doi.org/10.1037/rev0000022

Butterfill, S. A., & Sinigaglia, C. (2014). Intention and Motor Representation in Purposive Action. *Philosophy and Phenomenological Research*, *88*(1), 119–145. https://doi.org/10.1111/j.1933-1592.2012.00604.x

Butterworth, G., Verweij, E., & Hopkins, B. (1997). The development of prehension in infants: Halverson revisited. *British Journal of Developmental Psychology*, *15*(2), 223–236. https://doi.org/10.1111/j.2044-835X.1997.tb00736.x

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers. *Cerebral Cortex*, *15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Calvo-Merino, Beatriz, Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or Doing? Influence of Visual and Motor Familiarity in Action Observation. *Current Biology*, *16*(19), 1905–1910. https://doi.org/10.1016/j.cub.2006.07.065

Cannon, E. N., & Woodward, A. L. (2008). Action Anticipation and Interference: A Test of Prospective Gaze. *Annual Conference of the Cognitive Science Society*, 981–985.

Cannon, E. N., Woodward, A. L., Gredebäck, G., Hofsten, C. von, & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science*, *15*(1), 35–42. https://doi.org/10.1111/j.1467-7687.2011.01095.x

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

Carlson, S. M., Zelazo, P. D., & Faja, S. (2013). *The Oxford handbook of developmental psychology: Vol. 1. Body and mind*. Oxford, England: Oxford University Press.

Carlson, Stephanie M., Claxton, L. J., & Moses, L. J. (2015). The Relation Between Executive Function and Theory of Mind is More Than Skin Deep. *Journal of Cognition and Development*, *16*(1), 186–197. https://doi.org/10.1080/15248372.2013.824883

Carlson, Stephanie M., Mandell, D. J., & Williams, L. (2004). Executive Function and Theory of Mind: Stability and Prediction From Ages 2 to 3. *Developmental Psychology*, *40*(6), 1105–1122. https://doi.org/10.1037/0012-1649.40.6.1105

Carruthers, P. (2013). Mindreading in Infancy: Mindreading in Infancy. *Mind & Language*, *28*(2), 141–172. https://doi.org/10.1111/mila.12014

Carruthers, P. (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology*, *7*(1), 141–162. https://doi.org/10.1007/s13164-015-0259-y

Carruthers, P. (2017). Mindreading in adults: evaluating two-systems views. *Synthese*, *194*(3), 673–688. https://doi.org/10.1007/s11229-015-0792-3

Casile, A., Caggiano, V., & Ferrari, P. F. (2011). The Mirror Neuron System: A Fresh View. *The Neuroscientist*, *17*(5), 524–538. https://doi.org/10.1177/1073858410392239

Clayton, N. S., Dally, J. M., & Emery, N. J. (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 507–522.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377–395. https://doi.org/10.1016/0885-2014(94)90012-4

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Confais, J., Kilavik, B. E., Ponce-Alvarez, A., & Riehle, A. (2012). On the anticipatory precue activity in motor cortex. *Journal of Neuroscience*, *32*(44), 15359–15368.

Costantini, M., Ambrosini, E., Cardellicchio, P., & Sinigaglia, C. (2013). How your hand drives my eyes. *Social Cognitive and Affective Neuroscience*, *9*(5), 705–711.

Costantini, M., Ambrosini, E., & Sinigaglia, C. (2012). Does how I look at what you're doing depend on what I'm doing? *Acta Psychologica*, *141*(2), 199–204. https://doi.org/10.1016/j.actpsy.2012.07.012

Crivello, C., & Poulin-Dubois, D. (2017). Infants' false belief understanding: A non-replication of the helping task. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2017.10.003

Cross, E. S., Hamilton, A. F. de C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage*, *31*(3), 1257–1267. https://doi.org/10.1016/j.neuroimage.2006.01.033

Cross, E. S., Kirsch, L., Ticini, L. F., & Schütz-Bosbach, S. (2011). The Impact of Aesthetic Evaluation and Physical Ability on Dance Perception. *Frontiers in Human Neuroscience*, *5*, 1-10. https://doi.org/10.3389/fnhum.2011.00102

Cross, E. S., Stadler, W., Parkinson, J., Schütz-Bosbach, S., & Prinz, W. (2013). The influence of visual training on predicting complex action sequences. *Human Brain Mapping*, *34*(2), 467–486. https://doi.org/10.1002/hbm.21450

Csibra, G. (2007). Action mirroring and action understanding: An alternative account. *Sensorimotor foundations of higher cognition*, *22*, 427–451.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *107*(2), 705–717. https://doi.org/10.1016/j.cognition.2007.08.001

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1016/S0364-0213(02)00112-X

Cuevas, I., Plaza, P., Rombaux, P., De Volder, A. G., & Renier, L. (2009). Odour discrimination and identification are improved in early blindness. *Neuropsychologia*, *47*(14), 3079–3083. https://doi.org/10.1016/j.neuropsychologia.2009.07.004

de Klerk, C. C. J. M., Johnson, M. H., & Southgate, V. (2015). An EEG study on the somatotopic organisation of sensorimotor cortex activation during action execution and observation in infancy. *Developmental Cognitive Neuroscience*, *15*, 1–10. https://doi.org/10.1016/j.dcn.2015.08.004

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., & Bekkering, H. (2008). Complementary Systems for Understanding Action Intentions. *Current Biology*, *18*(6), 454–457. https://doi.org/10.1016/j.cub.2008.02.057

Devine, R. T., & Hughes, C. (2014). Relations Between False Belief Understanding and Executive Function in Early Childhood: A Meta-Analysis. *Child Development*, *85*(5), 1777–1794. https://doi.org/10.1111/cdev.12237

Disney, W. (Producer), & Swift, D. (Director). (1961). *The Parent Trap* [Motion Picture]. United States: Walt Disney Productions.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review*, *16*(5), 964–969.

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

Elsner, C., D'Ausilio, A., Gredebäck, G., Falck-Ytter, T., & Fadiga, L. (2013). The motor cortex is causally related to predictive eye movements during action observation. *Neuropsychologia*, *51*(3), 488–492. https://doi.org/10.1016/j.neuropsychologia.2012.12.007

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

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878–879. https://doi.org/10.1038/nn1729

Fizke, E., Butterfill, S., van de Loo, L., Reindl, E., & Rakoczy, H. (2017). Are there signature limits in early theory of mind? *Journal of Experimental Child Psychology*, *162*, 209–224. https://doi.org/10.1016/j.jecp.2017.05.005

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769–771. https://doi.org/10.1038/nature01861

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

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, *15*(1), 95–120. https://doi.org/10.1016/0010-0285(83)90005-1

Fodor, J. A. (1983) *The Modularity of Mind*. Cambridge, Mass: MIT Press.

Fodor, J. A. (1978). Propositional Attitudes. *The Monist*, *61*(4), 501–523.

Gagnon, L., Kupers, R., & Ptito, M. (2013). Reduced taste sensitivity in congenital blindness.

Chemical Senses, *38*(6), 509–517.

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

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501. https://doi.org/10.1016/S1364-6613(98)01262-5

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*(9), 396–403. https://doi.org/10.1016/j.tics.2004.07.002

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., & Keysers, C. (2007). Aplasics Born without Hands Mirror the Goal of Hand Actions with Their Feet. *Current Biology*, *17*(14), 1235–1240. https://doi.org/10.1016/j.cub.2007.06.045

Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey. *Experimental Brain Research*, *71*, 475–490.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

Gilet, E., Diard, J., & Bessière, P. (2011). Bayesian action–perception computational model: Interaction of production and recognition of cursive letters. *PloS One*, *6*(6), e20387.

Gopnik, A., & Meltzoff, N. (1997). *Words, Thoughts, and Theories*. Cambridge, Mass: MIT Press.

Gopnik, A., & Slaughter, V. (1991). Young Children's Understanding of Changes in Their Mental States. *Child Development*, *62*(1), 98–110. https://doi.org/10.1111/j.1467-8624.1991.tb01517.x

Gredebäck, G., & Kochukhova, O. (2010). Goal anticipation during action observation is influenced by synonymous action capabilities, a puzzling developmental study. *Experimental Brain Research*, *202*(2), 493–497. https://doi.org/10.1007/s00221-009-

2138-1

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*(2), 197–206. https://doi.org/10.1016/j.cognition.2009.09.004

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Developmental Science*, *20*(5), 1-15. https://doi.org/10.1111/desc.12445

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, *21*(2), 167–179. https://doi.org/10.1016/S0163-6383(98)90001-7

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. https://doi.org/10.1006/anbe.2000.1518

Harris, P. L. (1994). The Child's Understanding of Emotion: Developmental Change and the Family Environment. *Journal of Child Psychology and Psychiatry*, *35*(1), 3–28. https://doi.org/10.1111/j.1469-7610.1994.tb01131.x

Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A. O. (2005). Transmodal Sensorimotor Networks during Action Observation in Professional Pianists. *Journal of Cognitive Neuroscience*, *17*(2), 282–293. https://doi.org/10.1162/0898929053124893

He, Z., Bolz, M., & Baillargeon, R. (2011). False-belief understanding in 2.5-year-olds: evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*, *14*(2), 292–305. https://doi.org/10.1111/j.1467-7687.2010.00980.x

Hedger, J. A., & Fabricius, W. V. (2011). True belief belies false belief: Recent findings of competence in infants and limitations in 5-year-olds, and implications for theory of mind development. *Review of Philosophy and Psychology*, *2*(3), 429.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, *21*, 101-148. https://doi.org/10.1017/S0140525X98000703

Hogrefe, G.-J., Wimmer, H., & Perner, J. (1986). Ignorance versus False Belief: A Developmental Lag in Attribution of Epistemic States. *Child Development*, *57*(3), 567–582. https://doi.org/10.2307/1130337

Hugdahl, K., Ek, M., Takio, F., Rintee, T., Tuomainen, J., Haarala, C., & Hämäläinen, H. (2004). Blind individuals show enhanced perceptual and attentional sensitivity for identification of speech sounds. *Cognitive Brain Research*, *19*(1), 28–32. https://doi.org/10.1016/j.cogbrainres.2003.10.015

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

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*(6), 1447–1459. https://doi.org/10.1037/0012-1649.43.6.1447

Hughes, C., & Leekam, S. (2004). What are the Links Between Theory of Mind and Social Relations? Review, Reflections and New Directions for Studies of Typical and Atypical Development. *Social Development*, *13*(4), 590–619. https://doi.org/10.1111/j.1467-9507.2004.00285.x

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, *31*(1), 1–22. https://doi.org/10.1017/S0140525X07003123

Hutto, D. D. (2008). *Folk psychological narratives: the sociocultural basis of understanding reasons*. Cambridge, Mass: MIT Press.

Jacob, P. (2013). How from action-mirroring to intention-ascription? *Consciousness and Cognition*, *22*(3), 1132–1141. https://doi.org/10.1016/j.concog.2013.02.005

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, *9*(1), 21–25. https://doi.org/10.1016/j.tics.2004.11.003

Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford University Press.

Jeannerod, M. (1997). *The cognitive neuroscience of action* (Vol. 1997). Oxford: Blackwell, 1997.

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, *21*(17), 6917–6932.

Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *Journal of Neuroscience*, *29*(19), 6234–6238.

Kamawar, D., & Olson, D. R. (2009). Children’s Understanding of Referentially Opaque Contexts: The Role of Metarepresentational and Metalinguistic Ability. *Journal of Cognition and Development*, *10*(4), 285–305. https://doi.org/10.1080/15248370903389499

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, *105*, 211–221. https://doi.org/10.1016/j.anbehav.2015.04.028

Keren, G., & Schul, Y. (2009). Two Is Not Always Better Than One: A Critical Evaluation of Two-System Theories. *Perspectives on Psychological Science*, *4*(6), 533–550. https://doi.org/10.1111/j.1745-6924.2009.01164.x

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking Perspective in Conversation: The Role of Mutual Knowledge in Comprehension. *Psychological Science*, *11*(1), 32–38. https://doi.org/10.1111/1467-9280.00211

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41. https://doi.org/10.1016/S0010-0277(03)00064-7

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*, 522–525.

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, *7*(12), 1299–1301. https://doi.org/10.1038/nn1355

Kingstone, A., Tipper, C., Ristic, J., & Ngan, E. (2004). The eyes have it! An fMRI investigation. *Brain and Cognition*, *55*(2), 269–271. https://doi.org/10.1016/j.bandc.2004.02.037

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/10.1016/0364-0213(94)90007-8

Knudsen, B., & Liszkowski, U. (2012). 18-Month-Olds Predict Specific Action Mistakes Through Attribution of False Belief, Not Ignorance, and Intervene Accordingly. *Infancy*, *17*(6), 672–691. https://doi.org/10.1111/j.1532-7078.2011.00105.x

Kovács, Á. M., Kühn, S., Gergely, G., Csibra, G., & Brass, M. (2014). Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PloS One*, *9*(9), e106558.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The Social Sense: Susceptibility to Others' Beliefs in Human Infants and Adults. *Science*, *330*(6012), 1830–1834. https://doi.org/10.1126/science.1190792

Krauzlis, R. J. (2004). Recasting the Smooth Pursuit Eye Movement System. *Journal of Neurophysiology*, *91*(2), 591–603. https://doi.org/10.1152/jn.00801.2003

Kulke, L., & Rakoczy, H. (2018). Implicit Theory of Mind – An overview of current replications and non-replications. *Data in Brief*, *16*, 101–104. https://doi.org/10.1016/j.dib.2017.11.016

Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (2017). How robust are anticipatory looking measures of Theory of Mind? Replication attempts across the life span. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2017.09.001

Kulke, L., von Duhn, B., Schneider, D., & Rakoczy, H. (2018). Is Implicit Theory of Mind a Real and Robust Phenomenon? Results From a Systematic Replication Study. *Psychological Science*, *29*(6), 888–900. https://doi.org/10.1177/0956797617747090

Kupers, R., Beaulieu-Lefebvre, M., Schneider, F. C., Kassuba, T., Paulson, O. B., Siebner, H. R., & Ptito, M. (2011). Neural correlates of olfactory processing in congenital blindness. *Neuropsychologia*, *49*(7), 2037–2044. https://doi.org/10.1016/j.neuropsychologia.2011.03.033

Land, M., Mennie, N., & Rusted, J. (1999). The Roles of Vision and Eye Movements in the Control of Activities of Daily Living. *Perception*, *28*(11), 1311–1328. https://doi.org/10.1068/p2935

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

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind.'

Trends in Cognitive Sciences, *8*(12), 528–533. https://doi.org/10.1016/j.tics.2004.10.001

Libertus, K., & Needham, A. (2010). Teach to reach: The effects of active vs. passive reaching experiences on action and perception. *Vision Research*, *50*(24), 2750–2757.

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523–531. https://doi.org/10.1037/0012-1649.44.2.523

Low, J. (2010). Preschoolers' Implicit and Explicit False-Belief Understanding: Relations With Complex Syntactical Mastery. *Child Development*, *81*(2), 597–615. https://doi.org/10.1111/j.1467-8624.2009.01418.x

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive Architecture of Belief Reasoning in Children and Adults: A Primer on the Two-Systems Account. *Child Development Perspectives*, *10*(3), 184–189. https://doi.org/10.1111/cdep.12183

Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing How Rabbits Quack and Competitors Act: Limits on Preschoolers' Efficient Ability to Track Perspective. *Child Development*, *85*(4), 1519–1534. https://doi.org/10.1111/cdev.12224

Low, J., & Perner, J. (2012). Implicit and explicit theory of mind: State of the art. *British Journal of Developmental Psychology*, *30*(1), 1–13. https://doi.org/10.1111/j.2044-835X.2011.02074.x

Low, J., & Wang, B. (2011). On the Long Road to Mentalism in Children's Spontaneous False-Belief Understanding: Are We There Yet? *Review of Philosophy and Psychology*, *2*(3), 411-428. https://doi.org/10.1007/s13164-011-0067-y

Low, J., & Watts, J. (2013). Attributing False Beliefs About Object Identity Reveals a Signature Blind Spot in Humans' Efficient Mind-Reading System. *Psychological Science*, *24*(3), 305–311. https://doi.org/10.1177/0956797612451469

Luo, Y. (2011). Do 10-month-old infants understand others' false beliefs? *Cognition*, *121*(3), 289–298. https://doi.org/10.1016/j.cognition.2011.07.011

Luo, Y., & Baillargeon, R. (2005). Can a Self-Propelled Box Have a Goal?: Psychological Reasoning in 5-Month-Old Infants. *Psychological Science*, *16*(8), 601–608. https://doi.org/10.1111/j.1467-9280.2005.01582.x

Lurz, R. W. (2009). *The Philosophy of Animal Minds*. Cambridge University Press.

Lurz, R. W. (2011). *Mindreading Animals: The Debate Over what Animals Know about Other Minds*. Cambridge, Mass: MIT Press.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The Early Development of Inferences about the Visual Percepts of Others. *Child Development*, *45*(2), 357–366. https://doi.org/10.2307/1127956

Michael, J., & Christensen, W. (2016). Flexible goal attribution in early mindreading. *Psychological Review*, *123*(2), 219–227. https://doi.org/10.1037/rev0000016

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & Psychophysics*, *68*(2), 327–337. https://doi.org/10.3758/BF03193680

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-Analysis of the Relation Between Language Ability and False-belief Understanding. *Child Development*, *78*(2), 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*(1), F1–F9. https://doi.org/10.1111/j.1467-7687.2004.00315.x

Monroy, C. D., Gerson, S. A., & Hunnius, S. (2017). Toddlers' action prediction: Statistical learning of continuous action sequences. *Journal of Experimental Child Psychology*, *157*, 14–28. https://doi.org/10.1016/j.jecp.2016.12.004

Moses, L. J., Carlson, S. M., & Sabbagh, M. A. (2005). On the specificity of the relation between executive function and children's theories of mind: Interrelationships among executive functioning, wo. In W. Schneider, R. Schumann-Hengstelar, & B. Sodian (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, wo* (pp. 131-146). Mahwah, NJ: Erlbaum.

Moses, Louis J. (2001). Executive Accounts of Theory-of-Mind Development. *Child Development*, *72*(3), 688–690. https://doi.org/10.1111/1467-8624.00306

Mozuraitis, M., Chambers, C. G., & Daneman, M. (2015). Privileged versus shared knowledge about object identity in real-time referential processing. *Cognition*, *142*, 148–165. https://doi.org/10.1016/j.cognition.2015.05.001

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object Representation in the Ventral Premotor Cortex (Area F5) of the Monkey. *Journal of Neurophysiology*, *78*(4), 2226–2230. https://doi.org/10.1152/jn.1997.78.4.2226

Muthukumaraswamy, S. D., Johnson, B. W., & McNair, N. A. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research*, *19*(2), 195–201. https://doi.org/10.1016/j.cogbrainres.2003.12.001

Mylopoulos, M., & Pacherie, E. (2017). Intentions and Motor Representations: the Interface Challenge. *Review of Philosophy and Psychology*, *8*(2), 317–336. https://doi.org/10.1007/s13164-016-0311-6

Nanay, B. (2013). *Between perception and action*. Oxford: Oxford University Press.

Oktay-Gür, N., Schulz, A., & Rakoczy, H. (2018). Children exhibit different performance patterns in explicit and implicit theory of mind tasks. *Cognition*, *173*, 60–74. https://doi.org/10.1016/j.cognition.2018.01.001

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Patherie, E. (2000). The content of intentions. *Mind & Language*, *15*(4), 400–432.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind.' *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 731–744.

Perner, J. (1991). *Understanding the representational mind.* The MIT Press.

Perner, J. (2014). Commentary on Ted Ruffman's "Belief or not belief:…." *Developmental Review*, *34*(3), 294–299.

Perner, J., & Esken, F. (2015). Evolution of human cooperation in Homo heidelbergensis: Teleology versus mentalism. *Developmental Review*, *38*, 69–88. https://doi.org/10.1016/j.dr.2015.07.005

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344. https://doi.org/10.1016/S1364-6613(99)01362-5

Perner, J., Mauer, M. C., & Hildenbrand, M. (2011). Identity: Key to Children's Understanding of Belief. *Science*, *333*(6041), 474–477. https://doi.org/10.1126/science.1201216

Perner, J., Priewasser, B., & Roessler, J. (2018). The practical other: teleology and its development. *Interdisciplinary Science Reviews*, *43*(2), 99–114. https://doi.org/10.1080/03080188.2018.1453246

Perner, J., & Roessler, J. (2010). Teleology and causal understanding in children's theory of mind. *Causing Human Action: New Perspectives on the Causal Theory of Action*, 199–

228.

Perner, J., & Ruffman, T. (2005). Infants' Insight into the Mind: How Deep? *Science*, *308*(5719), 214–216. https://doi.org/10.1126/science.1111656

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: why alternative naming comes with understanding belief. *Cognitive Development*, *17*(3), 1451–1472. https://doi.org/10.1016/S0885-2014(02)00127-2

Pfurtscheller, G., Neuper, C., Flotzinger, D., & Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left-hand movement. *Electroencephalography and Clinical Neurophysiology*, *103*(6), 642–651. https://doi.org/10.1016/S0013-4694(97)00080-1

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). Young children show the bystander effect in helping situations. *Psychological Science*, *26*(4), 499–506.

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2017). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2017.10.004

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

Priewasser, B., Rafetseder, E., Gargitter, C., & Perner, J. (2017). Helping as an early indicator of a theory of mind: Mentalism or teleology? *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2017.08.002

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

Rakoczy, H. (2017). In defense of a developmental dogma: children acquire propositional attitude folk psychology around age 4. *Synthese*, *194*(3), 689–707. https://doi.org/10.1007/s11229-015-0860-8

Rakoczy, H., Bergfeld, D., Schwarz, I., & Fizke, E. (2015). Explicit Theory of Mind Is Even More Unified Than Previously Assumed: Belief Ascription and Understanding Aspectuality Emerge Together in Development. *Child Development*, *86*(2), 486–502. https://doi.org/10.1111/cdev.12311

Ramsey, R., & Hamilton, A. F. de C. (2010). Triangles have goals too: Understanding action representation in left aIPS. *Neuropsychologia*, *48*(9), 2773–2776. https://doi.org/10.1016/j.neuropsychologia.2010.04.028

Raos, V., Umiltá, M.-A., Murata, A., Fogassi, L., & Gallese, V. (2006). Functional Properties of Grasping-Related Neurons in the Ventral Premotor Area F5 of the Macaque Monkey. *Journal of Neurophysiology*, *95*(2), 709–729. https://doi.org/10.1152/jn.00463.2005

Repacholi, B. M., & Gopnik, A. (1997). Early Reasoning About Desires: Evidence From 14- and 18-Month-Olds. *Developmental Psychology*, *33*(1), 12–21.

Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the future: Three unsuccessful attempts to replicate Bem's 'Retroactive Facilitation of Recall'Effect. *PloS One*, *7*(3), e33423.

Rizzolatti, G, Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (2001). Functional organization of inferior area 6 in the macaque monkey. *Experimental Brain Research*, *71*, 491–507.

Rizzolatti, Giacomo, & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169–192.

Rizzolatti, Giacomo, Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*(9), 661–670. https://doi.org/10.1038/35090060

Rizzolatti, Giacomo, & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*(4), 264–274. https://doi.org/10.1038/nrn2805

Roby, E., & Scott, R. M. (2016). Rethinking the relationship between social experience and false-belief understanding: a mentalistic account. *Frontiers in Psychology*, *7*(1721), 1-7.

Rosnay, M. de, & Hughes, C. (2006). Conversation and theory of mind: Do children talk their way to socio-cognitive understanding? *British Journal of Developmental Psychology*, *24*(1), 7–37. https://doi.org/10.1348/026151005X82901

Rubio-Fernández, P., Jara-Ettinger, J., & Gibson, E. (2017). Can processing demands explain toddlers' performance in false-belief tasks? *Proceedings of the National Academy of Sciences*, *114*(19), E3750. https://doi.org/10.1073/pnas.1701286114

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, *34*(3), 265–293. https://doi.org/10.1016/j.dr.2014.04.001

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does Eye Gaze Indicate Implicit Knowledge of False Belief? Charting Transitions in Knowledge. *Journal of Experimental Child Psychology*, *80*(3), 201–224. https://doi.org/10.1006/jecp.2001.2633

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology*, *30*(1), 87–104. https://doi.org/10.1111/j.2044-835X.2011.02045.x

Russell, J. (1987). "Can we say…?" Children's understanding of intensionality. *Cognition*, *25*(3), 289–308.

Russell, J. (2013). *Agency: Its Role In Mental Development*. Psychology Press.

Saby, J. N., Meltzoff, A. N., & Marshall, P. J. (2013). Infants' somatotopic neural responses to seeing human actions: I've got you under my skin. *PLOS ONE*, *8*(10), e77905. https://doi.org/10.1371/journal.pone.0077905

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

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

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive Load Disrupts Implicit Theory-of-Mind Processing. *Psychological Science*, *23*(8), 842–847. https://doi.org/10.1177/0956797612439070

Schneider, D., Nott, Z. E., & Dux, P. E. (2014). Task instructions and implicit theory of mind. *Cognition*, *133*(1), 43–47. https://doi.org/10.1016/j.cognition.2014.05.016

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218. https://doi.org/10.1016/j.tics.2007.02.006

Scott, R. M., Setoh, P., & Baillargeon, R. (2017). Reply to Rubio-Fernández et al.: Different traditional false-belief tasks impose different processing demands for toddlers., Reply to Rubio-Fernández et al.: Different traditional false-belief tasks impose different processing demands for toddlers. *Proceedings of the National Academy of Sciences of the United States of America*, *Proceedings of the National Academy of Sciences of the United States of America*, *114*(19), E3751–E3752. https://doi.org/10.1073/pnas.1703665114, 10.1073/pnas.1703665114

Scott, R. M. (2014). Post hoc versus predictive accounts of children's theory of mind: A reply to Ruffman. *Developmental Review*, *34*(3), 300–304. https://doi.org/10.1016/j.dr.2014.05.001

Scott, R. M., & Baillargeon, R. (2009). Which Penguin Is This? Attributing False Beliefs About Object Identity at 18 Months. *Child Development*, *80*(4), 1172–1196. https://doi.org/10.1111/j.1467-8624.2009.01324.x

Scott, R. M., & Baillargeon, R. (2014). How fresh a look? A reply to Heyes. *Developmental Science*, *17*(5), 660–664. https://doi.org/10.1111/desc.12173

Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology*, *82*, 32–56. https://doi.org/10.1016/j.cogpsych.2015.08.003

Scott, R. M., & Roby, E. (2015). Processing Demands Impact 3-Year-Olds' Performance in a Spontaneous-Response Task: New Evidence for the Processing-Load Account of Early False-Belief Understanding. *PLOS ONE*, *10*(11), e0142405. https://doi.org/10.1371/journal.pone.0142405

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology*, *22*(2), 353–360. https://doi.org/10.1017/S0954579410000106

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind Eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science*, *325*(5942), 883–885. https://doi.org/10.1126/science.1176170

Setoh, P., Scott, R. M., & Baillargeon, R. (2016). Two-and-a-half-year-olds succeed at a traditional false-belief task with reduced processing demands. *Proceedings of the National Academy of Sciences*, *113*(47), 13360–13365. https://doi.org/10.1073/pnas.1609203113

Sinigaglia, C. (2010). Mirroring and making sense of others. *Nature Reviews Neuroscience*, *11*(6), 449. https://doi.org/10.1038/nrn2805-c2

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.

Sloman, S. A. (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin*, *119*(1), 3–22.

Sommerville, J. A., Hildebrand, E. A., & Crane, C. C. (2008). Experience matters: the impact of doing versus watching on infants' subsequent perception of tool-use events. *Developmental Psychology*, *44*(5), 1249.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*(1), B1–B11.

Song, H., Onishi, K. H., Baillargeon, R., & Fisher, C. (2008). Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition*, *109*(3), 295–315. https://doi.org/10.1016/j.cognition.2008.08.008

Southgate, V., Senju, A., & Csibra, G. (2007). Action Anticipation Through Attribution of False Belief by 2-Year-Olds. *Psychological Science*, *18*(7), 587–592. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Southgate, V. (2013). Do infants provide evidence that the mirror system is involved in action understanding? *Consciousness and Cognition*, *22*(3), 1114–1121.

Southgate, V., & Begus, K. (2013). Motor Activation During the Prediction of Nonexecutable Actions in Infants. *Psychological Science*, *24*(6), 828–835. https://doi.org/10.1177/0956797612459766

Southgate, V., Begus, K., Lloyd-Fox, S., di Gangi, V., & Hamilton, A. (2014). Goal representation in the infant brain. *NeuroImage*, *85*, 294–301. https://doi.org/10.1016/j.neuroimage.2013.08.043

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology*, *45*(6), 1794–1798. https://doi.org/10.1037/a0017197

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition*, *107*(3), 1059–1069. https://doi.org/10.1016/j.cognition.2007.10.002

Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009). Predictive motor activation during action observation in human infants. *Biology Letters*, *5*(6), 769-772.

https://doi.org/10.1098/rsbl.2009.0474

Southgate, V., & Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition*, *130*(1), 1–10.

Sprung, M., Perner, J., & Mitchell, P. (2007). Opacity and discourse referents: Object identity and object properties. *Mind & Language*, *22*(3), 215–245. https://doi.org/10.1111/j.1468-0017.2007.00307.x

Sterr, A., Müller, M. M., Elbert, T., Rockstroh, B., Pantev, C., & Taub, E. (1998). Perceptual correlates of changes in cortical representation of fingers in blind multifinger braille readers. *Journal of Neuroscience*, *18*(11), 4417–4423. https://doi.org/10.1523/JNEUROSCI.18-11-04417.1998

Stevenson, R. L. (1886). *The Strange Case of Dr. Jekyll and Mr. Hyde.* Longmans, Green & co.

Stiner, M. C., Barkai, R., & Gopher, A. (2009). Cooperative hunting and meat sharing 400–200 kya at Qesem Cave, Israel. *Proceedings of the National Academy of Sciences, 106*(32), 13207–13212.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of Beliefs by 13-Month-Old Infants. *Psychological Science*, *18*(7), 580–586. https://doi.org/10.1111/j.1467-9280.2007.01943.x

Surian, L., & Geraci, A. (2012). Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *British Journal of Developmental Psychology*, *30*(1), 30–44. https://doi.org/10.1111/j.2044-835X.2011.02046.x

Surtees, A., Apperly, I., & Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, *150*, 43–52. https://doi.org/10.1016/j.cognition.2016.01.014

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*(2), 452–460.

Teng, S., Puri, A., & Whitney, D. (2012). Ultrafine spatial acuity of blind expert human echolocators. *Experimental Brain Research*, *216*(4), 483–488. https://doi.org/10.1007/s00221-011-2951-1

Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology*, *30*(1), 172–187. https://doi.org/10.1111/j.2044-

835X.2011.02067.x

Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy*, *15*(4), 434–444. https://doi.org/10.1111/j.1532-7078.2009.00025.x

van der Wel, R. P. R. D., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, *130*(1), 128–133. https://doi.org/10.1016/j.cognition.2013.10.004

Vannuscorps, G., & Caramazza, A. (2016). Typical action perception and interpretation without motor simulation. *Proceedings of the National Academy of Sciences*, *113*(1), 86–91. https://doi.org/10.1073/pnas.1516978112

Wan, C. Y., Wood, A. G., Reutens, D. C., & Wilson, S. J. (2010). Early but not late-blindness leads to enhanced auditory perception. *Neuropsychologia*, *48*(1), 344–348. https://doi.org/10.1016/j.neuropsychologia.2009.08.016

Wang, B., Hadi, N. S. A., & Low, J. (2015). Limits on efficient human mindreading: Convergence across Chinese adults and Semai children. *British Journal of Psychology*, *106*(4), 724–740. https://doi.org/10.1111/bjop.12121

Warneken, F. (2013). Young children proactively remedy unnoticed accidents. *Cognition*, *126*(1), 101–108.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303. https://doi.org/10.1126/science.1121448

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271–294. https://doi.org/10.1111/j.1532-7078.2007.tb00227.x

Wasserman, E. A., Wasserman, S. P. of E. P. E. A., & Zentall, T. R. (2006). *Comparative Cognition: Experimental Explorations of Animal Intelligence*. Oxford University Press.

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. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants use statistical sampling to understand the psychological world. *Infancy*, *21*(5), 668–676.

Westra, E. (2017). Spontaneous mindreading: a problem for the two-systems account. *Synthese*, *194*(11), 4559–4581. https://doi.org/10.1007/s11229-016-1159-0

Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: How attentional selection is shaped by ascribing intentions to others. *PLoS ONE*, *7*(9), e45391. https://doi.org/10.1371/journal.pone.0045391

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. https://doi.org/10.1016/0010-0277(83)90004-5

Wong, M., Gnanakumaran, V., & Goldreich, D. (2011). Tactile spatial acuity enhancement in blindness: Evidence for experience-dependent mechanisms. *Journal of Neuroscience*, *31*(19), 7028–7037. https://doi.org/10.1523/JNEUROSCI.6461-10.2011

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34.

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, *22*(2), 145–160. https://doi.org/10.1016/S0163-6383(99)00007-7

Yazdi, A. A., German, T. P., Defeyter, M. A., & Siegal, M. (2006). Competence and performance in belief-desire reasoning across two cultures: The truth, the whole truth and nothing but the truth about false belief? *Cognition*, *100*(2), 343–368. https://doi.org/10.1016/j.cognition.2005.05.004

Yott, J., & Poulin-Dubois, D. (2012). Breaking the rules: Do infants have a true understanding of false belief? *British Journal of Developmental Psychology*, *30*(1), 156–171. https://doi.org/10.1111/j.2044-835X.2011.02060.x

Yott, J., & Poulin-Dubois, D. (2016). Are infants' theory-of-mind abilities well integrated? Implicit understanding of intentions, desires, and beliefs. *Journal of Cognition and Development*, *17*(5), 683–698. https://doi.org/10.1080/15248372.2015.1086771

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Zawidzki, T. W. (2011). How to interpret infant socio-cognitive competence. *Review of Philosophy and Psychology*, *2*(3), 483. https://doi.org/10.1007/s13164-011-0064-1

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, *1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198