Cognitive and Social Contributors in the Development of Intentional Action Understanding

Charlotte de Moor

Under the supervision of Dr Sarah Gerson and Dr Ross Vanderwert

A thesis submitted to the School of Psychology, Cardiff University, for the degree of Doctor of Philosophy

November 2021

Table of Contents

List of Figures

List of Tables

Summary

This thesis focuses on the role of comparison, motor skill, exposure to infant-directed action, and learned actions associated with objects and tools in facilitating development of intentional action understanding.

In chapter 2, we used an eye-tracking paradigm to investigate whether 12-month-olds can generalise understanding about intentional relations from comparison training to visually predict the outcome of a novel action that they had never seen before. Our findings revealed a marginally significant trend such that infants who underwent comparison training made a higher proportion of accurate visual predictions than infants in a control condition.

In chapter 3, we used a similar eye-tracking paradigm along with a motor task to investigate whether 10-month-olds can learn from comparison whilst accounting for individual differences in infants' motor ability. These findings revealed the reverse trend to that found in chapter 2. Also, infants in a control condition with low motor ability made a higher proportion of accurate predictions than infants in the same condition with high motor ability. Findings also imply that comparing an easier tool-use action with a more difficult one is more efficient for facilitating comparison than vice versa.

In chapter 4, we investigated whether mothers' use of infant-directed action (motionese) during toy demonstrations facilitates 10-month-olds' action prediction in our eye-tracking paradigm. More exchanges and less joint contact of toys were associated with infants' accurate predictions.

In chapter 5, we investigated how the meaningful features of objects and tools inform adults' prediction of how an action will unfold. When observing an actor eat from a spoon in an unusual manner, sensorimotor activation measured using EEG was greater than when observing an efficient spoon eating action, signifying greater prediction error.

Together these findings further our understanding of, and raise questions about, how intentional understanding comes about through different cognitive and social processes.

Statement of collaboration, publication, and dissemination of findings

A review of the literature that has previously been published as a book chapter (de Moor & Gerson, 2020) has been re-structured to conform to the accepted norms of a standard narrative thesis. This includes information presented in chapters 1 and 6, as well as background within the introductions of empirical chapters 2 to 4. The study presented in chapter 5 was conducted in collaboration with Dr Ní Choisdealbha as a replication of an original paradigm presented in her doctoral thesis (Ní Choisdealbha, 2015). This involved the use of existing stimuli as well as EEG data processing and analysis scripts adapted from those that had previously been written by Dr Ní Choisdealbha.

The work contained in this thesis has been presented at academic conferences throughout the period of my PhD. An early version of chapter 2 was presented in poster format at the Jean Piaget Society Annual Meeting in June 2018. A subsequent version of chapter 2 as well as plans for chapters 3 and 4 were presented within a symposium at the Society for Research in Child Development (SRCD) conference in March 2019. Most recently, the findings presented in chapter 3 were presented within a symposium at the SRCD conference in April 2021.

Acknowledgements

First, I would like to express my deepest appreciation to Dr Sarah Gerson for your unwavering support and guidance throughout my PhD, and for selflessly taking the time to give me advice whenever I have needed it. You were always there to make sure I was on track to achieve my goals both within and outside of my PhD. I couldn’t have had a better supervisor. I would also like to thank Dr Ross Vanderwert for giving me the opportunity to collect my data alongside yours, as well as the invaluable training in MATLAB, EEG, and E-Prime. You have both made my time at Cardiff University very enjoyable with our lab socials, coffee catchups, and letting me borrow Fudge! I’d also like to thank my advisor, Professor Merideth Gattis, for your insightful discussions during my annual mock vivas. Although these were challenging at first, your perspectives have helped me to think about my research in ways that I’d never considered, as well as helping me to step back and see the bigger picture at times when I was caught up in the detail. I also had a great pleasure of collaborating with Dr Áine Ní Choisdealbha on our adult EEG study. You taught me a great deal about EEG data analysis and were very encouraging and patient in doing so.

I am extremely grateful to all my PhD colleagues for their support and encouragement, particularly those who I have shared an office with. I would especially like to thank Kelsey Frewin for supporting me throughout my PhD. You have been the most welcoming and reassuring friend and colleague since before I even started my PhD. You have supported me in the lab, helped me with data collection, and have always been there to share the highs and lows of postgraduate life together. Thanks also to Kai Thomas for helping me with EEG data collection and analysis, our giggles in the office, as well as your amazing baking skills. I am also grateful to Salim Hashmi for your support as a fellow graduate teaching assistant early on in my PhD when I faced the daunting task of teaching my own seminars. An additional thank you to all my other office mates, Rachel Draper, Jenny Blackmore, Amy Hughes, Jonny Flint, Judith Wallis, Charlie Draper, Becky Saxton, and Charli Robinson for making our office undoubtedly the very best in CUCHDS.

Thanks should also go to my Psychology friends who have supported and assisted me all the way, particularly Natalia Zarzeczna for your advice on processing eye-tracking data, as well as sharing your genius statistical knowledge. Also, a big thanks to Isobel Ward and Lauren Revie for always listening over coffee and becoming friends for life.

I’d like to express huge gratitude to my parents, Delia & Steve, Duncan & Britt, as well as my husband, Amaury, for your unconditional support and encouragement to succeed. I could not have made it this far without my family. My mum has picked up the phone every morning to listen to my PhD woes, whilst Steve has shared his creative ideas with academic poster designs and presentation slides. Amaury has listened to my thoughts and ideas

throughout my academic life and has put up with the times that I’ve cancelled our social plans in order to meet deadlines or overcome setbacks. My dad and Britt have always believed that I could do well and have provided me with the support and encouragement to achieve the very best that I can. Importantly, they always made sure I had the perfect break from my work during trips to see them.

I also owe a great deal of thanks to all the students and interns who have helped with data collection and video coding, as well as bringing excitement and new perspectives to my research. Finally, I’d like to give huge thanks to all the infants, parents, and student participants who volunteered their time to take part in my studies.

Chapter 1

Introduction

The aim of this thesis is to investigate how infants come to understand other people's intentional actions within the first year of life. In particular, I investigate infants' prediction of goal-directed actions, such as reaching for a toy or eating from a spoon, as they unfold. As infants learn to produce actions themselves, they also begin to understand and predict the goal of the same action being performed by another person (e.g., Gerson & Woodward, 2014b, 2014c; Krogh-Jespersen & Woodward, 2018; Sommerville et al., 2005). However, it is known that infants can predict the outcome of actions that they have never produced, such as answering a telephone by bringing it to the ear (e.g., Hunnius & Bekkering, 2010). This raises an interesting question as to how infants come to understand the purpose of actions that are outside of their own motor repertoires. The research presented in this thesis examines how such intentional understanding can come about through experience with comparison between familiar and less familiar actions as well as exposure to infant-directed actions performed by their parents. I also consider whether learned knowledge of actions that are associated with certain objects and tools informs prediction of how the action will unfold.

Much of the literature that focuses on how infants come to understand other people's intentional actions suggests that first person experience performing an action provides an essential basis for intentional action understanding. That is, an infant must have the ability to actively produce an action, such as using cutlery to eat, to be able to understand the purpose of this action when watching another person perform it. Without this active experience, infants cannot recognise the relation between a person and their intentional goal (e.g., Woodward, 1998; Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018). However, other literature suggests that this is not the case (e.g., Csibra, 2008; Gergely et al., 1995; Luo, 2011), or at least there are additional cognitive and social factors that complement self-produced experience for facilitating action understanding (Gerson & Woodward, 2014a; Koterba & Iverson, 2009; Monroy et al., 2017a, b). The aim of this thesis is to investigate how factors other than self-produced experience play a role in facilitating infants' learning about the intentional goals of others' actions. In the current chapter, I first give an overview of the various factors that contribute to the development of intentional action knowledge. Next, I review evidence concerning infants' learning of intentional actions from active experience. I then go on to evaluate how this learning could also come about via comparison processes and use of behavioural cues such as object labelling and action-effects. I also review infant-directed action as a social factor that emerging evidence suggests is helpful for facilitating infants' understanding of other people's actions. Finally, I consider the extent to which these factors may interact with one another in different contexts and give an overview of the aims of investigation within each empirical chapter to follow.

Factors driving early intention understanding

Within our social environments, we are constantly surrounded by other people's actions. Our ability to interpret and understand the purpose of these actions is critical for social functioning. Action understanding refers to an individual's ability to understand another person's intention based on their action. When we understand another person's intention, we focus our attention to the relation between a person and their goal. This is prioritized over other aspects involved in a goal-directed action, such as physical movement or goal location (e.g., Gerson & Woodward, 2013). The development of action understanding during infancy is important for infants' communication within a social environment (e.g., Lizkowski et al., 2006, 2008), learning from others' actions (e.g., Brugger et al., 2007), and working with others during shared actions (e.g., Brandone et al., 2019; Brune & Woodward, 2007; Carpenter, 2009). This understanding of intentional actions emerges within the first year of life and continues to become more sophisticated throughout early childhood (e.g., Kayhan et al., 2019; Paulus et al., 2011). Such understanding is a precursor to fully-fledged theory of mind, which is classically thought to develop at around 4 years of age (Wellman et al., 2008, 2004; Woodward, 1998). Various motor, cognitive, and social factors that facilitate action understanding come into play during development, aiding infants in their understanding of increasingly complex intentional actions that they are exposed to within their social environments (e.g., Brand et al., 2002; Brandone et al., 2019; Cannon et al., 2016; Gerson & Woodward, 2014a, 2014b; van Elk et al., 2008).

Many studies have used visual habituation paradigms to assess infants' understanding of other people's intentional actions. These involve repeatedly presenting an infant with a stimulus whilst recording their looking time as a measure of attention. With repeated exposure, the infant's looking time decreases (i.e., the infant habituates). When presented with a novel stimulus during test trials, looking time increases (i.e., the infant dishabituates). A number of visual habituation studies have found that, as early as six months, when infants begin to produce reaching and grasping actions, they also understand other people's intentional grasping actions (Guajardo & Woodward, 2004; Woodward, 1998, 1999, 2003). In Woodward's (1998) seminal visual habitation paradigm, infants repeatedly viewed a hand reach and grasp one of two toys. During test trials, the toys swapped positions and the same toy was grasped (in its new location), or the alternative toy was grasped (in the old location). Infants looked longer at trials when the alternative toy was grasped, indicating that they understood the relation between the agent and her goal and expected her to grasp the same toy despite its new location. If infants did not represent the action in terms of intentional structure and prioritized the motion involved in the agent's reach, longer looking would be expected during new location trials. More recent advances in eye-tracking technology have enabled researchers to more precisely investigate when infants attend to certain features of observed actions (e.g., Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018). This allows for measurement of visual predictions, shedding light on how infants process and anticipate the outcomes of ongoing actions. For example, when an infant understands the relation between an agent and their goal, he or she may visually attend to an agent's goal before the grasping action is complete. In accordance with earlier habituation research (Guajardo & Woodward, 2004; Woodward, 1998, 1999, 2003), these studies found that infants younger than a year old show intentional understanding by making such predictions.

The first factor I review which is known to drive infants' understanding of others' intentional actions is active experience, also known as the 'humans-first' account. This proposes that infants need to have experience producing an action before they can represent the intentional structure of the same action when observing another person perform it. In other words, infants begin to recognise other people's actions as they begin to produce certain actions themselves (e.g., Gerson & Woodward, 2014b, c; Krogh-Jespersen & Woodward, 2018; Sommerville et al., 2005). One of the first goal-directed actions that infants learn to produce is to grasp a desired object by hand, meaning that they can also recognise other people's hand-grasping actions by at least 6-months of age (e.g., Woodward, 1998). Although young infants understand simple grasping actions, they struggle to make sense of inanimate goal-directed actions, such as when the agent's hand is masked by a glove (Guajardo & Woodward, 2004) or when a tool (such as a rod) is used to grasp an object (Woodward, 1998). Only older infants (12-month-olds as compared with 7- and 9-month-olds) in Guajardo and Woodward's (2004) research could infer the relation between a gloved agent and her goal. However, when infants were shown that the glove was attached to a person (Guajardo & Woodward, 2004), as well as when shown that a claw tool was operated by a person (Hofer et al., 2005) or given other behavioural cues to agency (Biro & Leslie, 2007), infants were more successful at inferring the relation between the agent and their goal. This implies that, as well as active experience, other cues are needed for infants to understand actions that they have not yet produced in early development.

A line of research that provides an alternate perspective to the humans-first account is the 'all-agents' account. This suggests that infants have an innate ability to understand the goal-directedness of any human or non-human agent. Infants can understand observed examples of unfamiliar intentional actions if provided with certain behavioural cues, such as self-propulsion and action outcomes (e.g., Csibra et al., 1999; Gergely et al., 1995). A number of studies that have assessed infants' understanding of non-human actions have involved habituating infants to an animated agent such as a self-propelled circle (Gergely et al., 1995) or a box (Luo, 2011; Luo & Baillargeon, 2005) taking either a rational or non-rational path around an obstacle to approach a goal. These animations have used varying levels of agentive cues such as greeting exchanges and two dimensional (Gergely et al., 1995) versus three dimensional (Csibra, 2008) displays to assess the richness of cues that infants require in order to view the animated agent as intentional. These studies are described in more detail later in the cue-based literature section below. In the same section, I also consider the extent to which the humans-first and all-agents views may not be directly contradictory. For example, Brand et al. (2015) and Gredebäck and Melinder (2010) provide support for both views, reasoning that infants can process the actions of all agents, including those of non-human objects, with the help of their own motor expertise to reinforce their interpretation of intentional structures.

Observational comparison could also help infants understand actions that they have never performed themselves. By comparing familiar actions with novel actions, infants could come to understand the intentional relations involved in never-before-seen actions. For example, an infant could draw upon his or her experience with a familiar goal-directed grasp (e.g., grasp a ball by hand) to understand a goal-directed tool-reach (e.g., grasp a ball using a claw) when presented with a context that encourages comparison of the goals of these actions. These actions both share the same goal, but a different means is used to achieve the goal in each (hand vs tool). Research has found that both active (Gerson & Woodward, 2012) and observational (Gerson & Woodward, 2014a) comparison can facilitate infants' representation of a complex novel action in terms of its intentional structure.

In addition to these cognitive factors, social factors are important for supporting infants' learning of action structures. As described above, infants must learn to parse intentional units of an action in order to understand its meaning. To help infants with this process, parents naturally tend to modify their actions when demonstrating objects to infants. This helps to maintain the infant's attention whilst highlighting the structure and purpose of the action. This is known as infant-directed action or 'motionese' (Brand et al., 2002, 2007). Motionese is derived from the language phenomenon of 'motherese,' which refers to modifications such as exaggerated intonation and higher pitch that adults naturally use when speaking to an infant relative to an adult (e.g., Fernald & Kuhl, 1987; Papoušek et al., 1991). Brand et al. (2002) first identified motionese in their study in which mothers demonstrated objects to either a familiar adult or their infant (age 6- to 8- or 11- to 13-months). Much like motherese in the auditory domain, mothers exaggerated their actions on a number of dimensions such as enthusiasm, range of motion, and repetitiveness when demonstrating to their infant relative to an adult. These results were replicated in a later study with both mothers and fathers (Rutherford & Przednowek, 2012). Some evidence supports the notion that motionese assists infants' processing and understanding of actions by capturing infants' attention (e.g., Brand & Shallcross, 2008) and promoting infants' success at producing demonstrated actions (e.g., Fukuyama et al., 2015; van Schaik et al., 2019). Infant-directed action is briefly touched upon again later in this chapter and the literature is evaluated in more detail in chapter 4.

Whilst various cognitive and social factors are important in aiding infants' action understanding, I argue that active experience provides a crucial underpinning at the origins of action understanding. However, factors additional to active experience are needed to facilitate understanding of never produced actions. The studies in this thesis investigate the role of cognitive and social factors, including comparison and motionese, that build upon the role of active experience in facilitating the development of action understanding. The next sections of this chapter review relevant literature and theory, providing background for the empirical chapters to follow.

Motor factors: The role of active experience

As mentioned previously, the active experience or 'humans-first' account proposes that infants need to have experience producing an action before they can represent the intentional structure of the same action when observing another person perform it. To establish a causal effect of active experience in facilitating action understanding, studies have trained 3-month-old infants, who cannot yet efficiently reach and grasp objects, to use Velcro mittens to manipulate a ball and a bear before assessing their understanding of a reach for one of these toys in a visual habituation paradigm (Gerson & Woodward, 2014b; Sommerville et al., 2005). In Sommerville et al.'s (2005) study, infants in a training condition received mitten training before the habituation procedure, whilst infants in a watch-first condition underwent the habituation procedure first, followed by mitten training. Infants in the reach-first condition looked longer on new goal trials, meaning that they recognised the relation between the agent and her goal. In contrast, infants in the watch-first condition looked equally to both new goal and same goal trials. Gerson and Woodward (2014b) later assessed whether the same mittened action understanding could come about through observational experience as well as active experience. In an observation condition, infants of the same age and gender were matched to those in an active condition. Infants in the observation condition observed the experimenter move each toy for approximately the same amount of time that the active infant had, thus matching the active infant's level of activity. This way, the authors could examine whether varied amounts of observational experience could also facilitate action understanding. Infants who had undergone active training looked longer on new goal trials, whereas infants who received only observational training or no training did not significantly differ in looking time between new goal and same goal trials. In addition to this, infants in each of these studies (Gerson & Woodward, 2014b; Sommerville et al., 2005) who had engaged in longer manual and visual contact during active training also looked longer to new goal trials. In other words, infants who had more experience producing reaching actions showed stronger evidence of goal recognition. From these findings, it seems that observational experience was not related to action understanding. Therefore, at this early stage of development when infants are less proficient at producing goal-directed actions, observational experience does not seem to facilitate action understanding to the same extent as motoric experience.

Within their natural experiences, infants are faced with a wide variety of complex actions. Tool-use actions, such as using a cane or pulling on a cloth to obtain an out-of-reach object, have been used as examples of complex two-step actions that come about after mastering simple grasping actions. Intervention studies have found that active experience performing tool-use actions facilitates action understanding (e.g., Gerson et al., 2015a, 2015b; Sommerville et al., 2008, 2005). This research has involved training infants to actively perform a tool-use action before measuring intentional understanding when observing another person perform the same action. For example, Sommerville et al. (2008) trained 10-month-olds to use a cane to retrieve a toy. Infants' success during training related to understanding the goal of the tool-use action in a visual habituation paradigm. Similar findings occurred in Gerson et al.'s (2015b) study that trained 8-month-olds to perform a cloth-pull action to retrieve a toy situated on the end of a cloth. In addition, the authors found that infants who were less successful during training attended to the proximal relation between the agent and the cloth that she was directly touching, implying a lack of understanding of the agent's overarching goal. As outlined previously, this aligns with the development of goal-directed action understanding. As infants become more proficient at performing an action, they undergo an attentional shift from attending to proximal means (i.e., the tool itself) to distal goals (Willatts, 1999). In a second between-subjects experiment, Gerson et al. (2015b) introduced an observation condition in which infants' level of experience was matched with those in the active condition (i.e., the matched infant in the observation condition observed the experimenter perform the cloth-pull action for the same duration that their active partner was exposed to the task). The level of exposure to the cloth-pull action in the observation condition did not have a significant effect upon looking time between old location and new location trials. Therefore, observational experience of this tool-use action was not beneficial in informing 8-month-olds' action understanding. Active experience performing the action was needed before infants could understand the intention of another person performing the same tool-use action.

Taken together, these studies have used active training to mimic natural development in a controlled way. As infants gain experience performing intentional actions, their understanding of other people's complex actions improves. From these intervention studies, we can draw conclusions about the causal factor of motor experience in facilitating action understanding. These results are consistent with neurophysiological findings, which suggest that execution and observation of matching actions involve mirrored neurocognitive representations. This is known as sensorimotor processing, when sensory input elicits a motor response. When observing an action that we know how to perform, our own sensorimotor system is also activated. Mu and beta desynchronisation patterns (desynchronisation implies more activation of the sensorimotor system) in the electrophysiological (EEG) signal response to an observed action are similar to those involved during movement. This shared neural representation, also known as 'mirroring,' has been observed in the sensorimotor system of both adults (e.g., Gerson et al., 2017; Järveläinen, et al., 2004) and infants (e.g., Cannon et al., 2016; van Elk et al., 2008). For example, in a study using EEG, van Elk et al. (2008) found that 14- to 16-month-olds' amount of crawling experience was related to mu and beta desynchronisation when observing videos of other infants crawling as compared with walking. The authors concluded that 14- to 16-month-olds who have more experience with crawling than walking have more established representations of crawling within their motor repertoires. This experience-based understanding has not only been found for locomotor experience but is also evident for manual actions in correlational EEG studies. Cannon et al. (2016) investigated specific changes in the reach-grasp competence of 9-month-olds, finding that more motorically efficient reach-grasp skills (e.g., hand pre-shaping and faster latency to complete the grasp) were related to greater mu desynchronisation during observed grasping actions.

Further to these correlational findings, an intervention study by Gerson et al. (2015a) also provided support for the role of motor experience as a causal factor influencing motor activity in the perception of others' actions. 10-month-olds received week-long training to perform a novel action that resulted in a sound effect. They also gained observational experience with another unfamiliar action that resulted in a different sound effect. Sensorimotor activity measured via EEG revealed that greater mu desynchronisation was present when infants heard the sound effect associated with the trained action, relative to the observed action that they had not motorically produced, or a novel sound. Therefore, the sensorimotor system was activated when perceiving the effect of a motorically familiar action. Like other studies concerning action understanding or sensorimotor system activation (e.g., Cannon et al., 2016; Sommerville et al., 2008; van Elk et al., 2008), individual differences in infants' degree of learning was related to the degree of mu desynchronisation such that those infants who better learned to perform the action showed more sensorimotor activity following training.

Findings from these EEG studies suggest that the sensorimotor system is modulated dependent on experience during the development of motor skills. This provides further support to suggest that active rather than observational experience is critical for infants' development of action understanding. Infants draw upon their familiarity with self-produced actions to predict the outcome of observed actions. Although active experience is beneficial, additional factors could build upon active experience to allow generalisation of understanding to actions outside of our motor repertoires. These are discussed in the following sections of this chapter.

The need for other factors: Cue-based literature

An alternative perspective suggests that infants have an innate ability to understand the goal-directedness of any human or non-human agent. This is also known as the 'all-agents view,' contrasting with the active experience or 'humans-first' view that suggests infants require self-produced experience to be able to understand intentional actions (Luo, 2011). As mentioned previously, the 'all-agents' perspective is of the view that infants can learn from observed examples of unfamiliar intentional actions if provided with certain behavioural cues, such as self-propulsion and action outcomes. The all-agents and humans-first views are often presented as alternatives, but in reality, infants may rely on various factors (independently or together) to understand intentional actions in different contexts. For example, an infant may rely more so on their sensorimotor system when observing a human action, whilst using behavioural cues to understand a non-human action (e.g., Brand et al., 2015).

In Gergely et al.'s (1995) classic paradigm, 12-month-olds in a rational action condition were habituated to an animated circle (the agent) jumping over an obstacle to get to a bigger circle. This animation exhibited cues including expansion and contraction in a 'greeting exchange' of each circle, self-propulsion to indicate agency, and an equifinal outcome to indicate the goal. In a non-rational action condition, infants were habituated to the same action sequence, except the obstacle was not in the same position, meaning that the circle's jumping action was not necessary. In test trials, all infants viewed a new action in which no obstacle was present, and the small circle approached the large circle using a direct pathway. They also viewed the old jumping action (as in habituation) without the obstacle present. Infants in the rational action condition dishabituated significantly more when seeing the same jumping action (now not rational) than the new direct path action (now rational). However, infants in the non-rational action condition showed the reverse, as they dishabituated more when seeing the novel direct path action. Therefore, infants who were not habituated to a rational action did not view the circle as being a rational, intentional agent. This research demonstrates that 12-month-olds can use behavioural cues exhibited by a non-human agent to infer its actions as goal-directed and expect the agent to use the most rational means to achieve its goal. These results are further supported by Luo and colleagues' findings that infants as young as 3-months of age understood the goals of a self-propelled box (Luo, 2011; Luo & Baillargeon, 2005). This suggests that agency cues can enable infants to recognise goals of agents earlier than expected by the experience-based, or ‘humans-first,’ perspective. It also seems that younger infants can represent non-human actions as intentional when provided with richer cues to agency. Csibra (2008) presented 6.5-month-olds with three-dimensional stimuli in which a box approached a target object using various routes (i.e., equifinal variations) to avoid an obstacle. As the stimuli were three-dimensional, the agent became momentarily occluded by the obstacle after passing it. When given this combination of cues, infants showed an understanding of the intentional structure of the agent’s action by dishabituating to test trials in which the agent took a non-rational path to approach the target object (but see Csibra et al., 1999 for counter evidence).

Biro and Leslie (2007) support the notion that young infants require rich cues to agency in a replication of Woodward’s (1998) paradigm (described in the first section of this chapter). In support of the all-agents view, they reasoned that adding behavioural cues to the non-human tool-reach action would enable young infants to interpret this complex action as goal-directed. These cues involved the goal object being poked using a rod from different angles (i.e., equifinal variations), self-propulsion of the rod, and lifting and moving the object (i.e., action effect). When infants were familiarised with all three of these cues, 6-, 9-, and 12-month-olds better understood the relation between the agent and their goal. When provided with two of these cues (self-propulsion and action effect), 12- and 9-month-olds encoded the agents’ goal, however 6-month-olds did not.

Although this habituation literature informs us of infants’ retrospective understanding of a visual scene, Southgate et al. (2010) pointed out the importance of measuring infants’ ability to make real-time predictions of how another person’s action will unfold. These predictions enable individuals to prepare an appropriate action in response, particularly during a collaborative activity. In a similar vein to online measures of visual predictions using eye-tracking (e.g., Cannon & Woodward, 2012), Southgate et al. (2010) used EEG to investigate whether motor activation is present when infants can predict the outcome of a goal-directed action. In an occlusion condition, 9-month-olds observed an experimenter’s hand reach down to the floor of a stage. The action outcome was not visible, meaning that infants saw part of an action and the outcome needed to be inferred (i.e., implied that the hand was contacting an object behind the occluder). In a mimed condition, infants viewed the same action without the presence of an occluder, thus the action was not goal-directed as no object was present when the gesture was carried out. In each of these conditions, infants viewed the hand in both a grasping and a back-of-hand posture. Southgate and colleagues found that mu desynchronisation was only present during the observation of a grasping hand action behind an occluder, as infants could predict that the hand was reaching for a hidden object. The authors concluded that motor activation is involved in the process of predicting an ongoing action, explaining the presence of mu desynchronisation when predicting a goal, but no desynchronisation during actions that were not goal-directed. This supports prior eye-tracking findings involving 12- and 14-month-olds' proactive gaze shifts when observing a reaching hand place objects in containers, as opposed to reactive gaze shifts when observing a fist posture move between objects, or when objects were self-propelled (Falck-Ytter et al., 2006; Gredebäck et al. 2009). Therefore, it seems that items present in the visual scene (e.g., occluder or container), as well as the function of movement (e.g., hand configuration for human agents or rational paths taken by non-human agents) serve as cues to the goal-directedness of an action and allow infants to make online predictions about ongoing actions.

Brand et al. (2015) bring the all-agents and humans-first views together, suggesting that infants can process the actions of all agents, whilst motor development further reinforces more specific interpretation of contexts and goals. Using Gergely et al.'s (1995) classic paradigm, Brand et al. (2015) investigated whether crawling and non-crawling 6- and 9-month-olds would interpret a circle's actions as goal-directed. The authors found that crawling infants of both ages interpreted non-human movements as goal-directed, but non-crawling infants did not. Brand et al. (2015) reasoned that the understanding of an agent navigating around obstacles is bolstered by infants' own ability to navigate obstacles when crawling. Whereas these findings support the claim that infants can interpret the actions of non-human agents, this still demonstrates the role of experience in infants' understanding of intentional actions.

Gredebäck and Melinder (2010) also provide support for both the all-agents and humans-first views in their dual process account of action understanding. 6- and 12-month-olds viewed feeding actions being performed in a rational or non-rational manner. The rational action involved one individual spoon feeding another by bringing the food to the other's mouth, whereas during the non-rational action the food was brought to the back of the hand. 12-month-olds made proactive visual fixations to the mouth before the food arrived at the mouth, whereas 6-month-olds made reactive fixations. This was dependent on 12-month-olds' life experience with being fed (from parent reports), as infants with more experience fixated to the mouth earlier than infants with little experience. Also, pupil dilation was found in response to non-rational feeding actions in both age groups. This implies that all infants found the non-rational action unusual. From this, the authors concluded that reasoning about rational actions requires less motor experience than does goal prediction. Although the all-agents view argues against a role of motor development in facilitating understanding of other's actions, it seems that motor factors, as well as behavioural cues, may facilitate infants' capacity to make sense of unfamiliar intentional actions. Below I discuss the role of comparison as a further observational factor that facilitates the development of intentional understanding. I hypothesise that this cognitive factor can also help infants bridge the gap between actions they can produce and more complex or unfamiliar actions.

Cognitive factors: Comparison processes

Even as adults, we cannot perform every action that we understand. We are eventually able to learn about actions through observational experience alone (e.g., Gerson et al., 2017). If infants could only understand others' actions that they have produced themselves, then they would be very restricted in what they would be able to learn through observing others. The cue-based section above cites some cases in which infants are able to learn through observation (e.g., phone to ear predictions in Stapel et al., 2010). Although active experience is an important factor in supporting action understanding, it is likely that additional observational factors come into play as infants become more proficient at performing goal-directed actions (e.g., Hunnius & Bekkering, 2014). One way that infants may move beyond understanding only actions that they have learned to perform themselves is through a process of comparison between familiar and novel exemplars.

Gentner and Medina's (1998) structure mapping theory suggests that learners can generalise relational information when they compare two instances that share a similar structure. This process of comparison could come about in contexts that provide the opportunity for infants to compare their own actions with another person's (Gentner & Medina, 1998; Meltzoff, 2007). Recall our example mentioned in the first section of this chapter, where an infant could draw upon his or her experience with a familiar goal-directed grasp (e.g., grasp a ball by hand) to understand a goal-directed tool-reach (e.g., grasp a ball using a claw) when presented with a context that encourages comparison of the goals of these actions. These actions both share the same relational structure in that the relation between the agent and the goal is the same. The only difference between these is the means by which the goal is achieved (hand vs tool). Literature has demonstrated that active (Gerson & Woodward, 2012) and observational (Gerson & Woodward, 2014a) comparison promotes understanding of intentional tool-use actions for infants as young as 7-months old (Gerson & Woodward, 2012, 2014a). In experimental training conditions, these studies have provided infants with the opportunity to compare their own familiar hand-grasp with another person's unfamiliar claw-reach action. Infants in control conditions did not have the same opportunity to compare between their own and another's actions. Infants' intentional understanding was measured using goal imitation procedures, in which infants could choose between the toy that the experimenter had previously reached for (i.e., the experimenter's goal), or a toy that had not been acted upon. Infants in experimental comparison conditions selected the experimenter's previous intentional goal significantly more often than infants in control conditions, demonstrating an intentional understanding that had come about through a process of comparison. This understanding develops through both observational comparison and active physical comparison between an infants' own grasp and a claw grasp when passed a toy using a claw. Overall, active comparison (i.e., being passed a toy using a claw) seems to be more beneficial than observational comparison (i.e., observing an experimenter passing toys to another experimenter using a claw). However, infants are still able to learn from observational comparison, especially when additional cues to compare between intentional actions, such as goal labelling, are present (Gerson & Woodward, 2014a). This literature is described in more detail in chapters 2 and 3, in which we investigate how infants' intentional understanding may come about through a process of observational comparison. In chapter 2, we investigated whether 12-month-olds, who are at the cusp of understanding and performing tool-use actions, were able to generalise intentional understanding from observational comparison training to understand a novel claw-reach action that they had never seen before. Chapter 3 used a similar paradigm to investigate whether younger 10-month-olds can also learn from observational comparison whilst taking individual differences in motor ability into account (see previous section for a discussion of the link between motoric learning and action understanding, e.g., Brand et al., 2015; Gredebäck & Melinda, 2010).

Social factors: Infant-directed action

When learning about new actions, infants see a continuous stream of movement and must learn to parse intentional units of an action in order to understand its meaning. For example, an infant may watch another person paint a picture by preparing the paint palette, applying the paint to the brush, and then applying the paint to paper. When demonstrating this series of actions to an infant, an adult may exaggerate these actions in a manner such that the infant may better understand the intentional sub-steps (e.g., applying paint to the brush vs brush to the paper). This is known as infant-directed action, or 'motionese' (Brand et al., 2002, 2007), derived from the language phenomenon of 'motherese' (e.g., Fernald & Kuhl, 1987, Papoušek et al., 1991). As mentioned earlier in this chapter, both mothers and fathers have been found to exaggerate their actions on a number of dimensions including enthusiasm, range of motion, and repetitiveness when interacting with infants relative to adults (Brand et al., 2002, 2007; Rutherford & Przednowek, 2012). Infants also prefer infant-directed action over adult-directed action (Brand & Shallcross, 2008). Modification in infant-directed action is associated with variation in infants' attention and subsequent exploratory behaviours (Koterba & Iverson, 2009) and is also thought to help infants understand actions (Brand et al., 2002, 2007). However, the potential relation between infants' understanding of intentional actions and their exposure to infant-directed actions has not yet been investigated directly. Chapter 4 evaluates the motionese literature in greater detail and also investigates how mothers' varying engagement in infant-directed action may facilitate 10-month-olds' understanding of complex intentional actions.

Conclusions and aims of investigation

In this chapter, I have presented motor, cognitive and social factors that contribute to infants' development of intentional action understanding. These factors complement one another and may come into play during different stages of development, as well as in different contexts. I began by addressing evidence suggesting that active motor experience is a critical factor in explaining how the development of action understanding comes about. This active experience or 'humans-first' view suggests that infants need to have experience producing actions before they can understand other people's (e.g., Woodward, 1998, 1999, 2003). Although motor skills are an important factor, they cannot fully explain the process by which infants come to understand other people's intentions, such as how infants understand actions that they have not self-produced (e.g., Gergely et al., 1995). Another factor that builds upon experience with self-produced actions is the process of comparison. Comparison may be useful in contexts where infants can compare their own actions with another person's, both actively (e.g., Gerson & Woodward, 2012) and through observation alone (e.g., Gerson & Woodward, 2014a). For example, at mealtime, an infant might compare their own action of picking up food by hand with their parent's action of using a fork. As touched upon earlier in this chapter, observational comparison can be supported using linguistic labelling of objects (Gerson & Woodward, 2014a). It is an open question as to whether consistent cues other than linguistic labelling could facilitate comparison between motorically familiar and unfamiliar actions. The process of comparison could also promote generalisation of understanding to novel actions that involve a different goal (Gentner & Medina, 1998). By familiarising infants with multiple examples of actions that share the same goal as a consistent cue, the first two empirical studies in this thesis investigate this question using an eye-tracking paradigm to measure visual prediction of a novel action. In the first empirical chapter (chapter 2), we examined whether experience with comparison is associated with 12-month-olds' ability to accurately predict the outcome of a novel action. As discussed earlier in this chapter, motor skill and observational cues seem to interact with one another, as infants with less efficient motor skills rely more so on observational cues to intention than infants with more efficient motor skills (e.g., Biro & Leslie, 2007; Csibra, 2008). In chapter 3, we replicated the paradigm used in chapter 2 to examine whether comparison training is associated with accurate predictions in a younger sample of 10-month-old infants. We also accounted for individual differences in infants' ability to make accurate predictions based on their motor skill.

Moving from a focus on cognitive factors that may facilitate action understanding to social factors; it is known that motionese increases infants' attention to an action (e.g., Brand & Shallcross, 2008) and modulates infants' subsequent exploration of the demonstrated object (e.g., Koterba & Iverson, 2009). It is assumed that motionese helps to facilitate infants' processing and understanding of actions, however evidence to support this notion is rare. Exposure to a high level of motionese could provide infants with a basis of intentional understanding to facilitate generalisation to novel actions. The study presented in chapter 4 attempted to address this question by measuring mothers' use of motionese whilst demonstrating toys to their 10-month-old infants. We assessed whether exposure to motionese was associated with infants' accurate predictions of the novel action in our eye-tracking paradigm.

Beyond goal-directedness alone, it is known that the type of object or tool that is being acted upon informs an individual's prediction of the action that will be performed. That is, certain objects and tools are associated with certain actions, such as using a spoon to eat food (e.g., Creem & Proffitt, 2001; De Bellis et al., 2016; van Elk et al., 2010). Individuals who are motorically experienced with this action hold a representation of the conventional way in which the action is performed (e.g., how to grasp and manipulate a spoon). This can be referred to as semantic knowledge. The onset of an action, such as the way in which an object or tool is grasped, informs an individual's prediction of the necessary motor commands to complete a functional action (e.g., Kilner et al., 2007; Ní Choisdealbha, 2015). In chapter 5, we investigated whether the meaningful features of objects and tools inform adults' action predictions within the sensorimotor system. In doing so, we assessed whether the sensorimotor system differentiates between stages of actions (i.e., onset and completion) that are carried out in a motorically conventional versus unconventional manner.

Together, the studies in this thesis aim to improve our knowledge of how intentional understanding develops through processes of comparison, motionese, and learned knowledge of actions. To preview our findings, results in chapter 2 revealed that observational comparison between familiar and less familiar actions may help to facilitate 12-month-olds' ability to understand and predict the outcome of a never-before-seen action. However, results in chapter 3 raise questions regarding the extent to which comparison is helpful for younger 10-month-olds with varying levels of motor skill. Findings in chapter 4 suggest that toy exchanges within motionese facilitate infants' accurate predictions of a novel action. Finally, findings in chapter 5 confirm that experienced individuals predict how a person will act based on learned knowledge of actions that are associated with tools and objects. Although various factors that contribute to the development of intentional action understanding have been investigated separately, the general discussion raises interesting questions regarding the extent to which these factors play a combined role in supporting intentional understanding.

Chapter 2

Infants' comparison and prediction of others' goal-directed actions

Abstract

By 12-months, an early understanding of others' intentional behaviour is emerging, and infants are at the cusp of understanding and performing tool-use actions. At around this age, infants look ahead to the end-goal of a person's incomplete reach when they understand the person's intentional goal (Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018). Active experience performing actions provides infants with an important underpinning for understanding other people's intentions (e.g., Woodward, 1998; Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018). However, infants also understand the intentional structure of some actions that they have never produced (e.g., Hunnius & Bekkering, 2010). We hypothesised that one factor which can facilitate infants' understanding of actions that they have never performed is comparison between familiar and novel actions (Gerson & Woodward, 2012). Study 1 examined whether observational experience comparing a familiar hand-reach with less familiar tool-reach actions supports 12-month-olds' understanding of a novel claw-reach action that they cannot typically understand and predict on their own. In our training intervention, infants in a comparison condition (*N* = 20) viewed a familiar hand-reach action followed by tool-reach actions that shared the same goal (i.e., toy object). Infants in a control condition (*N* = 20) viewed a tool-reach action without the opportunity to compare this with a familiar hand-reach. Test trials then required infants to generalise their intentional knowledge from training to visually predict the outcome of the novel claw-reach action. Our findings revealed a difference between conditions that was marginally significant. Infants who received comparison training made a higher proportion of accurate predictions (on 67% of trials) than controls (on 43% of trials). Infants in the comparison condition systematically made accurate predictions. However, accurate versus inaccurate predictions made by infants in the control condition did not significantly differ from chance. This evidence indicates that by comparing different ways of achieving the same goal, infants can generalise their intentional understanding to predict the outcome of a new action without knowing how to perform it themselves.

Introduction

Within their social environments, infants are surrounded by other people's actions that they do not understand the purpose of. A challenge that infants face in learning to understand actions is to extract the relation between a person and their intentional goal. A fundamental theory for the development of intentional action understanding is the active experience or 'humans-first' account. This proposes that infants need to have active experience producing an action before they can represent the intentional structure of the same action when observing another person perform it. Literature supporting this account has trained infants to perform new actions, then has subsequently assessed their understanding of the same action using visual habituation paradigms (e.g., Gerson & Woodward, 2014a, 2014b; Gerson et al., 2015b; Sommerville et al., 2005, 2008), as well as eye-tracking paradigms measuring infants' visual anticipation of an unfolding action (e.g., Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018). Findings from these studies confirm that active experience better facilitates infants' action understanding than does observational experience or no prior experience, particularly in very young infants (e.g., 3-month-olds) whose own skill at performing actions is beginning to emerge (Gerson & Woodward, 2014b; Sommerville et al., 2005).

In a seminal study, Woodward (1998) showed that infants as young as 6-months of age demonstrated an understanding of goal-directed reaches in a habituation paradigm. That is, they looked longer when they observed a hand reach for a new object (i.e., an unexpected goal) as opposed to when a hand reached for a previously familiarised goal object. More recently, a study by Cannon and Woodward (2012) used an eye-tracking paradigm to replicate Woodward's (1998) original paradigm. In this study, 11-month-olds were familiarised to videos of a hand reaching and grasping one of two objects. The objects then swapped locations, and in a single test trial, the hand made an incomplete reach between the two objects. Using eye-tracking, Cannon and Woodward (2012) measured whether infants looked from the hand to the goal object that they had been familiarised with, now in its new location (i.e., an accurate goal prediction). Looks from the hand to the non-familiarised object were inaccurate, as this demonstrated that infants did not understand the person's intentional goal and were simply anticipating that the hand would reach to the familiarised location. In a separate condition, infants observed the same events, except a less familiar claw-tool was used to reach and grasp the objects. The authors found that infants in the hand condition made accurate predictions to the prior goal object, whereas infants in the claw condition made inaccurate predictions to the prior location. Consistent with Woodward's (1998) original findings and in support of the active experience account, these results suggest that infants can understand other people's intentional actions that are familiar (i.e., the hand grasp) but struggle to understand less familiar actions that they have never performed (i.e., the claw grasp).

In a more recent series of experiments, Krogh-Jespersen and Woodward (2018) varied the level of action priming that 8-month-olds received before completing an eye-tracking paradigm similar to that used by Cannon and Woodward (2012). Infants either received active priming that involved reaching and grasping objects that they would later see in experimental videos, observational priming that involved watching a person sitting next to them reaching for the objects, or no priming at all. Infants then viewed one pre-familiarisation trial in which an actor reached for a single toy. In two familiarisation trials, the actor reached and grasped one of two toys. In two test trials, the toys had swapped locations and the actor raised her hand and paused mid-air between the two toys. The authors found that infants who had only received observational priming, or no priming at all, did not make a higher proportion of accurate predictions (to the prior goal) than inaccurate predictions (to the prior location). In contrast, infants whose own familiar reaching and grasping actions were primed made a higher proportion of accurate than inaccurate predictions. These findings suggest that without active priming, anticipatory predictions of unfolding actions are difficult for young infants to make, even when the action is familiar. Online prediction also seems to be more cognitively challenging than responding to intentional goal structures in visual habituation paradigms in which younger infants have shown sensitivity to intentional goals (e.g., Woodward, 1998, 1999). Rather than training infants to perform a new action as in other literature (e.g., Gerson et al., 2015a, 2015b; Sommerville et al., 2008, 2005), motor priming helps infants to recall their own experience of reaching and grasping when predicting the outcome of the same action performed by another person.

Although the active experience account highlights an important role for self-produced experience in the development of action understanding, it is known that by 6-months of age infants can predict the outcome of some actions that they cannot yet produce, such as using cutlery to eat or bringing a phone to the ear (e.g., Hunnius & Bekkering, 2010). As infants begin to produce actions themselves, comparison is a factor that could build upon infants' active experience in contexts where infants can compare familiar actions with less familiar actions. For example, when eating together at mealtime, an infant could compare his or her own action of picking up food and bringing it to their mouth by hand with an adult's action of picking up the food and eating it with a fork. These actions both achieve the same intentional goal, yet the means by which the goal is achieved is different. Comparison is one way in which infants could come to understand actions, such as using a fork, before they are able to perform them. Whilst some literature suggests that comparison is helpful (e.g., Gerson & Woodward, 2012, 2014a), further empirical evidence is needed to determine whether infants can learn from this process.

To assess infants' understanding of actions that they cannot yet perform, research has used two-step tool-use actions to assess infants' understanding of actions that are more complex than a simple reach with a hand (e.g., Gerson & Woodward, 2012, 2014a; Sommerville et al., 2008). For example, pulling on a cloth or blanket that supports an out-of-reach toy is a more complex action that comes about after mastering simple hand-grasp actions (Gerson et al., 2015b). An even more challenging tool-use action that infants may struggle to understand and cannot execute is a grasp using a claw tool (e.g., Gerson & Woodward, 2012, 2014a; Hofer et al., 2005). To be able to represent this action in terms of its intentional structure, infants must attend to the distal relation between the agent and their intended goal, rather than the proximal relation between the agent and the tool that he or she is directly touching (i.e., the handle of a claw). This ability typically emerges by around 12-months of age but could come about as early as 7-months (Willatts, 1999).

A study by Gerson and Woodward (2012) assessed the role of comparison in infants' action understanding using a training paradigm that enabled 7-month-olds to actively compare their own familiar hand-grasp with another person's unfamiliar claw-reach action. In a physical comparison condition, infants grasped toys that were held by an experimenter using a claw. This condition was compared with three control conditions in which infants could not directly compare their own goal-directed action with the experimenter's, but other behavioural cues that were controlled for to varying extents, such as a movement of the claw to demonstrate its functional affordance. Infants' understanding of the experimenter's intention was assessed in a subsequent goal imitation procedure, in which all infants observed the experimenter grasp one of two toys using the claw and then could choose between the toys. Infants in the physical comparison condition imitated the experimenter's goal choice significantly more often than those in control conditions. This implies that contexts in which infants can compare the goals of their own and another's actions are helpful in facilitating infants' representation of a novel action in terms of its intentional structure. Gerson and Woodward (2012) also investigated whether active physical comparison is more beneficial than comparison between observed exemplars in slightly older infants. 10-month-olds underwent the previously described physical alignment condition, or an observation condition in which they watched the experimenter pass toys to another experimenter using the claw. Going back to my earlier fork example, the alignment condition could mimic a social context in which a parent and infant eat from the same plate using a fork and their hands, respectively. Contrary to this, the observation condition might mimic a context in which an infant watches their parent using the fork to feed another child. Thus, in the alignment condition, infants had the opportunity to compare an observed familiar (hand-reach) and unfamiliar (claw-reach) action without engaging in either action themselves. Infants in the physical comparison condition imitated the experimenter's goal choice more often than those in the observation condition. In support of the active experience account and the literature described above (Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018; Woodward 1998), this study demonstrates that young infants can learn about the intentional structure of a new action by actively comparing it with their own familiar action but are less successful when comparing between observed actions.

To promote the process of observational comparison alone at this stage, additional cues to compare may be required, such as verbal goal labelling. In a study by Gerson and Woodward (2014a), 10-month-olds observed an experimenter use a claw to grasp and pass toys to another experimenter whilst labelling the object. For example, one experimenter passed a toy to the other and said, “A boat, here, a boat.” This gave infants the opportunity to visually compare an unfamiliar tool-grasp with a familiar hand-grasp whilst hearing a matched label intended to draw attention to the shared goal between the actions. In a non-word alignment condition, the experimenters performed the same interaction, except they made a positive vocalization (“Ooh”) that did not draw attention to the shared goal. In a non-alignment condition, the experimenter labelled the toys as she moved them. However, she did not pass the toys to another experimenter, meaning that infants did not have the opportunity to compare the claw-grasp with a hand-grasp action. Goal imitations did not differ between the non-word and non-alignment conditions, whereas infants in the labelling condition made the highest proportion of goal imitations. This suggests that comparing a familiar and unfamiliar action whilst hearing the same label facilitates infants’ understanding of the novel action as goal directed. Object labelling prompted infants to align the two action types and view them as alike, allowing them to infer relational similarity. This provides evidence that, if comparison is facilitated via a consistent cue, infants do not need to have active experience performing an action in order to understand it.

It is untested whether cues other than verbal labelling can facilitate comparison between familiar and unfamiliar actions. For example, multiple observed instances of actions that share the same goal object (i.e., structurally comparable actions) might also facilitate comparison. According to Gentner and Medina (1998), the process of comparison promotes generalisation of similar instances to new, less similar, instances. We propose that relational comparison is an important cognitive factor that can advance infants’ understanding of actions beyond what they have actively produced. The current study investigated whether 12-month-old infants could transfer understanding from comparison to predict the outcome of a novel action. At this age, infants have a clear understanding of hand-reach actions (e.g., Cannon & Woodward, 2012) and are at the cusp of understanding and performing more complex means-end actions (e.g., Sommerville & Woodward, 2005). In an experimental comparison condition, infants observed videos in which an agent’s hand moved onto a screen and grasped a toy car. This was followed by two tool-use actions in which the agent achieved the same goal of grasping the toy car. This involved a cloth-reach action in which the toy car was situated on the end of a cloth which the agent pulled close to grasp the car, as well as a magnet-reach action in which the agent used a magnet to connect and pull the toy close then grasp it. Infants in the comparison condition therefore had the opportunity to compare the structural similarity between a familiar hand-reach followed by less familiar tool-reach actions that shared the same goal. In contrast, infants in a control condition viewed a series of cloth or magnet-reach actions and thus did not have the opportunity to compare this with a familiar hand-reach. In an adaptation of Cannon and Woodward's (2012) paradigm, all infants then underwent familiarisation trials in which the agent used a claw to obtain one of two objects. The objects then swapped locations, and in test trials the hand moved onto the screen and froze, making an incomplete action. During these test trials, we used eye-tracking to measure infants' in-the-moment visual predictions of the incomplete action. This enabled us to assess infants understanding of an unfolding action as opposed to more responsive measures of understanding as in previous imitation or habituation paradigms (e.g., Gerson & Woodward, 2014a; Woodward, 1998, 1999). We decided to use four test trials in an attempt to minimise the risk of missing data due to lack of attention or distraction during only one or two test trials. We hypothesised that infants who were successful at identifying and extracting information about the same overarching goal structure of different, yet comparable, actions would generalise this intentional understanding to accurately predict the outcome of the novel claw-reach action.

Method

Participants

11- to 13-month-olds were recruited in this study. 61 infants completed the study, however data from 21 of these infants were discarded due to failure to make any predictive anticipations across all four test trials (see data coding section below for exclusion criteria). Data from the remaining 40 infants (13 females, *range:* 10.52 – 13.81, M_{age} = 12.08, *SD* = 0.69) were analysed. All infants who participated were full term and reported no developmental delays. Participants were recruited via a participant database and social media. Participants came from a mid-sized city in the United Kingdom and were largely of a White British ethnicity. In a between-subjects design, 20 infants were allocated to a comparison condition (M_{age} = 12.11) and 20 were allocated to a control condition (M_{age} = 12.02). Our dependent variable was infants' proportion of accurate predictions of an action outcome in eye-tracking test trials.

Procedure

Setup and equipment

Data were collected using a Tobii Pro X3-120 eye-tracker, with a 120Hz sampling rate. The Tobii default fixation filter was used to define eye fixation: a stable gaze (within 0.75 visual degrees) for a minimum of 200ms. The eye-tracker was attached to a 23" (58.4cm) monitor (1920 x 1080p). The monitor was mounted onto the wall via a moveable arm that was adjusted for optimal height and distance of the infant sitting on the parent's lap (approximately 60cm). A webcam was attached to the monitor so that the mother and infant could be viewed by the experimenter from another room. Tobii Studio (Tobii Technology, Stockholm, Sweden) eye-tracking software was used for calibration, to record and integrate eye gaze data, and to present stimuli. Prior to the experiment, infants received a five-point calibration to Tobii Studio's pre-set locations. All infants underwent this calibration with the exception of two who underwent nine-point calibration due to experimenter error. We required data to be valid for five calibration points for both eyes before beginning the experiment. Throughout the eye-tracking task, parents were asked to hold the infant securely and orient them toward the screen but not to influence the infant's looking to objects on the screen. At the end of the testing sessions, parents completed a short questionnaire regarding their infant's reaching, grasping and tool-use ability. This was an additional variable that was measured as part of a student project and is not considered in this chapter.

Observational training phase

During the observational training phase, infants in the comparison condition viewed a series of videos in which structurally comparable reaching actions (i.e., actions that shared the same goal structure) were performed (with each video showing a single action, see figure 1a). These videos were approximately four seconds long with some natural variation. All stimuli were filmed using a metronome to aid similar approximation in timing. For example, the time at which the agent moved her hand onto the screen, grasped the tool (cloth or magnet) followed by pulling the object closer were similar across videos. The natural variation between videos was not significantly different between different kinds of videos (i.e., hand-reach, cloth-reach and magnet-reach). Also, infants saw different versions of the same action (e.g., cloth-reach) being performed in each set of three videos. Stimuli did not contain any audio. The first set of three videos showed a hand reaching and pulling a toy car close, followed by three videos of the same toy car located on the end of a cloth being drawn close and a final three showing the car being pulled close using a magnet. Infants in this condition always viewed three hand-reach actions first, followed by cloth-reaches and magnet-reaches counterbalanced between participants. Infants in the control condition only observed one tool-use action (cloth-pull or magnet-pull counterbalanced between participants). They viewed this action three times, with a different colour toy car in each set of three videos (see figure 1b). The chronological order in which infants viewed the different colours of toy car (white, yellow, and red) was counterbalanced between participants. Each set of three videos within both the comparison and control training were interleaved with attention-grabbing animations accompanied with sounds.

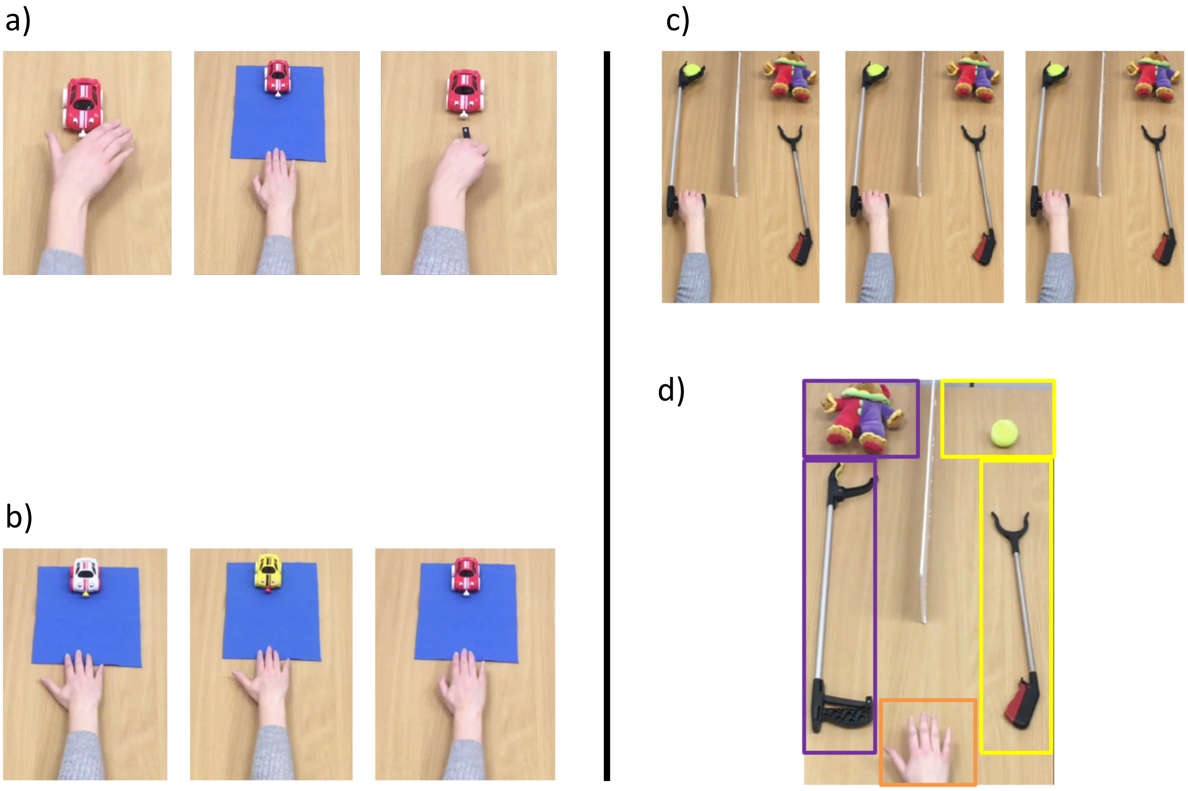

Figure 1. Schematic of eye-tracking paradigm. Still frames depict videos presented. a) Comparison condition: Structurally comparable tool-reach actions. b) Control condition: Tool-reach actions (cloth or magnet) with changing goal object. c) Familiarisation: The same goal object (bear or ball) is consistently reached for and drawn close using corresponding claw. d) Test trials: Incomplete reaching action with defined AOIs. In this example, the orange AOI captured fixations to the hand. Purple AOIs captured fixations to the left claw and left toy (inaccurate goal). Yellow AOIs captured fixations to the right claw and right toy (accurate goal). These AOIs were not visible to infants whilst viewing the videos.

Familiarisation phase

After observational training, infants underwent an adaptation of Cannon and Woodward's (2012) paradigm in which they were familiarised with videos of a claw-reach action involving an agent using a hand-operated claw to reach, grasp, and pull a goal-object closer. A teddy bear and a tennis ball were presented in the top corners of the scene on opposite sides of a wall. Below the bear and ball, two claws were also presented on either side of the wall, which could be used to act upon the corresponding toy. Infants viewed three videos in which the same goal-object (bear or ball) was consistently acted upon with the claw (see figure 1c). The agent's hand entered the bottom centre of the screen and used the left or right claw (counterbalanced) to retrieve the toy (bear or ball) on the corresponding side. The toy that was grasped using the claw was also counterbalanced. These videos were approximately five seconds long with some natural variation.

Eye-tracking freeze screen

In-between familiarisation and test trials, all infants viewed a single still frame in which the bear and ball had swapped positions, but the claws remained in the same position. This image was presented for five seconds to allow time for infants to encode the swap by fixating on at least one of the two toys (see data processing section below).

Eye-tracking test trials

With the bear and ball in their new positions, infants then viewed four test trials in which the agent's hand moved onto the centre of the screen and froze, making an incomplete action (see figure 1d). The time at which the hand moved onto the screen varied slightly between videos, appearing after an average of 1.15 seconds. These videos were approximately five seconds long with some natural variation.

Data Coding

Eye-tracking data

Eye-tracking data from both the freeze screen and test trials were processed using R (R Core Team, 2013). During the freeze screen, eye fixations to both the bear and ball were measured to infer whether infants encoded the swap in toy positions. AOIs (matched in size, 345 x 240p) were defined over the bear and ball to ensure that infants fixated on at least one of these objects to encode the swap in locations. If infants did not fixate on either one of these objects, we concluded that infants did not encode the swap during this phase. If an infant did not encode the swap during the freeze screen, we used the same method to investigate whether the swap was encoded during initial test trials. This is because infants may not have noticed the swap in toy positions during the freeze screen but subsequently had the opportunity to notice the swap during test trials. No trials were coded for predictions before the swap had been encoded, as fixations before this point were not likely to be anticipatory of the agent's intention. The majority of infants encoded the swap during the freeze screen, whilst two infants encoded the swap during an initial test trial.

In test trials, AOIs for the bear and ball (matched in size, 311 x 228p), as well as the left and right claw (matched in size, 230 x 739p) were defined along with an AOI covering the hand (see figure 1d for representation of AOIs). Fixations were coded from the exact moment the hand entered the screen on each test trial. This was precisely defined for each trial. Timestamps of fixation to each AOI determined the order of fixations and thus allowed us to code predictive looks. After fixating on the hand, we coded the first object that infants fixated on (hand-to-object prediction). As in previous visual prediction paradigms (Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018), fixations to a toy or claw without first fixating on the hand were not considered predictive anticipations of the familiarised action. This is because looks to objects without first looking to the hand first were likely to be arbitrary rather than anticipatory. When infants made hand-to-object fixations on the familiarised goal object or its newly corresponding claw, this was considered an *accurate prediction.* When infant made hand-to-object fixations on the familiarised claw, or the non-familiarised object (in the familiarised location) this was considered an *inaccurate prediction.* Fixations to claws were included as predictive looks because infants were familiarised with the action trajectory of the hand using the claw to grasp the toy. If infants do not understand the agent's two-step action of using the tool to reach the goal-object, we expect them to look from the hand to the claw that the agent manipulated during familiarisation (i.e., an inaccurate prediction in which infants are only predicting a one-step action). In contrast, infants who understand the agent's two-step action of using the claw to achieve a preferred toy should look to the claw that can reach for the familiarised toy (i.e., an accurate prediction). If infants did not look at the hand or did not fixate on another AOI following fixation on the hand, these trials were considered as *no anticipation* and were excluded from analyses. If infants did not make predictive anticipations on any test trial their data were excluded from analyses (*N* = 21). This attrition rate is comparable to other research that has used similar remote (as opposed to head-mounted) eye-tracking technology (Corbetta et al., 2012). Saccades through AOIs were not coded as predictive looks. Each infant's proportion of predictive fixations was calculated separately for accurate and inaccurate predictions, with number of predictive trials as the denominator.

Manually coded data

As an initial method for analysing our data, we also manually coded predictive looks using a gaze replay overlay presented using Tobii Studio Replay. This was the same technique as that used by Cannon and Woodward (2012), in which infants' accurate and inaccurate predictions were coded using a similar method of gaze trace overlay. Using the same prediction criteria as detailed above (i.e., hand-to-object predictions), all test trials were coded for accurate and inaccurate visual predictions or were coded as no anticipation if no predictive fixations were made. Each video recording was viewed at 25 frames per second with a gaze trace visible for manual coding. AOIs were not visible in gaze trace coding. For this reason, a second trained observer coded 23% of all recruited participants' videos (i.e., 61 participants before exclusion due to no anticipations) and proportion of accurate predictions were compared for inter-rater agreement. There was moderate agreement between coders, κ = .55, *p* < .001. Due to an error in video presentation, videos of a smaller size were presented to two infants in the comparison condition, meaning that the specified AOIs did not accurately represent objects in the visual scene. Therefore, data from manual coding were used in place of the AOI extracted data for these two infants in order to maintain a larger sample size. Due to the possibility that these infants may interpret the videos differently due to their smaller size, we checked that the data were descriptively similar both with and without data from these two participants included.

Analyses and Results

Analyses

Before running our analyses, we checked for any inconsistencies between prediction data that had been processed from eye-tracking and our manually coded data. Out of a total of 152 trials (including those in which no anticipations had been made) across the 38 infants for whom we had both eye-tracking and manually coded data, we identified 2 trials in which there was a mismatch between our two coding formats. In both of these trials, predictions had been defined as accurate during manual coding, whereas these had been defined as inaccurate in the data processed from eye-tracking. This was likely due to ambiguity over fixations that were between AOIs. For example, looks between the bear and ball, or what looked like a fixation to the hand which was actually within a claw AOI. All subsequent analyses included prediction data that had been extracted from eye-tracking with the addition of manually coded prediction data from two infants (described above).

Preliminary analyses were run to check for any potential effect of age, sex, or familiarised goal object (bear or ball) upon proportion of accurate predictions. As proportion of accurate predictions were calculated across a maximum of four test trials, this variable ranged ordinally (i.e., 0%, 25%, 33%, 50%, 67%, 100%). We therefore used generalised linear models (GLZMs) as non-parametric alternatives to ANOVAs. In a preliminary analysis, we ensured that there was no difference in proportion of accurate predictions between infants in the control condition who had been exposed to either the cloth or the magnet as the tool used during training trials. In our main analysis, we examined our hypothesis that infants in the comparison condition would make a higher proportion of accurate predictions than infants in the control condition. We then assessed whether predictions made in each condition were significantly different from chance level (i.e., 50%). In addition to our main analyses, we investigated whether there was any difference between conditions for the first prediction (accurate or inaccurate) that each infant made across all four test trials. For example, if an infant did not make a prediction in the first test trial but did in the second, we used the prediction data from the second trial. This was in line with Cannon and Woodward's (2012) paradigm that used a single test trial to assess accurate visual prediction of an incomplete action. An ANOVA was used for analysis concerning latency to make predictions. Table 1 shows the proportion of infants in each condition who made accurate predictions, inaccurate predictions, and no anticipations within each test trial. Within the comparison condition, proportions of accurate predictions declined throughout test trials, particularly in the final trial, whilst no anticipations increased throughout test trials likely due to a decline in infants' attention. Such trends were not seen in the control condition, with more variation in predictive performance throughout test trials. Table 2 represents the various categories of performance patterns shown by individual infants across test trials. That is, the extent to which infants' prediction accuracy improved, declined, or varied from the first to last test trial in which accurate or inaccurate predictions were made. Infants are also categorised into those who made accurate predictions only or inaccurate predictions only. Average numbers of predictive trials indicate the extent to which infants made predictions (i.e., accurate or inaccurate) as opposed to no anticipations. Within the comparison condition most infants made accurate predictions only (45%), whereas within the control condition most infants made inaccurate predictions only (35%).

Table 1

Proportion of infants in each condition who made accurate predictions, inaccurate predictions, and no anticipations within each test trial.

	Trial 1	**Trial 2**	**Trial 3**	**Trial 4**
Comparison condition (*N* = 20)				
Accurate	40%	35%	35%	13%
Inaccurate	15%	10%	13%	13%
No anticipation	45%	55%	53%	75%
Control condition (*N* = 20)				
Accurate	30%	35%	30%	20%
Inaccurate	40%	30%	30%	45%
No anticipation	30%	35%	40%	35%

Table 2

Patterns of prediction accuracy from first to last predictive test trial, as well as the average number of test trials in which infants made a prediction (i.e., accurate or inaccurate) as opposed to no anticipation.

Accuracy across predictive test trials	**%**	**Average number of predictive test trials**
Comparison condition (*N* = 20)		
Accuracy improved (inaccurate to accurate)	15%	2.67
Accuracy declined (accurate to inaccurate)	25%	2.40
Accuracy varied	5%	4.00
Accurate predictions only	45%	1.80
Inaccurate predictions only	10%	2.00
Control condition (*N* = 20)		
Accuracy improved (inaccurate to accurate)	10%	3.50
Accuracy declined (accurate to inaccurate)	20%	2.30
Accuracy varied	10%	3.50
Accurate predictions only	25%	2.60
Inaccurate predictions only	35%	2.30

Results

In a preliminary GLZM, proportion of accurate predictions was entered as the dependent variable. Sex, familiarised goal object (bear or ball), and goal location (left or right) were entered as predictor variables. Age was entered as a continuous variable. This revealed a significant effect of sex, Wald χ^2 (1) = 6.05, *p* = .01, β = .37, such that females made a higher proportion of accurate predictions (65% accurate) than males (50% accurate). This may have been due to the unequal number of females (*N* = 13) and males (*N* = 27). There was also a marginally significant effect of goal object, Wald χ^2 (1) = 3.45, *p* = .06, β = .33, such that infants who were familiarised with the ball made a higher proportion of accurate predictions (59% accurate) than infants who were familiarised with the bear (52% accurate). As well as this, there was a significant interaction between sex and goal object, Wald χ^2 (1) = 3.97, *p* = .05, β = -.47. Males who were familiarised with the ball made a higher proportion of accurate predictions (64% accurate) than males who were familiarised with the bear (39% accurate). However, a follow-up Mann-Whitney test revealed that this difference was not significant, *U* = 59.50, *p* = .12. Females showed the reverse pattern in that those who were familiarised with the bear made a higher proportion of accurate predictions (75% accurate) than females who were familiarised with the ball (48% accurate). However, a follow-up Mann-Whitney test revealed that this difference was not significant, *U* = 11.50, *p* = .20. The GLZM revealed no other significant effects or interactions between variables (*ps* > .17). Due to these effects, we controlled for sex and goal object in all subsequent analyses that included proportion of accurate predictions as the dependent variable.

A cloth-reach action may be more familiar to 12-month-olds than a magnet-reach action, as infants are less likely to interact with magnets in their everyday lives whereas a cloth is like a blanket. So, we considered whether infants in the control condition who observed the cloth-reach during training may have better generalised their understanding to make accurate predictions than infants who saw the magnet-reach. Proportion of accurate predictions was entered as the dependent variable in a GLZM. Tool (cloth vs magnet), sex (control variable), and familiarised goal object (control variable) were entered as predictors. This revealed no significant main effect of the tool used during training, Wald χ^2 (1) = .10, *p* = .76, β = .06.

Hand-to-object predictions

Our next analysis examined our main hypothesis that infants in the comparison condition would make a higher proportion of accurate predictions (to the goal object's new location) than inaccurate predictions (to the goal object's old location) during test trials. Proportion of accurate predictions was entered as the dependent variable in a GLZM. Condition (comparison vs control), sex (control variable), and familiarised goal object (control variable) were added as predictors. This revealed a marginally significant effect of condition, Wald χ^2 (1) = 3.61, *p* = .06, β = -.22. Infants in the comparison condition made a higher proportion of accurate predictions (on 67% of trials) than infants in the control condition (on 43% of trials; see figure 2). If infants were performing at chance level, then predictions were equally likely to be accurate or inaccurate. Planned comparisons against chance indicated that infants in the comparison condition systematically made accurate predictions, Wald χ^2 (1) = 4.80, *p* = .03, β = -.17. However, predictions made by infants in the control condition did not significantly differ from chance, Wald χ^2 (1) = .58, *p* = .45, β = .07. There were no main effects of, or interactions between, sex or goal object in these analyses (*ps* > .40).

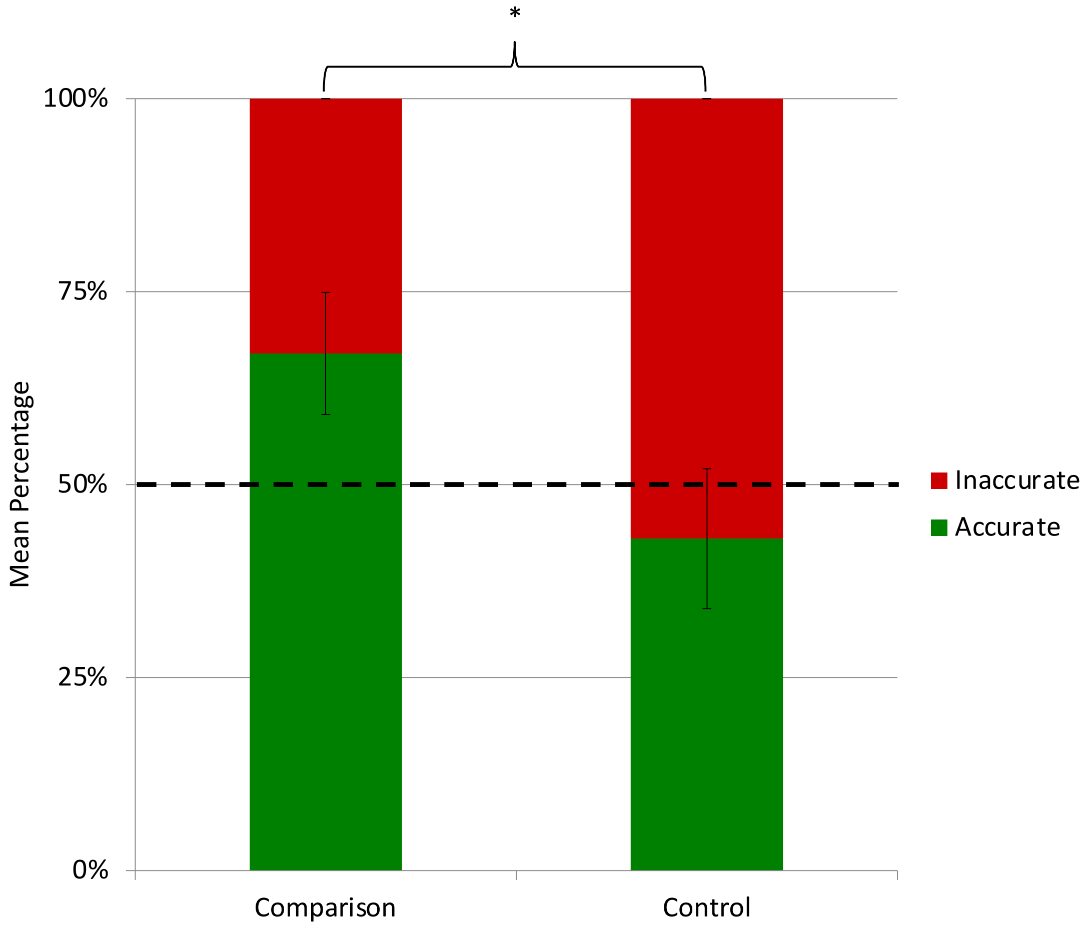

Figure 2. Proportion of accurate and inaccurate hand-to-object predictions in comparison and control conditions. Error bars represent standard error. Chance = 50%. **p* = .06.

We also investigated whether there was any difference between conditions for the first prediction (accurate or inaccurate) that each infant made across all four test trials. For example, if an infant did not make an anticipation in the first test trial but did make a prediction in the second trial, we used the prediction data from the second trial. First prediction was entered as the binary dependent variable in a GLZM. Condition, sex (control variable) and familiarised goal object (control variable) were added as predictors. This revealed no significant effect of condition, Wald χ^2 (1) = 1.41, *p* = .24, β = -.17, or goal object, Wald χ^2 (1) = 1.19, *p* = .28, β = .16. However, there was a significant main effect of sex, Wald χ^2 (1) = 6.28, *p* = .01, β = -.38, such that males made a higher number of accurate first predictions (12) than females (10). Like the effect of sex that we saw in preliminary analysis, this is likely because of the unequal number of males (*N* = 27) and females (*N* = 13).

Prediction latency

Next, we investigated the hypothesis that infants would make accurate predictions more slowly than inaccurate predictions due to an increase in cognitive effort (Krogh-Jespersen &

Woodward, 2014, 2018). We were interested in whether prediction speed differed both across and between conditions. Latency to predict was entered as the dependent variable in an ANOVA. Prediction type (accurate vs inaccurate) and condition were entered as predictor variables. There were no significant main effects or interactions between variables (*ps* > .27).

Discussion

Previous literature has shown that infants can learn to understand new intentional actions by comparing them with familiar actions (e.g., Gerson & Woodward, 2012, 2014a). The current study investigated whether 12-month-olds are able to generalise intentional understanding from comparison to successfully anticipate the outcome of a never-before-seen action. Our findings confirm that infants can learn about the goals of new actions through this process. When provided with a cue to compare (a familiar hand-reach) before seeing similar tool-use actions that share the same goal structure, infants were able to gain insight into the intentional structure of these actions. Infants then applied this intentional understanding to accurately predict the outcome of the novel claw-reach action with a different goal. Although marginally significant, infants in the comparison condition also made a higher proportion of accurate predictions than infants in a control condition who did not have the opportunity to compare a familiar hand-reach with tool-reach actions. Infants in the control condition performed at chance level with regards to generating predictions to either the familiarised toy (i.e., accurate predictions) or to the familiarised location (i.e., inaccurate predictions). Despite previous findings that accurate predictions are made more slowly than inaccurate predictions due to heavier cognitive processing (Krogh-Jespersen & Woodward, 2014, 2018), we found no difference in latencies between prediction types in the current study. Unexpectedly, we found significant main effects of sex and goal object across conditions, as well as an interaction between sex and goal object. Females made a higher proportion of accurate predictions (65% accurate) than males (50% accurate), and infants who were familiarised with the ball made a higher proportion of accurate predictions (59% accurate) than infants who were familiarised with the bear (52% accurate). Also, males who were familiarised with the ball made a higher proportion of accurate predictions (64% accurate) than males who were familiarised with the bear (39% accurate). Females showed the reverse, such that those who were familiarised with the bear made a higher proportion of accurate predictions (75% accurate) than females who were familiarised with the ball (48% accurate). It is difficult to determine whether these findings are meaningful or spurious.

To some extent, our main findings challenge the active experience account of action understanding (e.g., Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018; Woodward 1998), as infants in the current study were able to accurately predict the outcome of a novel claw-reach action that they had not produced previously. However, our results still highlight the importance of infants' previous motor experience in that they rely on knowledge of actions that they can perform when making comparisons between familiar and less familiar actions. Our findings also provide further reinforcement for Gerson and Woodward's (2012) research, which found that 10-month-olds are able to learn an agent's goal preference through a process of observational comparison in which verbal labelling of the goal object served as a cue for infants to compare structural similarity between a claw-grasp and a hand-grasp. Our results demonstrate that 12-month-olds, who are at the cusp of understanding complex actions, can use multiple observed instances of the same goal (i.e., the cue to compare) reached using different tools to extract and generalise intentional information to a novel action. However, we do not yet know if infants younger than a year old (e.g., 10-months of age) are able to generalise understanding from comparison in the same way. In the current study, we expected that infants in the control condition, who did not compare between actions, would systematically make visual predictions to the familiarised location in which the goal object had been reached for. However, as the infants performed at chance level, it is possible that some infants within this group were able to make accurate predictions without the need for observational training. That is, more motorically advanced infants might be able to draw upon their existing motor experience to make sense of the novel claw action. A younger, less motorically advanced, sample of infants may make a higher proportion of inaccurate predictions in the control condition. In the next chapter we used a similar paradigm to investigate 10-month-olds' ability to understand a new action after undergoing comparison training. We also added a behavioural measure of motor ability to investigate the influence that individual differences in motor skill have on infants' capacity to make accurate goal predictions.

To conclude, the active experience account suggests that infants begin to understand other people's actions as they learn to produce them themselves (e.g., Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018; Woodward, 1998). This may stand true during early development, as infants struggle to recognise another person's intention after receiving observational training alone (e.g., Gerson & Woodward, 2014b; Krogh-Jespersen & Woodward, 2018; Sommerville et al., 2005). However, our findings suggest that as infants become more proficient at producing goal-directed actions, they are able to call to mind their own experience when seeing a familiar action and compare this with examples that they are less proficient at performing. Therefore, it seems that active experience provides a basis for learning about other actions through observational comparison. In the development of intentional understanding, infants move from relying on active experience alone to a combination of social and cognitive factors, such as motor experience and cues for comparison.

Chapter 3

The effects of observational comparison and active motor training on infants' action understanding

Abstract

At a young age, active experience facilitates action understanding over and above matched observational experience (e.g., Gerson et al., 2015b). However, when cues for comparison are present, infants can also learn about new actions through observation (e.g., Gerson & Woodward, 2014a). By comparing actions that they cannot perform with actions that they can self-produce, infants could develop an intentional understanding of completely novel actions (Gerson & Woodward, 2012). In contrast with the 12-month-old infant sample in study 1, study 2 examined whether observational experience with hand- and tool-reach comparisons supports 10-month-olds' predictions of a novel action that they have never produced. In our observational comparison intervention (*N* = 22), infants observed a hand-reach followed by a comparable cloth- and magnet-reach (counterbalanced) upon the same goal, whereas infants in a control condition (*N* =17) viewed the cloth- and magnet-reaches (counterbalanced) before the hand-reach. Test trials required all infants to generalise their intentional knowledge from training to visually predict the outcome of a novel claw-reach action. After completing the eye-tracking task, infants completed a motor task to assess their ability to actively perform the cloth- and magnet-reach action in a planful manner. We assessed whether individual differences in infants' existing motor ability, as well as their propensity to learn to perform tool-use actions, related to differences in accurate predictions. Our main findings revealed that within the control condition, infants who showed low motor ability made a higher proportion of accurate predictions (on 70% of trials) than infants showing high motor ability (on 24% of trials). In the comparison condition, infants who showed low motor ability made a slightly higher proportion of accurate predictions (on 36% of trials) than infants showing high motor ability (on 30% of trials). However, this small difference did not reach significance. We did not find any significant effects of motor ability alone or condition alone on infants' proportion of accurate visual predictions. Across both comparison and control conditions, infants who observed the cloth-reach action first during our observational training intervention made a higher proportion of accurate predictions (49%) than infants who observed the magnet-reach action first (31%). We speculate that infants in the control condition were able to compare between tool-reach actions and the familiar hand-reach even when viewing the hand-reach last. Our results suggest that, for infants with less efficient motor skills, comparing between

actions in this order may be more beneficial than viewing a familiar action first. We also speculate that the cloth-reach action is more familiar to 10-month-old infants than a magnet-reach, and thus the cloth-reach is an easier basis for comparison with the more complex magnet-reach.

Introduction

In the previous chapter, we found that 12-month-olds were able to generalise intentional understanding that they learned from observational comparison to accurately predict the outcome of a never-before-seen action. In a comparison condition, infants viewed a familiar hand-reach action, involving an agent reaching for a toy car. They then saw two comparable tool-reach actions, where the agent reached for the same toy by pulling on a cloth (that the toy was situated on the end of), as well as a magnet-reach. Infants in the control condition did not have the same opportunity to compare the tool-reach with a familiar hand-reach and instead saw the agent reach for different toy cars (a changing goal-object) using either a cloth or magnet. When viewing a novel claw-reach action after observational training, infants in the comparison condition made a higher proportion of accurate visual predictions than infants in the control condition. However, this difference between conditions did not reach significance. We speculated that if younger infants (e.g., 10-month-olds) underwent the same observational training, we may see a larger difference between conditions. We reason this because some older infants (12-month-olds) who are more motorically advanced may make sense of a novel action and accurately anticipate the outcome without the need to learn from intentional cues such as comparison (Sommerville & Woodward, 2005). Biro and Leslie (2007) provide support for this idea, as infants as young as 6-months of age could make sense of an unfamiliar tool-use action if a number of intentional cues were present, whereas older 9- and 12-month-olds with more efficient motor skills required fewer cues in order to represent the action in terms of its intentional structure. In the current study, we investigated 10-month-olds' ability to generalise learning from observational comparison training to a novel claw-reach action. We also took individual differences in infants' motor ability into account by measuring infants' skill at performing the cloth- and magnet-reach actions seen during observational training.

As infants become more proficient at performing actions themselves, they also begin to recognise the intentions of other people's actions (e.g., Gerson et al., 2015; Gerson & Woodward, 2014b, 2014c; Sommerville et al., 2005, 2008). Experience performing actions seems to be more helpful for intentional understanding than simply observing actions, particularly during early development. Findings to support this come from Gerson and Woodward's (2014b) research in which 3-month-old infants were given either active or observational experience using Velcro mittens to manipulate a ball and a bear. This enabled infants in the active condition to reach and grasp objects, which is an action that they cannot yet efficiently perform. In a visual habituation paradigm, infants who had undergone active training showed strong evidence of goal recognition (i.e., longer looking when a different goal object was grasped), whereas goal recognition did not differ between infants who had received observational training or no training at all. Other intervention studies have trained older infants to perform more complex two-step actions, such as pulling on a cloth to retrieve a toy situated on the end of it, before measuring infants' understanding of the same action when watching someone else perform it. For example, Gerson et al. (2015b) trained 8-month-olds to perform a cloth-reach action and found that infants' success during training related to goal recognition in a habituation paradigm. As well as this, infants who were less successful during training attended to the proximal relation between the agent and the cloth that she was directly touching (i.e., the first step of the action), implying a lack of understanding of the agent's distal goal of retrieving the toy (i.e., the complete two-step action). Much like the findings of Gerson and Woodward (2014b, 2014c), experience observing the cloth-pull action did not facilitate 8-month-olds' goal recognition. From these intervention studies, it is clear that motor experience is useful for enabling infants to understand other people's intentional actions. However, other evidence suggests that cognitive factors, such as comparison, also help to facilitate infants' understanding of actions that they cannot yet perform (e.g., Gerson & Woodward, 2012; 2014a).

As described in chapter 1, structure mapping theory (Gentner & Medina, 1998) proposes that learners can generalise relational information when they compare two instances that share a similar structure, such as a matching goal. Infants could apply such a process of comparison when coming across a less familiar goal-directed action that is comparable to a self-produced action. Knowledge and experience gained from extracting the relation between a person and their goal via comparison could also help infants to understand completely novel intentional actions that involve different goals. In the context of learning through structural comparison, Gerson and Woodward (2014a) have investigated the role of linguistic goal labelling as a cue to highlight the intentional structure of an action. 10-month-olds watched an experimenter use an unfamiliar claw-tool to pass toys (e.g., boat, train, ring etc.) to another experimenter whilst verbally naming the object. The second experimenter then took the toy from the claw using a familiar hand-grasp whilst verbally labelling the object to draw infants' attention to the goal. When hearing the matched label, infants were more successful at identifying the shared goal of the claw- and hand-reach, and thus understanding the intention of the claw action. Infants in this comparison condition made a higher proportion of goal imitations than infants in control conditions who did not hear the matched label or did not have the opportunity to compare the two actions. These findings suggest that, when provided with this cue to compare, 10-month-olds can learn about intentional actions through observation alone. From our findings in the previous chapter, as well as those of Gerson and Woodward (2012, 2014a) it seems evident that with age, infants can learn from observational comparison alone. When cues for comparison are present, observational training may be helpful for 10-month-olds. Because infants' own proficiency at producing actions facilitates understanding of other people's actions (e.g., Gerson et al., 2015b; Gerson & Woodward, 2014b, c; Sommerville et al., 2005), individual differences in infants' motor skills could also help explain any variability in accurate visual predictions.

In the current study we investigated the role of observational comparison, motor ability, and propensity to learn to produce an action upon infants' accurate visual predictions of a novel action. Tool use naturally emerges by approximately 12-months of age (e.g., Sommerville & Woodward, 2005). Motor ability also becomes more efficient by the end of the first year (e.g., Santos et al., 2001), meaning that infants may make accurate predictions without relying on cues to compare. To ensure that infants were at the cusp of learning to perform and understand tool-use actions but were not yet particularly efficient as a group, we recruited 9- to 11-month-olds for the current study. As motor efficiency varies at this age, we hypothesised that infants who were less motorically advanced may rely on our comparison intervention for facilitating intentional understanding, rather than previous motor experience. Infants underwent a similar observational training and eye-tracking paradigm to that described in the previous chapter. However, some changes were made to improve the validity of our observational training. In study one (see chapter 2), infants in the control condition only observed one of two tool-use actions during observational training (cloth-reach or magnet-reach) with a changing goal object (a different colour toy car). In contrast, infants in the comparison condition viewed a hand-reach followed by both of the tool-reach examples with the same toy car as the goal object. This meant that the initial hand-reach served as a basis for comparison with the tool-reaches. In addition to varying in terms of the possible number of comparisons to be made between conditions, these conditions also varied perceptually between trials. The control condition consisted of videos in which the same action was carried out repeatedly. The only variability was the colour of the toy car that was acted upon. This may have unintentionally primed infants in the control condition to compare the changing goal-object with the constant tool, thus highlighting the tool as being more important than the goal (Gerson & Woodward, 2013). To reduce the level of variability between conditions and better test the effects of our comparison intervention in the current study, infants in the control condition observed actions that were identical to those seen in the comparison condition. However, infants in the control condition viewed the videos in a different order, with the hand-reach shown last instead of first. This way, a matching goal object was highlighted in each of the actions, but infants in the control condition did not have the hand as an initial basis for comparison with the subsequent tool-reaches. Observational training in the comparison condition remained the same as that described in the previous chapter. That is, infants were primed by observing a familiar hand-reach first, followed by less familiar tool-reaches, to facilitate comparison between structurally similar actions. After training, infants in both conditions were familiarised with a novel action in which an agent used a claw to reach for one of two toys. The toys then swapped positions and in test trials we measured infants' visual predictions of the incomplete claw-reach action. Unlike previous literature that measured the effects of a motor intervention for facilitating infants' understanding of intentional actions (e.g., Gerson et al., 2015b; Gerson & Woodward, 2014b, 2014c; Sommerville et al., 2005, 2008), infants in the current study completed a motor task *after* the eye-tracking task. The purpose of this motor task was to measure individual differences in infants' motor ability without motorically priming infants before they underwent our observational training intervention. This enabled us to measure the effects of our observational training intervention rather than the effects of active motor training. To account for the potential confound of learning from 'seeing' followed by 'doing', we also checked that there were no differences in performance on the motor task between infants who had undergone our observational comparison versus control training interventions. Using a motor task inspired from previous work (Gerson et al., 2015b; Sommerville et al., 2008; Sommerville & Woodward, 2005), we investigated individual differences in infants' existing ability to perform the same cloth-reach and a magnet-reach action that they saw during observational training. We also assessed infants' propensity to learn to perform these new actions after undergoing active motor training (i.e., learning from practice rather than learning from observing videos). We expected that infants in the control condition who were motorically efficient in the motor task would make accurate predictions during the eye-tracking task without the need to make comparisons. We also expected that drawing upon motor experience would not be necessary to learn from comparison, thus infants who were not efficient in the motor task would still make accurate predictions in the comparison condition.

Method

The procedure and initial analysis pipeline for this study was preregistered on Open Science Framework (Findlay & Gerson, 2019, https://osf.io/cp9zj/)

Participants

9- to 11-month-olds were recruited in this study. 56 infants completed the study, however data from 17 of these infants were discarded. Within the eye-tracking task, this was because of fussiness (N = 1), failure to attend to training or familiarisation trials for an adequate proportion of time (N = 7), failure to encode the swap in toy locations (N = 1), or failure to make any type of prediction during test trials (*N* = 7). One infant also failed to complete the motor task. See the data processing section below for detailed descriptions of exclusion criteria within each phase of the study. Data from the remaining 39 infants (22 females, *age range* = 9.1 – 11.8 months, M_{age} = 10.08 months, *SD* = 0.7) were analysed. In a between-subjects design, infants were allocated to a comparison (*N* = 22; M_{age} = 10.18) or control condition (*N* = 17; M_{age} = 9.94) for the eye-tracking task. All infants then completed the same motor task. Infants who were excluded from analyses were spread relatively equally across conditions (*N* = 9 excluded from comparison and *N* = 7 excluded from control). All infants who participated were full term and reported no developmental delays. Participants were recruited via a participant database as well as social media and came from a mid-sized city in the United Kingdom. Information regarding ethnicity was collected for 37 of the infants who took part in the study. Of these, the majority were White British (*N* = 28), with the remainder being of another White ethnicity (*N* = 4), mixed Chinese and British (*N* = 3), mixed White and Asian (*N* = 1), and mixed Kurdish Turkish and British (*N* = 1).

Procedure

Setup and equipment

Data were collected using a Tobii Pro X3-120 eye-tracker, with a 120Hz sampling rate. The Tobii default fixation filter was used to define eye fixation: a stable gaze (within 0.75 visual degrees) for a minimum of 200ms. The eye-tracker was attached to a 23” (58.4cm) monitor (1920 x 1080p). The monitor was mounted onto the wall via a moveable arm that was adjusted for optimal height and distance of the infant sitting on the parent’s lap (approximately 60cm). A webcam was attached to the monitor so that the mother and infant could be viewed by the experimenter from another room. Tobii Studio (Tobii Technology, Stockholm, Sweden) eye-tracking software was used for calibration and to record eye-gaze data whilst E-prime (Psychology Software Tools, Pittsburgh, PA) was used to present stimuli. Prior to the experiment, all infants underwent a five-point calibration to Tobii Studio’s pre-set locations. We required data to be valid for five calibration points for both eyes before beginning the experiment. Throughout the eye-tracking task, parents were asked to hold the infant securely and orient them toward the screen but not to influence the infant’s looking to objects on the screen. Infants underwent the same procedure as that reported in study 1, except trials were progressed manually using a key press. Using another key press, the experimenter could also play an animation to gain the infant’s attention. There was also a change to the control training condition, as described in the next section.

Eye-tracking training phase

During the observational training phase, infants in the comparison condition viewed a series of videos in which structurally comparable reaching actions (i.e., actions that shared the same goal structure) were performed (with each video showing a single action, see figure 3a). These videos were approximately four seconds long with some natural variation. All stimuli were filmed using a metronome to aid similar proximation in timing. For example, the time at which the agent moved her hand onto the screen, grasped the tool (cloth or magnet) followed by pulling the object closer were similar across videos. The natural variation between videos was not significantly different between different kinds of videos (i.e., hand-reach, cloth-reach and magnet-reach). Also, infants saw different versions of the same action (e.g., cloth-reach) being performed in each set of three videos. Stimuli did not contain any audio. The first set of three videos showed a hand reaching and pulling a toy car close, followed by three videos of the same toy car located on the end of a cloth being drawn near and a final three showing the car being pulled close using a magnet. Infants in this condition always viewed three hand-reach actions first, followed by cloth-reaches and magnet-reaches counterbalanced between participants. Infants in the control condition viewed the tool-reaches first (magnet or cloth first counterbalanced between participants), followed by the hand-reach (see figure 3b). Each set of three videos within both the comparison and control training were interleaved with attention-grabbing animations accompanied with sounds. A fixation cross was displayed preceding each set of three videos (i.e., three hand- cloth- and magnet-reach examples were shown consecutively). The length of time that the fixation cross was displayed was controlled manually by the experimenter. This enabled the experimenter to wait until the infant was attending to the screen before progressing to the next trial.

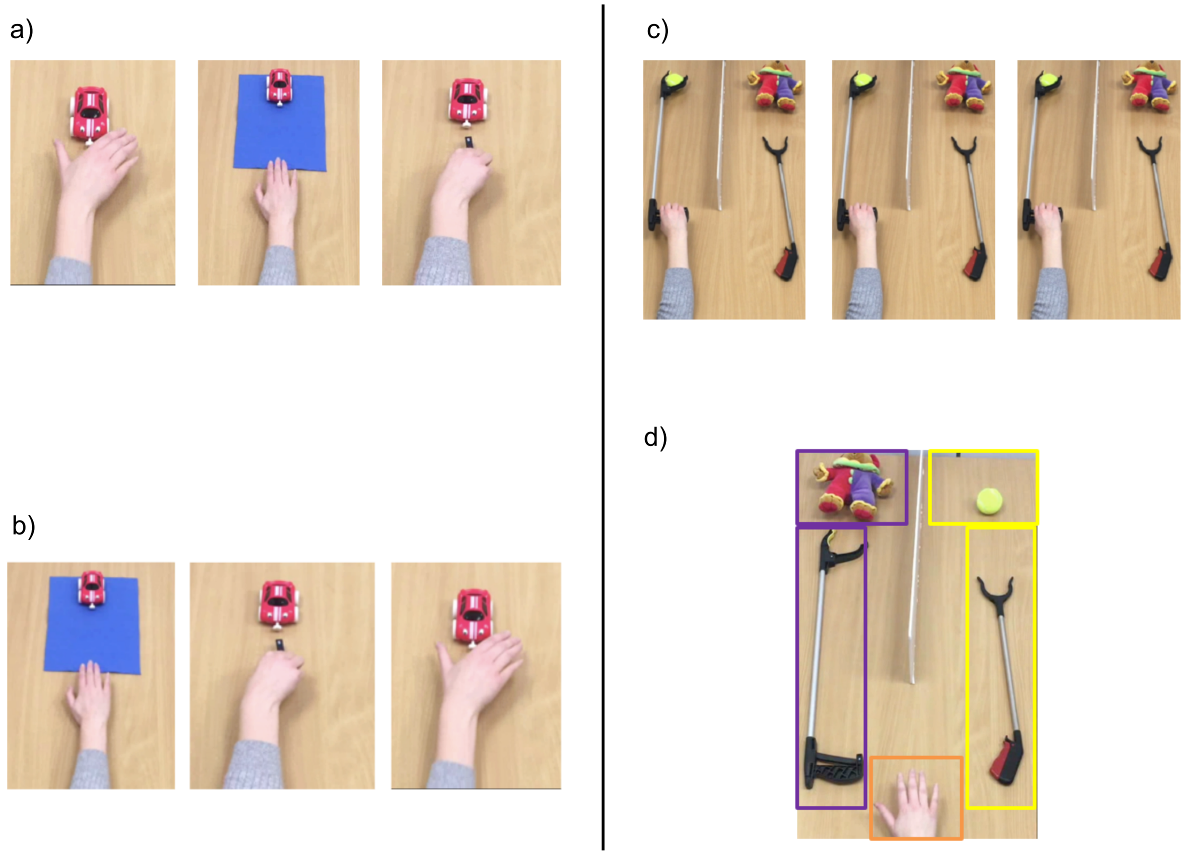

Figure 3. Schematic of eye-tracking paradigm. Still frames depict videos presented. a) Comparison condition: Hand-reach action shown first. b) Control condition: Hand-reach action shown last. c) Familiarisation: The same goal object (bear or ball) is consistently reached for and drawn close using corresponding claw. d) Test trials: Incomplete reaching action with defined AOIs. In this example, the orange AOI captured fixations to the hand. Purple AOIs captured fixations to the left claw and left toy (inaccurate goal). Yellow AOIs captured fixations to the right claw and right toy (accurate goal). These AOIs were not visible to infants whilst viewing the videos.

Eye-tracking familiarisation phase

After observational training, infants underwent an adaptation of Cannon and Woodward's (2012) paradigm in which they were familiarised with videos of a claw-reach action involving an agent using a hand-operated claw to reach, grasp, and pull a goal-object closer. A teddy bear and a tennis ball were presented in the top corners of the scene on opposite sides of a wall. Below the bear and ball, two claws were also presented on either side of the wall, which could be used to act upon the corresponding toy. Infants viewed three videos in which the same goal-object (bear or ball) was consistently acted upon with the claw (see figure 3c). The agent's hand entered the bottom centre of the screen and used the left or right claw (counterbalanced) to retrieve the toy (bear or ball) on the corresponding side. The toy that was grasped using the claw was also counterbalanced. These videos were approximately five seconds long with some natural variation. Within this phase, as well as the later training phase (see below), a fixation cross was displayed preceding each trial and trials were progressed manually by the experimenter. Upon an alternate key press, the experimenter could also choose to present an attention-grabbing animation to re-gain the infant's attention before proceeding with the trial.

Eye-tracking freeze screen

In-between familiarisation and test trials, all infants viewed a single freeze screen in which the bear and ball had swapped positions, but the claws remained in the same position. This image was presented for five seconds to allow time for infants to encode the swap by fixating on at least one of the two toys (see data processing section below).

Eye-tracking test trials

With the bear and ball in their new positions, infants then viewed four test trials in which the agent's hand moved onto the centre of the screen and froze, making an incomplete action (see figure 3d). The time at which the hand moved onto the screen varied slightly between videos, appearing after an average of 1.15 seconds. These videos were approximately five seconds long with some natural variation. After completing this task, infants underwent a second eye-tracking task as part of a separate project that is beyond the scope of this thesis.

Motor trials: Pre- and post-training

After completing the eye-tracking tasks, infants completed a motor task inspired from previous work (Gerson et al., 2015b; Sommerville et al., 2008; Sommerville & Woodward, 2005) that consisted of pre-training, training, and post-training trials in which infants had the chance to perform a cloth-reach and a magnet-reach (in a counterbalanced order) to retrieve a magnetic toy train. Infants used the same cloth and magnet that were seen during the eye-tracking training phase. However, goal objects differed between these two phases of the experiment, as infants had previously observed a toy car being reached for in the training videos, whereas they reached for a toy train in the motor task. The colour of the magnetic train paired with each tool was counterbalanced between infants. For example, to help keep the task interesting and maintain attention, an infant may have completed the cloth trials with a yellow train, followed by magnet trials with a red train. Infants sat on a parent's lap facing the table. An experimenter sat next to the infant and the task was recorded by a second experimenter for later coding. During three pre- and post-training trials, infants were given the opportunity to perform the tool-reach to retrieve the toy train without any guidance from the experimenter (see figure 4a). The experimenter set up the task in front of the infant and then looked down at the table. If the infant was distracted or uninterested in the task, the experimenter tapped on the train to draw the infant's attention to the problem but did not provide more specific cues to prompt the infant's actions. The trial ended when the infant obtained the toy or when 30 seconds had elapsed. When the infant obtained the toy, the experimenter praised them by saying "well done" enthusiastically. If the infant did not obtain the toy, the experimenter did not say anything to the infant and proceeded with the next trial.

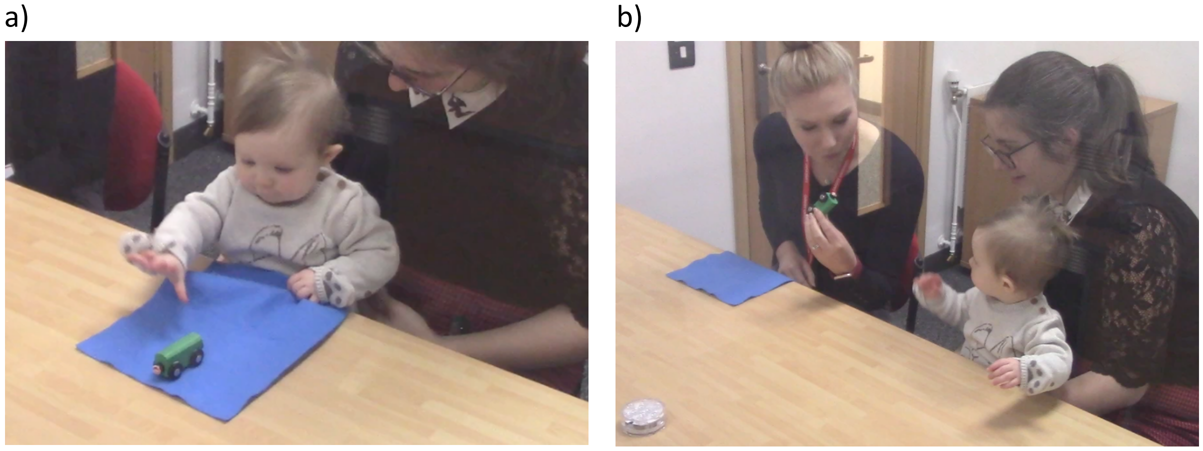

Figure 4. Experimental setup for motor task. a) Infants' tool-reach (cloth and magnet, pre- and post-training). Note that only the cloth-reach example is shown in this figure. b) Training demonstration.

Motor training trials

Between the pre- and post-training trials, infants completed training trials in which they received help from the experimenter if they were struggling to perform the action. Before initiating the training trials, the experimenter demonstrated the tool-reach action twice whilst highlighting the end goal. That is, the experimenter attracted the infant's attention by saying "look", then performed the tool-reach action whilst watching the toy as it drew near, followed by grasping the toy and examining it whilst expressing interest by saying "Ooh" (see figure 4b). In three training trials, the task was set up in front of the infant and, if the infant did not initiate the action on their own, the experimenter helped them to perform the tool-reach. For example, this could include helping the infant to pull on the cloth, aiding the infant in connecting the magnet to the toy train, or guiding the direction of the tool-reach. All trials ended after the infant either obtained the toy or did not act on either the tool or toy for longer than 30 seconds. The second experimenter timed each trial using a stopwatch and illuminated a remote-controlled LED light (attached to the table in the testing room) to mark the end of each trial.

Data Processing

Eye-tracking data

A limitation with study one (see chapter 2) is that we did not collect eye-tracking data during training or familiarisation trials. However, we decided that it was important to check that infants were paying attention to these trials. Therefore, in the current study an AOI covering the entire screen was used to assess whether infants attended to each video during training and familiarisation trials. Proportion of time spent looking to training and familiarisation trials was initially assessed for 55 infants who had provided fixation data in all phases of the eye-tracking task (i.e., training, familiarisation, and test trials). The average proportion of time that infants spent attending to training and familiarisation trials was 64% (*range*: 12% - 93%). Figure 5 shows a step change in infants' proportion of time looking to training and familiarisation trials around the 40% mark, with most infants looking to trials for 40% of the time. Based on this distribution of the data, infants who did not attend to at least one video of each action type (i.e., the three actions during training and the claw-reach during familiarisation) for a minimum of 40% of the time were excluded from all subsequent analyses (*N* = 7). 48 infants' data were remaining for assessment of fixation data during the freeze screen.

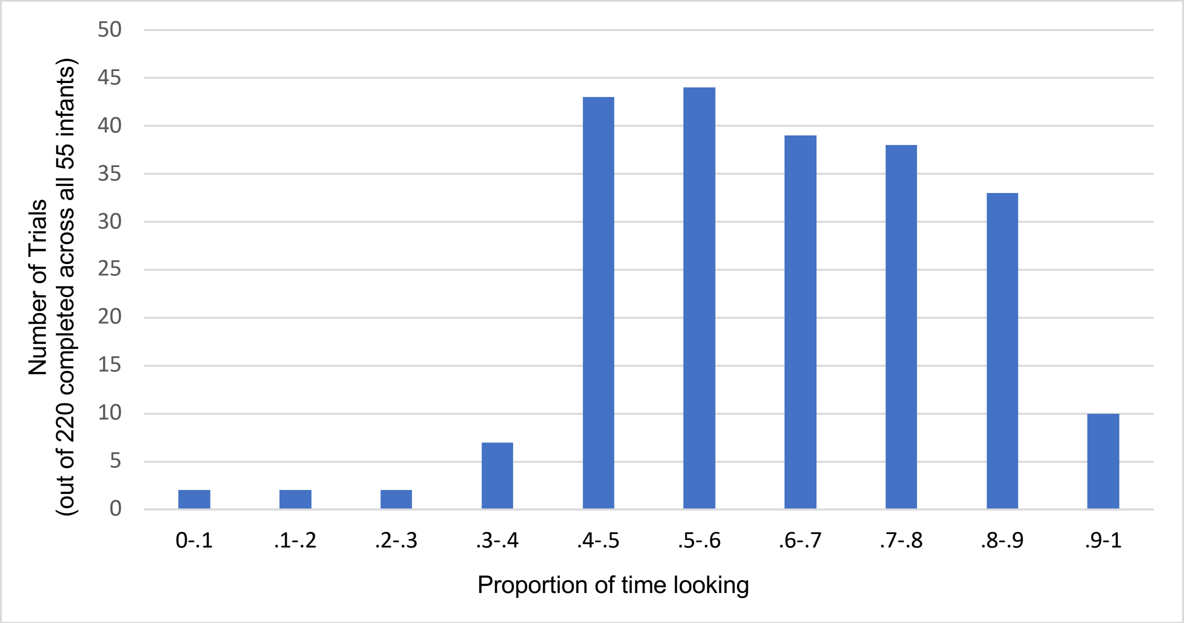

Figure 5. Proportion of infants' looking time across training and familiarisation trials. If an infant did not attend for more than 40% of any trial, their data were excluded from subsequent analyses.

Using the same criteria that we used in the previous chapter (see chapter 2), eye fixations to both the bear and ball were measured to infer whether infants encoded the swap in positions during the freeze screen. AOIs (matched in size, 170 x 210p) were defined over the bear and ball to ensure that infants fixated on at least one of these objects to encode the swap in locations. If infants did not fixate on either of these objects, we concluded that they did not encode the swap during this phase. If an infant did not encode the swap during the freeze screen, we used the same method to investigate whether the swap was encoded during initial test trials. This is because infants may not have noticed the swap in toy positions during the freeze screen but subsequently had the opportunity to notice the swap during test trials. If the swap was encoded during test trials, subsequent test trials were then coded for predictions. No trials were coded for predictions before the swap had been encoded, as fixations before this point were not likely to be anticipatory of the agent's intention. If we could not confirm that an infant encoded the swap after these investigations, their data were excluded from analyses as fixations recorded during test trials were likely to be arbitrary rather than predictive (*N* = 1).

In test trials, AOIs for the bear and ball (matched in size, 150 x 210p), as well as the left and right claw (matched in size, 90 x 650p) were defined along with an AOI covering the hand (120 x 240p, see figure 3d for a representation of AOIs). Fixations were coded from the exact moment the hand entered the screen on each test trial (after an average of 1.15 seconds, with natural variation between videos). This was precisely defined for each trial. Timestamps of fixation to each AOI determined the order of fixations and thus allowed us to code predictive looks. After fixating on the hand, we coded the first object that infants fixated on (hand-to-object prediction). As in previous visual prediction paradigms (Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018), fixations to a toy or claw without first fixating on the hand were not considered predictive anticipations of the familiarised action. When infants made hand-to-object fixations on the familiarised goal object or its newly corresponding claw, this was considered an *accurate prediction.* When infants made hand-to-object fixations on the familiarised claw, or the non-familiarised object (in the familiarised location) this was considered an *inaccurate prediction.* Fixations to claws were included as predictive looks because infants were familiarised with the action trajectory of the hand using the claw to grasp the toy. If infants do not understand the agent's two-step action of using the tool to reach the goal-object, we expect them to look from the hand to the claw that the agent manipulated during familiarisation (i.e., an inaccurate prediction in which infants are only predicting a one-step action). In contrast, infants who understand the agent's two-step action of using the claw to achieve the preferred toy should look to the claw that can reach for the familiarised toy (i.e., an accurate prediction). If infants did not fixate on the hand or did not fixate on another AOI following fixation on the hand, these trials were considered as *no anticipation* and were excluded from analyses. If infants did not make predictive anticipations on any test trial their data were excluded from analyses (*N* = 7). Saccades through AOIs were not coded as predictive looks. All looking time and fixation data were output by E-Prime (Psychology Software Tools, Pittsburgh, PA) and processed using Matlab (MATLAB, 2019). Each infant's proportion of predictive fixations was calculated separately for accurate and inaccurate predictions, with number of predictive trials as the denominator. Table 3 shows the proportion of infants in each condition who made accurate predictions, inaccurate predictions, and no anticipations within each test trial. Table 4 represents the various categories of performance patterns shown by individual infants across test trials. That is, the extent to which infants' prediction accuracy improved, declined, or varied from the first to last test trial in which accurate or inaccurate predictions were made. Infants are also categorised into those who made accurate predictions only or inaccurate predictions only. Average numbers of predictive trials indicate the extent to which infants made predictions (i.e., accurate or inaccurate) as opposed to no anticipations. Across both conditions most infants made inaccurate predictions only.

Table 3

Proportion of infants in each condition who made accurate predictions, inaccurate predictions, and no anticipations within each test trial.

	Trial 1	**Trial 2**	**Trial 3**	**Trial 4**
Comparison condition (*N* = 22)				
Accurate	14%	18%	36%	23%
Inaccurate	45%	45%	23%	27%
No anticipation	41%	36%	41%	50%
Control condition (*N* = 17)				
Accurate	35%	18%	18%	24%
Inaccurate	35%	35%	18%	47%
No anticipation	29%	47%	65%	29%

Table 4

Patterns of prediction accuracy from first to last predictive test trial, as well as the average number of test trials in which infants made a prediction (i.e., accurate or inaccurate) as opposed to no anticipation.

Accuracy across predictive test trials	**%**	**Average number of predictive test trials**
Comparison condition (*N* = 22)		
Accuracy improved (inaccurate to accurate)	18%	2.75
Accuracy declined (accurate to inaccurate)	14%	2.33
Accuracy varied	14%	3.33
Accurate predictions only	14%	2.33
Inaccurate predictions only	41%	1.78
Control condition (*N* = 17)		
Accuracy improved (inaccurate to accurate)	12%	3.50
Accuracy declined (accurate to inaccurate)	6%	2.00
Accuracy varied	6%	3.00
Accurate predictions only	29%	1.80
Inaccurate predictions only	47%	2.25

Motor data

Across pre- and post-training trials, infants' cloth- and magnet-reach actions were coded for the extent to which infants engaged in well-structured planful actions. An action was coded as planful if the infant maintained visual attention to the toy while using the tool to draw it near in one continuous movement, then obtained or touched the toy within three seconds of completion of the pull (Gerson et al., 2015b; Sommerville & Woodward, 2005). For each tool, every infant's motor efficiency during pre-training was operationalised by calculating the proportion of planful actions made across the three trials. Propensity to learn was operationalised by calculating infants' improvement in planful actions between pre- and post-training (i.e., proportion planful at post-training – proportion planful at pre-training). To check for inter-rater agreement, 30.77% of infants' motor data was double coded by a second coder. There was moderate agreement between coders for pre-training trials, $\kappa = .77$, $p < .001$, as well as post-training trials, $\kappa = .65$, $p < .001$. No infants made planful magnet-reach actions during pre- or post-training, so only the cloth-reach action was included in analyses. One infant did not complete the cloth-reach task.

Analyses and Results

Analyses

Preliminary analyses were run to check for any potential correlation between age and our dependent variables (proportion of accurate predictions, motor ability, or propensity to learn). We used a generalised linear model (GLZM) to test for any effects of sex, goal objects acted upon during observational training (i.e., white, red, or yellow car), the order of tool-use actions observed during training (i.e., cloth first or magnet first), as well as location of the goal object during test trials (i.e., left or right), upon proportion of accurate predictions as the ordinal dependent variable. We also used a GLZM to test for any influence of observational training condition upon infants' performance in the motor task. Additional GLZMs replicated analyses carried out in chapter 2, to clarify whether the recruitment of younger infants in the current study, as well as the change in our control training paradigm, led to similar findings. We then used Mann Whitney tests to assess whether predictions made in each condition were significantly different from chance level (i.e., 50%). In addition to investigating proportion of accurate predictions across all four test trials, we used a chi square test to assess the first type of prediction that infants made between conditions. This was in line with our analyses in chapter 2 as well as Cannon and Woodward's (2012) paradigm that used a single test trial to assess accurate visual prediction of an incomplete action.

Planful actions made by infants for each trial of the motor task were coded as binary (i.e., planful or not planful), so we were unable to examine changes in planfulness across trials using a repeated measures ANOVA. A generalised estimating equation (GEE; Ballinger, 2004; Gerson & Woodward, 2013; Zeger et al., 1988) is a more appropriate technique to test for any potential differences among repeated observations. GEEs are an extension of GLZMs that can be used to analyse binary or ordinal repeated measures. We used a GEE to estimate predicted probability of changes in planfulness between pre- and post-training trials. Because each infant received a binary code for each trial (i.e., three pre-training trials and three post-training trials), predicted probability in each trial translated to the estimated percent of infants who were predicted to be planful in their actions within pre- and post-training. For our binary dependent variable (i.e., planful or not planful), we specified a binomial probability distribution and a logit link function with an unstructured working correlation matrix structure. The output of a GEE includes Wald χ^2 values for main effects and interactions in a model, as well as estimated marginal means that can be examined with pairwise comparisons.

Our main analyses consisted of three separate GLZM analyses to test for the effects of condition, motor ability, and propensity to learn upon proportion of accurate predictions. These types of models are comparable to ANOVA, except with an ordinal rather than continuous outcome measure. This was more appropriate for the nature of our data, as proportion of accurate predictions were calculated across a maximum four trials, meaning that proportion ranged ordinally (i.e., 0%, 25%, 33%, 50%, 67%, 100%). As this ordinal data were not normally distributed, GLZMs were conducted. Any significant interactions were followed up using ordinary least square (OLS) regression models. This output also included estimated means for proportion of accurate predictions. Although these analyses investigated the same hypotheses, this analysis pipeline deviated from that described in our registered report (Open Science Framework, Findlay & Gerson, 2019) due to learning about these more appropriate techniques to investigate our research questions.

Next, we used a GLZM to assess the first type of prediction that infants made between conditions whilst also taking motor ability into account. In this case, first prediction type was a binary dependent variable (i.e., accurate vs inaccurate). Final analyses of variance (ANOVAs) were conducted to investigate differences in infants' latency to make accurate versus inaccurate predictions both across and between conditions, as well as assessing the potential effect of infants' motor ability on speed to predict.

Results

Prior to our main analyses, a Spearman's correlation assessed whether age was correlated with any of our dependent variables (proportion of accurate predictions, motor ability, and propensity to learn). This revealed that age was positively correlated with infants' motor ability, *r*(37) = .40, *p* = .01. We therefore controlled for age in all subsequent analyses that included infants' motor ability at pre-training or propensity to learn (i.e., difference in proportion of planful actions from pre- to post-training) as a variable.

Next, we assessed whether sex, goal objects acted upon during observational training (i.e., white, red, or yellow car), the order of tool-use actions observed during training (i.e., cloth first or magnet first), or location of the goal object during test trials (i.e., left or right) had any effect on infants' proportion of accurate predictions. Proportion of accurate predictions was entered as the dependent variable in a GLZM. Sex, goal object during training, tool-use order, and goal location were entered as predictor variables. No significant effects of sex, goal objects during training, or goal location emerged (*ps* > .13). However, a significant main effect of tool-use order was revealed, Wald χ^2 (1) = 10.16, *p* = .001, β = -.51. Infants who observed the cloth-reach first made a higher proportion of accurate predictions (49%) than infants who observed the magnet-reach first (31%). However, due to the way in which the task was counterbalanced, infants who saw the cloth-reach first were also familiarised with the bear as the goal object. In contrast, infants who saw the magnet-reach first were familiarised with the ball as the goal object. Therefore, it is unclear whether this effect is entirely due to the order in which infants observed the tool-use actions, or potentially due to a preference for looking at the bear over the ball. We therefore controlled for this effect in subsequent analyses that included proportion of accurate predictions as the dependent variable.

To test for any potential effect of observing tool-use actions during the eye-tracking task followed by performing the same actions in the motor task, we checked whether there was any effect of observational training condition upon infants' motor ability. Proportion of planful actions at pre-training was entered as the ordinal dependent variable in a GLZM. Condition was entered as the predictor variable and age was entered as a continuous covariate to control for the relation between age and infants' motor ability. This revealed no significant effect of condition, Wald χ^2 (1) = .34, *p* = .56, β = .07.

Before assessing the effects of infants' motor ability or propensity to learn upon proportion of accurate predictions, we first decided to re-produce our main analysis from chapter 2 to assess whether we find similar results in the current study. As previously discussed, the current study recruited a younger age group and involved an adaptation of the control training paradigm. As our previous analysis found an effect of tool-use order on proportion of accurate predictions, we used a GLZM to control for this effect whilst investigating the difference in proportion of accurate predictions between conditions. Proportion of accurate predictions was entered as the dependent variable. Condition and tool-use order (control variable) were added as predictors. The main effect of condition was marginally significant, Wald χ^2 (1) = 2.91, *p* = .088, β = .30. Surprisingly, this trend was the reverse of that seen in chapter 2, such that infants in the control condition made a higher mean proportion of accurate predictions (43%) than those in the comparison condition (35%). Comparisons against chance (i.e., 50%) indicated that infants in the comparison condition systematically made location-based predictions (rather than goal-based), *U* = 165.00, *p* = .04. However, predictions made by infants in the control condition did not significantly differ from chance, *U* = 136.00, *p* = .75. The GLZM also revealed a marginally significant interaction between condition and tool-use order, Wald $\chi^2(1) = 3.13$, *p* = .08, β = -.42. We investigated the nature of this interaction using an ordinary least square (OLS) regression model with condition and tool-use order as predictor variables. For infants in the control condition, this model revealed a significant difference in the estimated mean proportion of 24% accurate predictions made by infants who saw the magnet-reach first, compared with an estimated mean of 65% accurate predictions made by infants who saw the cloth-reach first. For infants in the comparison condition, this model revealed an estimated mean of 36% accurate predictions made by infants who saw the magnet-reach first, compared with a mean proportion of 34% accurate predictions made by infants who saw the cloth-reach first. This difference did not reach significance. See table 5 for coefficient factors.

Table 5

Coefficients for accurate predictions made by infants in the control condition who observed the magnet-reach first (a) and the mean differences with the comparison condition (b_1) and infants who saw the cloth-reach first ($b_2 + b_3$).

	β coefficient	*SE*	t-value	p-value (2-tailed)
Main effects				
a (constant)	.24	.13	1.87	.07
comparison (b_1)	.12	.17	.71	.49
cloth-reach first (b_2)	.41	.19	2.16	.04
Interaction effects				
condition [comparison] * tool-use order [cloth-reach first] (b_3)	-.42	.25	-1.68	.10

Like in chapter 2, we also assessed whether there was any difference between conditions for the first trial in which each infant made a prediction (accurate or inaccurate). For example, if an infant did not make a prediction in the first test trial but did make a prediction in the second, we used the prediction data from the second trial. A chi square test assessed the frequencies of accurate and inaccurate first predictions between conditions and found no significant differences between conditions, χ^2 (1, *N* = 40) = .22, *p* = .64.

In our next analysis, we conducted a GEE to investigate whether infants' planful actions improved between pre- and post-training. Recall that this analysis only included infants' cloth-pulling actions, as no infants made planful magnet-reach actions during pre- or post-training. Planfulness (planful or not planful) was entered as a binary dependent variable. Time (pre- vs post-training) and trial number (1 to 3) were entered as predictor variables. Age was entered as a covariate. No significant main effects of, or interactions between, time and trial emerged (*ps* > .37). However, there was a significant main effect of time, Wald χ^2 (1) = 4.21, *p* = .04. Unexpectedly we found that across conditions, infants made a higher estimated mean proportion of planful actions at pre-training (55%, *SE* = .06, 95% CI, .44 to .66) than post-training (49%, *SE* = .07, 95% CI, .36 to .62). There was also a significant interaction between time and age, Wald χ^2 (1) = 3.99, *p* = .05. We investigated the nature of this interaction using an ordinary least square (OLS) regression model with time (pre- vs post-training) and age as predictor variables. For the purpose of this analysis, age was re-coded into a dummy variable split into younger than the mean (*N* = 20) and older than the mean (*N* = 19; M_{age} = 10.08). This enabled us to investigate the interaction between time and age at two levels. For younger infants, this model revealed a difference in the estimated mean of 32% planful actions at pre-training, compared with an estimated mean of 53% planful actions at post-training that approached significance. For older infants, this model revealed an estimated mean of 60% at pre-training, compared with an estimated mean of 50% at post-training, which also approached significance (see figure 6). Therefore, younger infants improved in their performance of planful actions from pre- to post-training, whereas older infants' performance decreased slightly, potentially due to boredom. These findings further justify our decision to control for age in all analyses that include motor ability and propensity to learn (i.e., difference in proportion of planful actions from pre- to post-training). See table 6 for coefficient factors.

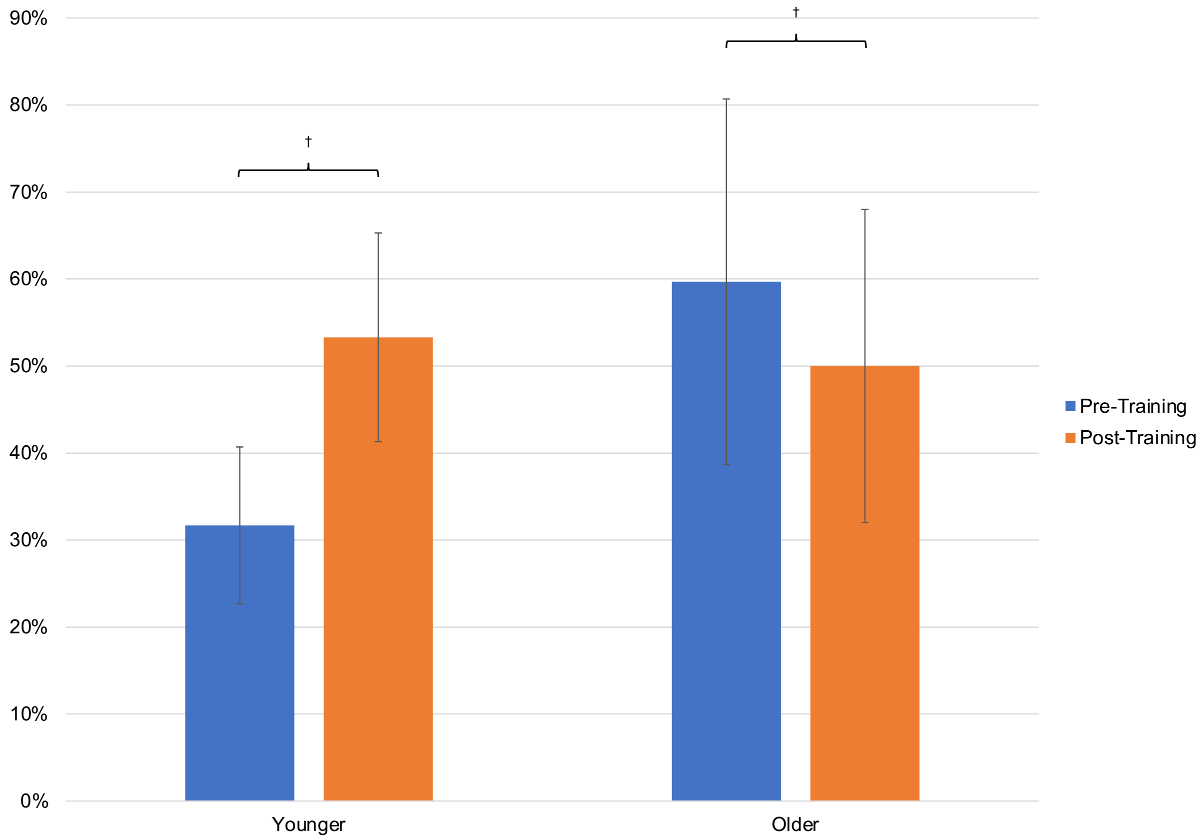

Figure 6. Proportion of planful actions for younger and older infants at pre- and post-training. Error bars represent standard error. $^{\dagger}p = .08$.

Table 6

Coefficients for proportion of planful actions performed by younger infants during pre-training (a) and the mean differences with post-training (b_1) and older infants ($b_2 + b_3$).

	β coefficient	*SE*	t-value	p-value (2-tailed)
Main effects				
a (constant)	.32	.09	3.75	.000
post-training (b_1)	.22	.21	1.81	.08
older (b_2)	.28	.12	2.31	.02
Interaction effects				
time [post-training] * age [older] (b_3)	-.31	.18	-1.79	.08

Effects of condition and motor ability on accurate predictions

In the following analyses, we investigated the possibility that infants' existing motor ability may help to facilitate accurate predictions. Proportion of accurate predictions was entered as the dependent variable in a GLZM. Condition (comparison vs control), motor ability at pre-training, and tool-use order during observational training (control) were entered as predictor variables. Age was also entered as a covariate. This revealed a main effect of condition that was marginally significant, Wald χ^2 (1) = 3.12, *p* = .08, β = -4.25. There was also a significant interaction between condition and motor ability, Wald χ^2 (1) = 5.05, *p* = .03, β = 1.13, as well as an interaction between condition and age that approached significance, Wald χ^2 (1) = 3.27, *p* = .07, β = .38. We investigated the main effect of condition using an ordinary least square (OLS) regression model with condition and age (centred to the mean) as predictor variables. This revealed a non-significant trend such that infants in the control condition made a higher estimated mean of 44% accurate predictions, compared with an estimated mean of 34% accurate predictions made by infants in the comparison condition, β = -.10, *SE* = .13, *p* = .45. To investigate the nature of the interaction between condition and motor ability, we next added motor ability to the OLS model. For the purpose of this analysis, motor ability at pre-training was re-coded into a dummy variable split above the median (i.e., 33% planful) into a high ability group (*N* = 19) and at or below the median into a low ability group (*N* = 20). This enabled us to investigate the interaction between condition and the two levels of motor ability. Within the control condition, this model revealed a significant difference between the estimated mean of 70% accurate predictions for infants in the low ability group, compared with an estimated mean of 24% accurate predictions for those in the high ability group. Within the comparison condition, the estimated mean proportion of accurate predictions within the low ability group was 36%, compared with 30% for infants in the high ability group. This difference did not reach significance (see figure 7). See table 7 for coefficient factors.

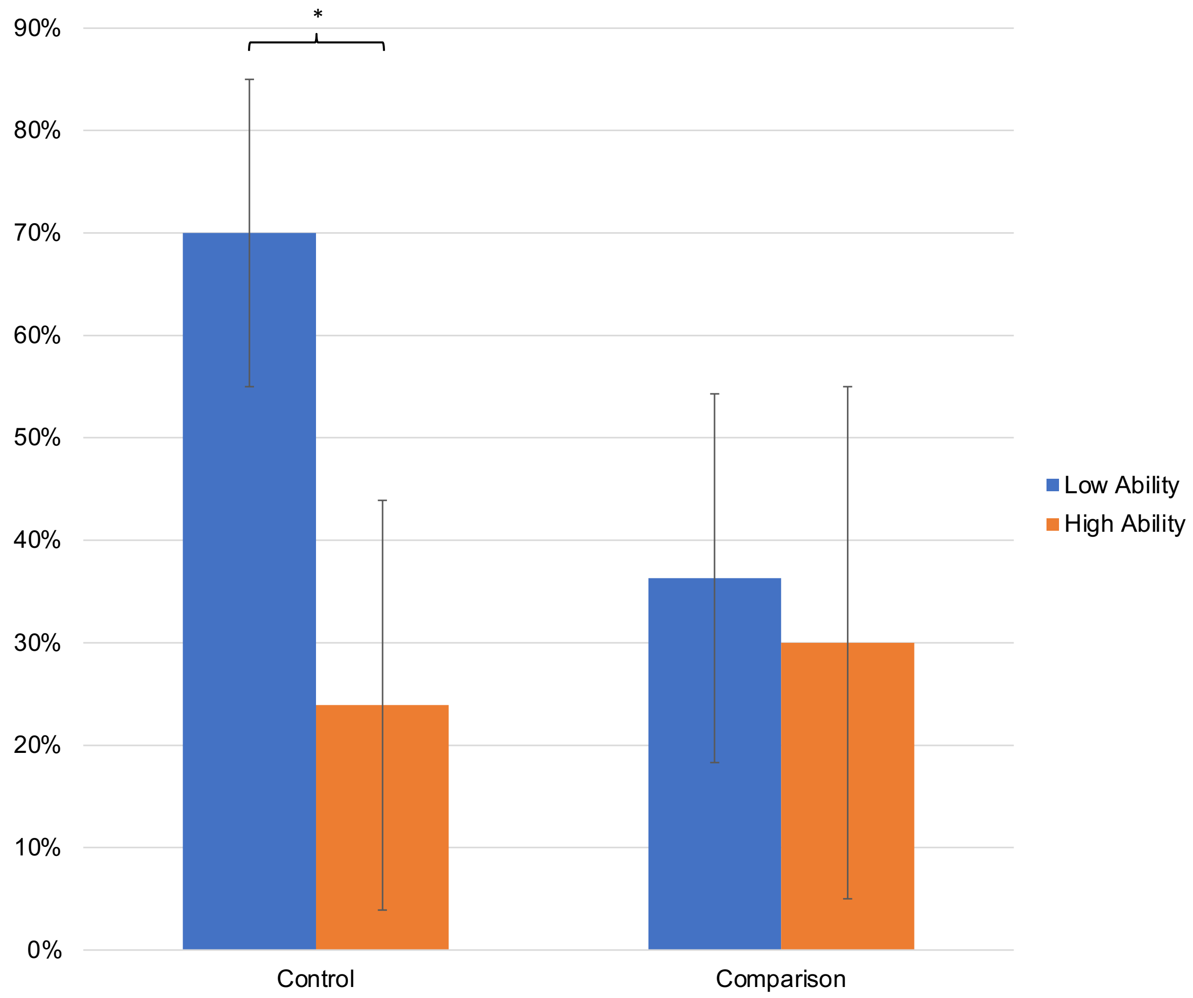

Figure 7. Proportion of accurate predictions for infants showing high and low motor ability within control and comparison conditions. Error bars represent standard error. **p* < .05.

Table 7

Coefficients for accurate predictions made by infants in the control condition with low motor ability (a) and the mean differences with the comparison condition (b_1), age (b_3), and high motor ability ($b_2 + b_4$).

	β coefficient	*SE*	t-value	p-value (2-tailed)
Main effects				
a (constant)	.70	.15	4.82	.000
comparison (b_1)	-.34	.18	-1.85	.07
high motor ability (b_2)	-.46	.20	-2.30	.03
age (centered) (b_3)	.18	.10	1.82	.08
Interaction effects				
condition [comparison] * motor ability [high] (b_4)	.40	.25	1.58	.13

To investigate the nature of the interaction between condition and age that was marginally significant in our GLZM, we used an ordinary least square (OLS) regression model with condition and age (split into older and younger than the mean) as predictor variables and proportion of accurate predictions as the dependent variable. This revealed no significant main effects or interactions between variables (*ps* > .10). However, there was a non-significant trend such that in the comparison condition, older infants made a higher proportion of accurate predictions (47%) than younger infants (22%). The reverse trend was seen in the control condition, with younger infants making a higher proportion of accurate predictions (48%) than older infants (36%).

We also analysed our data for the first prediction (accurate or inaccurate) that each infant made across all four test trials whilst also taking motor ability into account. First prediction was entered as the binary dependent variable in a GLZM. Condition, motor ability at pre-training, and tool-use object during observational training (control) were entered as predictor variables. Age was also entered as a covariate. No significant main effects or two-way interactions emerged (*ps* > .50).

Effects of condition and propensity to learn on accurate predictions

In our next analysis we investigated the possibility that infants' propensity to learn to perform an action (i.e., difference in planful actions from pre- to post-training) may aid infants in their ability to make accurate predictions. As there was a lack of overall improvement from pre- to post-training trials, we split our sample into two groups of infants who did improve (*N* = 17) or did not improve (*N* = 19). Proportion of accurate predictions was entered as the dependent variable in a GLZM. Condition and motor improvement (did vs did not improve) were entered as predictor variables. Given our previous finding that younger infants improved from pre- to post-training and older infants' performance decreased, age was also included as a covariate in this model. No significant main effects of, or two-way interactions between, any of these variables emerged (*ps* > .39).

Prediction latency

Next, we investigated the possibility that infants may make inaccurate predictions more slowly than accurate predictions due to an increase in cognitive effort (Krogh-Jespersen & Woodward, 2014, 2018, also see chapter 2). We were interested in whether infants' speed at making accurate and inaccurate predictions differed both across and between conditions, as well as whether infants' motor ability had any potential effect. As prediction latency data were non-normally distributed with skewness of 3.11 (*SE* = .33) and kurtosis of 12.16 (*SE* = .64), a natural log transformation was performed on the data before running our analyses. Latency to predict was entered as the dependent variable in an ANOVA. Prediction type (accurate vs inaccurate), motor ability at pre-training (high vs low), and condition (comparison vs control) were entered as predictor variables. There were no significant main effects or interactions between variables (*ps* > .3).

Discussion

Within the action understanding literature, there is a well-documented finding that as infants learn to produce actions, their understanding of other people's intentions also develops (e.g., Gerson et al., 2015b; Gerson & Woodward, 2014b, c; Sommerville et al., 2005). Experience producing actions provides a crucial underpinning for making sense of new actions that infants come across, especially young infants with less motor experience (e.g., Gerson & Woodward, 2012). By the end of the first year, infants begin to master more complex actions such as tool-use (e.g., Sommerville & Woodward, 2005). Whilst motor experience is important, other cognitive factors such as comparison can also help advance infants' learning about intentional actions, including actions that infants cannot self-produce (Gerson & Woodward, 2014a). Although caution should be taken as results were marginally significant, our findings in chapter 2 provide some support for the role of observational comparison in infants' understanding of a novel action. That is, 12-month-olds were able to generalise intentional understanding that they had learned from observational comparison training to accurately predict the outcome of a never-before-seen action. The current study replicated the paradigm used in chapter 2 with a younger sample of 10-month-olds and a slight change to the control condition to improve validity. That is, infants in the control condition viewed a cloth-reach and a magnet-reach action, followed by a familiar hand-reach action. These stimuli were identical to those in the comparison condition, except the familiar hand-reach was viewed first in the comparison training intervention. Also, the same goal-object was obtained in every example. This differed from our original control training paradigm with 12-month-olds, who only observed one single example of a tool-reach action (i.e., a cloth-reach or a magnet-reach) which was repeated three times. The colour of the goal object also changed in our original control condition. Therefore, the stimuli observed between the control and comparison conditions in the current study were better matched in terms of perceptual and relational similarity. We also included an additional motor task to measure infants' existing ability to perform a cloth-reach action as well as their propensity to learn to perform planful cloth-reach actions from a training intervention. Infants completed the motor task after the eye-tracking task so as not to motorically prime them before undergoing our observational training intervention. We hypothesised that we would see a larger difference in the proportion of accurate predictions of the novel action made between comparison and control conditions in 10-month-olds, as they are likely to be less motorically advanced than 12-month-olds. Because of this, younger infants may be more reliant on cues for comparison than older, more motorically efficient infants, who may accurately predict the outcome of a novel action without the need for comparison (Sommerville & Woodward, 2005). Importantly, we were able to consider the possibility that variability in infants' accurate predictions may be characterised by individual differences in motor skill or propensity to learn.

When comparing our results from the current study to those that we gained in the previous chapter, our findings revealed the reverse trend to that which we expected based on our findings with older infants. That is, infants in the control condition made a higher proportion of accurate predictions than infants in the comparison condition. Also, infants in the comparison condition systematically made inaccurate predictions to the familiarised location (i.e., location-based predictions), rather than the familiarised goal object (i.e., goal-based predictions), whereas infants in the control condition performed at chance level. In addition to this, our main finding revealed that infants in the control condition who were less motorically efficient at pre-training made a higher proportion of accurate predictions than infants in the same condition who were more motorically efficient. These findings lead us to question whether infants in the control condition were somehow learning from the training trials. We speculate that observing both the cloth-reach and the magnet-reach action before seeing the familiar hand-reach may have enabled infants to compare between these actions. In addition, this control condition was particularly helpful for infants who were less motorically efficient than others and thus could only rely on observed intentional cues rather than drawing upon previous motor experience (Biro & Leslie, 2007). It seems that the order in which infants observed the hand-reach action last in the control condition may be more optimal for comparison between actions than we initially expected. It could be that observing two tool-use actions first provides infants with a foundation for comparison with the familiar hand-reach, rather than vice versa. Previous work by Gerson & Woodward (2012) found that 7-month-olds are able to learn about complex tool-use actions through a process of comparison between an unfamiliar tool-reach and their own familiar hand-reach examples. This involved passing the infant a toy using a claw-tool. This meant that infants observed the unfamiliar tool-use action prior to comparing it with their own familiar hand-reach action. Despite differences in design, it is possible that older 10-month-olds in the current study could learn from a similar order of unfamiliar versus familiar actions through a process of observational comparison. Our findings suggest that this may be more beneficial for intentional learning than when observing the hand-reach first. However, further research is needed to confirm this speculation.

The current study also investigated whether infants' proportion of planful cloth-reach actions improved between pre- and post-training trials as a result of our active training intervention. Interestingly, we found that the proportion of younger infants' planful actions increased from pre- to post-training, whereas the proportion of older infants' planful actions decreased from pre- to post. Therefore, it seems that younger infants who were less motorically advanced were learning and improving from the motor training. On the other hand, older infants who were more motorically advanced did not benefit from the training and became less engaged with the task. Another interesting finding that came to light in the current study is that across conditions, infants who observed the cloth-reach first during observational training made a higher proportion of accurate predictions than infants who observed the magnet-reach first. However, due to the way that the task was counterbalanced, the effects of seeing the cloth first versus the magnet first could not be split apart from the effects of the familiarised goal object. That is, all infants who viewed the cloth first were familiarised with the bear, whereas all infants who viewed the magnet first were familiarised with the ball. We speculate that the cloth-reach action may be familiar enough for infants to use as a basis for comparison with the magnet-reach. That is, the cloth-reach is similar to a blanket such that the means of pulling on a cloth or blanket is much more familiar to infants than using a magnet. To some extent, this idea is supported by infants' inability to perform the magnet-reach action in the current study. Our findings imply that seeing the cloth-reach action before the magnet-reach action enabled infants to compare the more familiar tool-reach with the less familiar tool-reach action, respectively. In contrast, seeing the magnet-reach first is too unfamiliar for 10-month-olds to compare with the cloth-reach action. This interpretation could raise interesting questions for further work that could investigate the optimal order for comparison between structurally similar actions that differ in complexity. There was a marginally significant difference in accurate predictions between infants in the control condition who had seen the cloth-reach first (65% accurate) and infants who had seen the magnet-reach first (24% accurate). This gives us reason to believe that viewing the cloth-reach action first is more helpful than viewing the hand-reach action first. It may be that infants are more sensitive to learning about the more familiar cloth-reach action at this age. Or, by seeing the hand-reach before the cloth-reach, infants may learn that the experimenter can reach far enough to obtain the toy without the need to pull on the cloth in subsequent training trials.

As detailed earlier in this chapter, infants' ability to perform and propensity to learn a magnet-reach action, in addition to the cloth-reach action, was also assessed. However, there were floor effects such that 10-month-old infants could not perform the magnet-reach action. This was likely due to the complex structure involved in the magnet-reach as compared to the cloth-reach. The magnet-reach action could be defined as a three-step structure. That is, infants were required to pick up the magnet, connect it to the toy train, and pull the train close. This action also came with an added level of complexity such that the magnet must remain in the correct orientation in order to attract and connect with, rather than repel, the toy train. In contrast, the cloth-reach action can be defined as a two-step structure. Infants were simply required to grasp the cloth (that already supported the toy train) and pull it close. Future research investigating infants' motor ability should assess and scaffold the production of actions that are similar in complexity. For example, pulling a hook-shaped cane to retrieve a toy that is located in the crook of the cane arguably involves a two-step grasp and pull action that is similar to a cloth-reach action (Sommerville et al., 2008). This further leads us to question whether 10-month-olds are able to compare a magnet-reach action with a familiar hand-reach. To some extent, our findings suggest that this may be the case, as infants who viewed the magnet-reach action before the cloth-reach did not make as many accurate predictions as infants who observed the cloth-reach first. It may be that 10-month-olds are too young to compare between these examples, whereas 12-months-olds could successfully compare between them (see chapter 2).

In sum, the current study obtained unexpected results that were the opposite to what we hypothesised. That is, infants who were not efficient at performing tool-use actions and who did not undergo observational comparison training made the highest proportion of accurate predictions. These findings do not support literature that suggests a role for existing motor skill in facilitating infants' understanding and prediction of other's intentional actions (e.g., Gerson et al., 2015b; Gerson & Woodward, 2014b, 2014c; Sommerville et al., 2005, 2008), but provide new insight into alternative mechanisms. Although we did not expect our control training intervention to promote comparison between complex and familiar actions, this intervention was more effective for facilitating infants' intentional learning and brings interesting questions for future directions. Further work is needed to investigate the possibility that infants who rely on cues for comparison due to their lack in motor skill may learn better when observing complex actions followed by a familiar action, instead of vice versa.

Chapter 4

More exchanges and less joint contact of toys in infant-directed action are associated with infants' prediction of a novel intentional action

Abstract

The previous chapters of this thesis have investigated the role of cognitive comparison for facilitating infants' development of intentional understanding. In addition to this, social factors are important for supporting infants' learning of action structures. To help infants understand actions, parents often modify their actions in an exaggerated way, known as infant-directed action, or 'motionese' (Brand et al., 2002, 2007). It is believed that infant-directed actions are a social facilitator that help infants understand observed actions. However, evidence to support this notion is rare. In study 3, we investigated the possibility that motionese can support infants' understanding of a novel intentional action. The current study utilised the same infant sample as that in chapter 3. However, a larger number of infants were included (N = 43) as exclusion criteria for our observational training intervention did not apply. Mothers demonstrated challenging toys to their 10-month-old infants. We coded motionese on a global scale across eight dimensions (adapted from Brand et al., 2002) as well as on a fine-grained scale (adapted from Brand et al., 2007). Infants then completed an eye-tracking task to measure their ability to visually predict the outcome of an unrelated goal-directed action. Our main findings revealed that frequent exchanges and less joint contact of toys (features of motionese) were associated with infants' ability to make accurate visual predictions in our eye-tracking paradigm. These results imply that exchanges of objects could be used to mark the boundaries between sub-steps of actions (e.g., repeatedly stacking rings on top of one another) that contribute to an overarching goal (e.g., completing a tower of rings). Although the reason for the negative association between joint contact of toys and accurate predictions is less clear, we speculate that joint contact of toys may be characteristic of mothers interfering with their infants' exploratory behaviour. To our knowledge, this is the first study to demonstrate that these interactive dimensions of mother-infant interaction support infants' broader understanding of observed intentional actions.

Introduction

When observing a continuous stream of action, infants must identify the meaningful components of the action to understand its purpose. Streams of actions are present in various day-to-day interactions that infants have with their parents. To reiterate my example of painting a picture in chapter 1, this involves a series of action steps including preparing the paint palette, applying paint to the paint brush, then applying the paint to paper. To help infants understand the meaning of such processes, an adult might exaggerate each action, as well as transitions between actions, so that infants can better identify the intentional sub-steps (e.g., applying paint to the brush vs brush to paper). Infant-directed action modifications, also known as 'motionese', help to highlight the structure of the action whilst maintaining infants' attention (e.g., Brand et al., 2002, 2007). Motionese is derived from the language phenomenon of 'motherese,' which refers to infant-directed speech modifications such as exaggerated intonation and higher pitch (e.g., Fernald & Kuhl, 1987; Papoušek et al., 1991). The first study to assess the phenomenon of motionese identified infant-directed modifications across eight dimensions: range of motion, rate, enthusiasm, proximity, repetitiveness, punctuation, interactiveness, and simplification (Brand et al., 2002). Much like motherese in the auditory domain, such modifications were significantly more exaggerated when mothers demonstrated toys to their infants as compared with another adult.

Brand et al. (2007) further investigated two rich dimensions of infant-directed action: interactiveness and simplification. These dimensions had initially been coded using a global rating scale that did not capture the full breadth of these dimensions (Brand et al., 2002). In a more precise and fine-grained manner, Brand et al. (2007) subsequently quantified interactiveness through fine-grained coding of eye gaze bouts, number of object exchanges and joint contact. Simplification was operationalised as the number of different action types between exchanges. More object exchanges, eye gaze, and fewer action types per turn were found in demonstrations to infants compared with adults. Brand et al.'s (2002, 2007) research indicates that the way in which mothers naturally modify their actions when interacting with their infant helps to maintain the infant's attention whilst highlighting the structure and purpose of the action.

Another study by Brand and Shallcross (2008) found that both 6- to 8-month-olds as well as 11- to 13-month-olds prefer infant-directed action over adult-directed action, paying more attention to infant-directed actions when simultaneously presented with videos of both action types. In addition to this, Koterba and Iverson (2009) investigated how individual dimensions of motionese may play varying roles in capturing 8- to 10-month-olds' attention as well as subsequent exploratory behaviour. The authors specifically assessed the individual and combined roles of repetition and range of motion, which they termed *amplitude.* They found that infants attended longer to caregiver displays involving enhancement of at least one of these parameters (e.g., high in repetition, high in amplitude, or both) compared with still displays. Interestingly, displays that were high in both repetition and amplitude were no more effective at gaining infant attention than displays containing just one enhanced parameter.

Also, displays that were high in repetition led infants to spend more time banging and shaking the object, whereas low repetition led to more turning and rotating behaviours. The authors speculated that infants may have been attempting to imitate the behaviour after seeing more repetitions, thus focusing on the movement involved in their caregiver's display. In contrast, infants who saw low repetition displays engaged in examining behaviours (turning and rotating). Based on attentional differences as well as infants' varying attempts to imitate actions that differ in terms of motionese modulations, we have some theoretical reason to believe that infant-directed action could help infants to understand actions, at least in the long-term. However, this has yet to be investigated directly. As acknowledged by Brand and Shallcross (2008), attention to infant-directed actions does not imply that infants necessarily understand the intentions behind actions. Therefore, the current study investigates the relation between parents' use of motionese and infants' broader understanding of goal-directed actions.

Further to the broad modulations of motionese, research also demonstrates that the structure of an action is highlighted differently dependent on whether the goal or the means (i.e., the tool or manner in which the goal is achieved) is important. In a study by Nagai and Rohlfing (2008), mothers and fathers demonstrated cup stacking to their 8- to 11-month-olds, in which the end-goal of nested cups was important. In contrast, parents also demonstrated a means-crucial salt sprinkling task in which the sprinkling action was important. The authors found that during the goal-crucial task, parents emphasised the initial and final states of the stacking cups by using long pauses before and after the demonstration. In the means-crucial task, parents highlighted the shaking action of the salt dispenser. In related action demonstration research, Fukuyama et al. (2015) investigated how mothers' action demonstration is modified depending on their infants' action skill. In a cup-nesting demonstration, mothers with infants aged 11- to 13-months increased the length and variance of their movement (measured using motion tracking) after infants performed a task irrelevant manipulation (e.g., banging the cup), whereas these modifications decreased after infants performed a relevant manipulation (e.g., nesting one cup inside another or attempting to do so). These modifications were not found when mothers demonstrated the cup nesting to 6- to 8-month-old infants, who did not have the motor skill to perform task-relevant manipulations. Another recent study that also took evaluations of motor skill into account was by van Schaik et al. (2019), in which optical motion tracking was used to measure modulations of parents' infant-directed actions compared with adult-directed actions. Parents demonstrated novel objects with opaque action effects to their 14-month-old infants as well as an adult confederate. All cylindrical objects were similar in appearance, meaning their action effects were opaque if one was not familiar with them. The difference between each object was only identifiable by colour and each object produced a single action effect, such as shaking to create a rattling sound, or pressing to turn on a light. The authors coded infants' success at producing the objects' action effects during the demonstration, as well as their memory of the action after the demonstration. In accordance with other motionese research (e.g., Brand et al., 2002; Rutherford & Przednowek, 2012), parents performed the actions in closer proximity, demonstrated for longer, and repeated demonstrations more often when showing the objects to their infant compared with an adult. In addition, infants learned better from actions performed closer than adult demonstrations, but not too close, identifying a 'just right' modulation in proximity. This modulation in proximity was related to infants' learning and memory of the demonstrated actions, as well as parents' prior evaluations of their infants' motor abilities. Both demonstrations that preceded infants' successful attempts at producing actions and parents who evaluated their infant's motor ability to be average performed demonstrations further away from the infant; as opposed to demonstrations that preceded failed attempts or parents who evaluated their infant to have high motor ability. The authors speculated that this could be so that, for infants who need more assistance, parents had the space to demonstrate actions that were larger than those demonstrated in close proximity. This 'just right' level of proximity could help maintain infants' attention and allows for large modifications to highlight action functions. Taken together, the findings from van Schaik et al. (2019) and Fukuyama et al. (2015) suggest that parents dynamically modify their infant-directed actions dependent on their infants' motor proficiency. Therefore, social and motor factors may complement one another to differentially facilitate infants' learning based on the infant's ability. Using the motor task described in chapter 3, the current study examines how infants' motor proficiency may relate to differential modulation in mother's demonstrations of toys.

It is possible that infants can understand other people's actions that they cannot produce themselves (e.g., Hunnius & Bekkering, 2010) as well as actions that they have never seen before (e.g., Gerson & Woodward, 2012). It is likely that social factors like infant-directed actions help infants to understand the intentional structure of actions outside of their own motor repertoire. For example, an adult may help an infant to understand the end-goal of stacking a tower of cups by highlighting each stacking action as well as the completion of the tower. The infant may understand the intentional goal of stacking all the cups, yet not have the refined motor skill to produce such an action. Differences in the extent to which parents modify their infant-directed actions could be associated with infants' general ability to understand and predict goal-directed actions. For example, exposure to motionese-rich social interactions could provide infants with a basis of intentional understanding that could be generalised to new actions. The current study investigated how various dimensions of parental motionese relate to 9- to 11-month-olds' ability to predict the outcome of a complex tool-use action that they have never seen before. At this age, infants are beginning to perform and understand tool-use actions but are not yet efficient as a group, with varying motor proficiency. Visual predictions of an incomplete action were measured using the eye-tracking paradigm previously described in chapter 3. Mothers' use of motionese during toy demonstrations were measured on both a global scale adapted from Brand et al. (2002) as well as a fine-grained scale adapted from Brand et al. (2007). We hypothesised that infants who were exposed to more motionese would make a higher proportion of accurate predictions in our eye-tracking paradigm than infants who were exposed to less motionese. We also explored whether infants' general motor proficiency might relate to differences in the way in which mothers exaggerated features of motionese, such as proximity (van Schaik et al., 2019).

Method

Participants

The current study utilised the same infant sample as that in chapter 3. 56 mothers participated in a two-visit study with their 9- to 11-month-old infants. The first visit involved an interaction session in which mothers demonstrated toys to their infants. Approximately one week later, (*mode* = 7 days; *mean* = 8.65 days; *range* 1 - 48 days), mothers and infants returned to complete an eye-tracking task followed by a motor task. Data from 13 infants were discarded due to insufficient eye-tracking data. This was because of fussiness (*N* = 1), failure to attend to familiarisation trials for an adequate proportion of time (*N* = 4), failure to encode the swap in toy locations (*N* = 1), or failure to make any type of prediction during test trials (*N* = 7). See the data processing section below for detailed descriptions of exclusion criteria within each phase of the eye-tracking study. In a between-subjects design, the remaining sample of 43 (22 female infants. Time 1: *age range* = 8.88 – 11.74 months, M_{age} = 9.77. Time 2: M_{age} = 10.05 months, *age range* = 9.1 – 11.8 months) infants were allocated to a comparison (*N* = 24, M_{age} at time 1 = 9.92, M_{age} at time 2 = 10.15) or control condition (*N* = 19, M_{age} at time 1 = 9.57, M_{age} at time 2 = 9.91) for the eye-tracking task. 4 infants failed to complete the motor task and thus motor data was analysed in relation to motionese data for a total of 39 infants. All infants who participated were full term and reported no developmental delays. Participants were recruited via a participant database and social media and came from a mid-sized city in the United Kingdom. All but three mothers completed a demographics questionnaire to obtain information including the infant's ethnicity, as well as the mother's age and highest level of education. Of the 40 mothers who completed this questionnaire, the majority reported their infants to be of White British ethnicity (*N* = 31), with the remainder being of another White ethnicity (*N* = 4), mixed Chinese and British (*N* = 3), mixed White and Asian (*N* = 1), and mixed Kurdish Turkish and British (*N* = 1). Mothers' ages ranged from 24 to 42 years, with the majority between 29 and 39 years (*N* = 35). The majority of mothers had obtained a Bachelor's degree (*N* = 19), with others holding an NVQ/Diploma (*N* = 1), A Levels/Further Education Qualification (*N* = 2), Master's Degree or equivalent (*N* = 13), or a MD/PhD or equivalent (*N* = 5).

Time 1 Procedure: Mother-infant Interaction Task

Setup and equipment

During the interaction task, mothers demonstrated four toys to their infants: a detachable caterpillar, stacking cups, shape sorter, and a ring stacker (see figure 8 for photos of each toy). These toys were challenging in that they afforded actions that 10-month-olds cannot yet perform, such as inserting a shape through the correct hole in a shape sorter. These toy demonstration tasks were included as part of a larger study that involved additional tasks that all infants underwent, such as a separation event and still face event in counterbalanced orders. Throughout the interaction session, mothers and infants were sitting across a table from one another. The infant sat in a highchair attached to the table and the mother sat in a chair approximately 75cm away from the infant. Mothers could reach across the table when demonstrating toys. An LED light was fixed to one side of the table using Velcro. A strip of black tape marked the centre of the table to aid proximity coding (see global coding and reliability section below). Sessions were recorded using three video cameras providing visual angles of the table, the infant, and the mother. Two of these cameras were placed in the room with the mother and infant (figure 9a and figure 9b), and the third (figure 9c) was operated by an experimenter on the other side of a two-way mirror. Each toy demonstration was timed for four minutes using a stopwatch operated by an experimenter in the observation room. The LED light was switched off to signal the beginning of a toy demonstration and switched on when the demonstration ended.

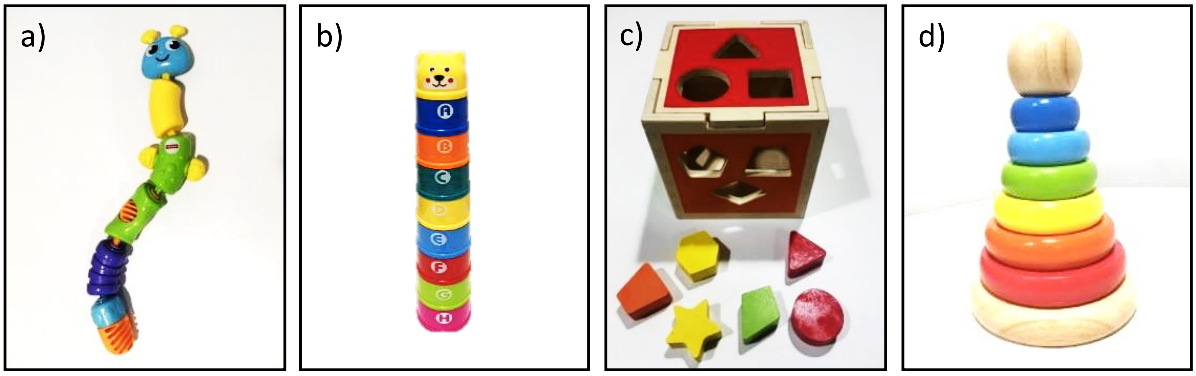

Figure 8. Toys demonstrated during the mother-infant interaction task. a) caterpillar toy, b) stacking cups, c) shape sorter, d) stacking rings.

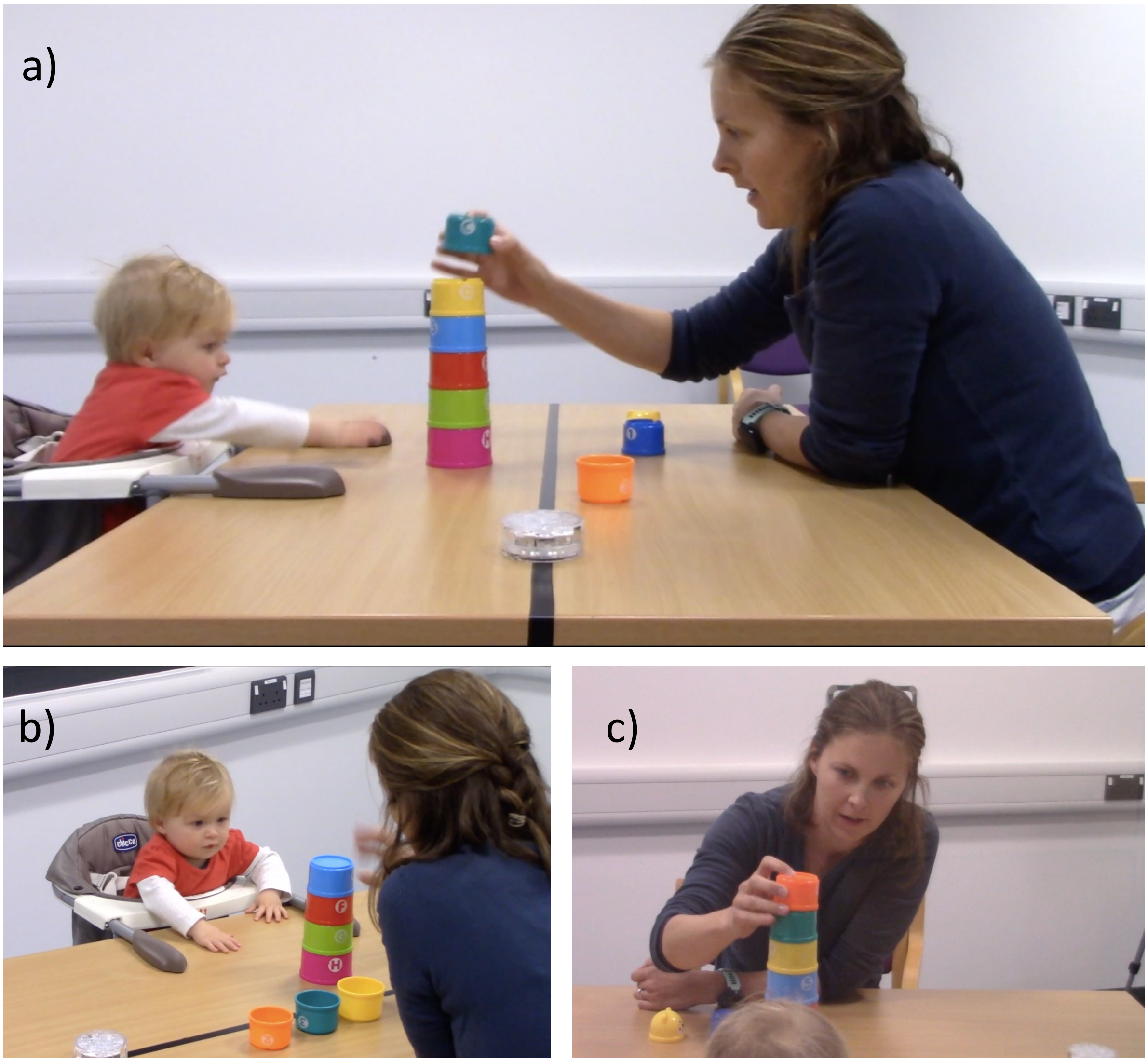

Figure 9. Camera angles. a) visual angle of the table with both mother and infant in view, b) visual angle of the infant, c) visual angle of the mother. An LED light was mounted on the table. This was switched off to mark the beginning and switched on to mark the end of toy demonstrations.

Toy demonstration tasks

Mothers were informed that the aim of this study was to investigate how toys are demonstrated to infants. The session began with a five-minute free-play session to help acclimatise the mother and infant to the experimental room. During free play, the mother and infant were provided with a variety of toys that were laid out on the table. After the free-play session, an experimenter cleared the toys away and introduced the first toy to be demonstrated to the infant. The mother was asked to demonstrate the toy to her infant and “show her/him how it works.” To ensure that the demonstration and play was as naturalistic as possible, the mother was informed that the infant can also play with the toy and to demonstrate and play together as they would at home. All four toys were demonstrated in separate four-minute sessions. Between demonstrations, the experimenter entered the room to exchange toys and repeat the demonstration instructions. During the first half of the play session, mothers and infants were provided with the caterpillar toy followed by the stacking cups in a counterbalanced order. In the second half, mothers and infants were provided with the stacking rings and shape sorter in a counterbalanced order. As part of a separate larger study beyond the scope of this thesis, infants were also provided with a pacifier during either the stacking rings or shape sorter demonstration (counterbalanced). We ensured that the presence of the pacifier was not associated with any of our variables of interest (see results section below). In addition, the larger study involved separation, still face, and a positive emotion event (e.g., peek-a-boo) that were interspersed within this procedure in counterbalanced orders. During the separation event, the mother and infant played together with a stuffed toy for two minutes. After two minutes, the LED light was switched on as the mother's cue to leave the room for two minutes. The still face event followed the same format, with the light as the mother's cue to look at the infant with a neutral expression for two minutes. During the positive emotion event, mothers and infants played with a stuffed toy. The light signalled the mother's cue to make their baby laugh using a technique of their choice (e.g., tickling, peek-a-boo, etc.).

Global coding and reliability

The recorded interaction sessions were coded on eight broad dimensions based upon Brand et al.'s (2002) coding scheme. For each of the four toy demonstrations, coders provided a global rating (0 – 4) on each dimension: *proximity, interactiveness, enthusiasm, range of motion, rate, repetitiveness, punctuation,* and *simplification.* In their instructions, coders were given descriptions of specific behaviours to include in their global ratings. See table 8 for our global coding scheme. For coding *proximity*, the mother and infant's space were defined by a strip of tape across the centre of the table. Higher ratings on each dimension contributed to a higher motionese score, except *rate* which was reverse scored due to slower movements expected within infant-directed actions (e.g., Brand et al., 2002; Rutherford & Przednowek, 2012). For each toy demonstration, coding began from the moment that the mother touched the toy and continued until the LED light signalled the end of the four-minute interaction. The six dimensions that were specific to the mothers' action on the object (i.e., proximity, range of motion, rate, repetitiveness, punctuation, and simplification) could not be meaningfully rated when the mother was not in contact with the toy (e.g., when the toy, or part of the toy such as a cup, was placed on the table or only the infant was in possession of the toy). However, enthusiasm toward the toy and interactiveness with the infant could still be coded during these periods. After coding a number of videos, it became apparent that many mothers demonstrated the toy at the beginning of each session, before passing the toy to their infant and engaging in play. This meant that mothers were mostly in contact with the toy at the beginning of each session. Because of this, we coded the dimensions *proximity, range of* *motion, rate,* and *punctuation* during the initial demonstration (i.e., until the mother had moved the toy into the infant's space and was no longer manipulating the toy), whereas the entire 4-minute demonstration was coded for *enthusiasm* and *interactiveness.* This also gave mothers control over the length of the initial demonstration as in Brand et al. (2007) and Rutherford and Przednowek (2012). For each participant, an average score was calculated for each of the eight dimensions across all four toys. If the visit ended early (i.e., not all toys were demonstrated), the average was calculated across three of the toys (*N* = 2). One mother and infants' data were removed for the stacking rings demonstration because of the pacifier causing distraction from the task, resulting in a lack of engagement with the task.

Table 8

Global motionese coding scheme.

Motionese Dimension	**Scale Points**	**Rated Period**	**Behavioural Description**
Proximity to partner	0 = demonstration always or almost always in mother's space 1 = demonstration more often in demonstrator's space 2 = demonstration between demonstrator's and infant's space or equal amounts in both 3 = demonstration more often in infant's space 4 = demonstration always or almost always in infant's space	Initial demonstration	
Interactiveness	0 = very low interaction 1 = low interaction 2 = neither low nor high interaction 3 = high interaction 4 = very high interaction	Entire interaction	Gaze checking (e.g., seeking eye contact), joint attention, object exchanges
Enthusiasm	0 = very low enthusiasm 1 = low enthusiasm 2 = neither low nor high enthusiasm 3 = high enthusiasm 4 = very high enthusiasm	Entire interaction	Amplified facial expressions and body movements
Range of motion	0 = very small, restricted movements 1 = small, restricted movements 2 = neither small nor broad movements 3 = broad, expansive movements 4 = very broad, expansive movements	Initial demonstration	
Rate	0 = very slow 1 = slow 2 = neither slow nor fast 3 = fast 4 = very fast	Initial demonstration	
Repetitiveness	0 = no repetitions 1 = few repetitions 2 = some repetitions 3 = many repetitions 4 = extremely repetitive	Entire interaction	Repeating individual actions (e.g., repeatedly stacking the same two cups is considered more repetitive than repeating the same action sequence, such as repeatedly stacking the tower of cups)
Punctuation	0 = gentle, continuous, very fluid actions 1 = somewhat continuous, fluid actions 2 = neither fluid nor punctuated actions 3 = somewhat abrupt, punctuated actions 4 = sharp, abrupt, very punctuated actions	Initial demonstration	Introducing more pauses and enacting sharper, more abrupt movements
Simplification	0 = complex combinations of many actions 1 = somewhat complex combinations of actions 2 = neither complex nor simple actions 3 = somewhat small, simple units of action 4 = small, simple units of action	Entire demonstration	Presenting shorter sequences and less complex combinations of actions

Coders were the first author of this study, as well as trained undergraduates and interns. To avoid any potential bias, all videos were coded without knowledge of infants' performance on the action prediction task. Videos were also coded with the sound muted to avoid any influence of speech upon action coding. To check for inter-rater agreement, 20% of participants were coded by two coders. Reliability was checked by comparing average scores on each dimension. To compare average scores between coders, we calculated the percentage of ratings that were within .5 of a point from one another. Percentage of agreement was high across all dimensions (>77%), except for *simplification* (22.22%). We reasoned that this may be due to the limitation of using a global scale to code this dimension, which is better captured using a fine-grained coding scale (Brand et al., 2007). Therefore, we excluded the dimension of simplification in main analyses of our global motionese data. A composite score for each participant was also calculated by averaging across motionese dimensions (excluding simplification).

Fine-grained coding and reliability

To better capture and quantify mothers' use of simple actions during toy demonstrations as well as to gain a rich measure of interactiveness, we also coded our video data on a fine-grained scale adapted from Brand et al. (2007). Using ELAN video coding software (Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands), we quantified various aspects of the mother-infant interaction for each toy demonstration. Interactiveness was operationalised by coding (a) the number of mother's eye-gaze bouts (gaze shifts from elsewhere to the face of the infant) per minute, (b) the percentage of the demonstration spent gazing at the infant, (c) the average length of each gaze bout, and (d) the number of exchanges per minute (i.e., each time the mother relinquished a toy from her own hand to the infant's hand). Simplification was quantified as (e) the number of different action types that the mother performed during each possession of a toy. We calculated the number of distinct action types that mothers performed during each possession of the toy, or part of the toy, and divided this by the number of times that the mother was in possession throughout the interaction. Because our toys were detachable, meaning both the infant and the mother could be in possession with a part of the toy at one time, distinct action types were quantified differently to the way that they were by Brand et al. (2007). That is, Brand and colleagues quantified the number of action types performed between exchanges of a single toy, thus making mother and infant turn-taking distinguishable. This meant that number of action types were divided by the number of mother's turns. As we were interested in different types of actions rather than repetition of actions, repeated action types were not coded within a single possession. A list of action types for each toy was generated after watching demonstrations of the toy from at least four mothers. Table 9 shows one example of potential action types that mothers may perform when demonstrating the stacking cups. We also coded the proportion of time that the mother and infant were in possession of the toy, or any part of it, as well as the proportion of time that both the mother and infant were in joint contact with the toy. See appendix for our fine-grained coding scheme.

Table 9

Potential action types that mothers may perform when demonstrating the stacking cups.

Action type	Definition
Stacks	Stacking a cup on top of another to make a tower of at least two cups.
Unstacks	Takes stacked cups apart. Knocking cups over does not count.
Nests cup(s) inside	Mother places smaller cup(s) inside a bigger cup. Rims of cups are facing upwards.
Nests cup(s) over	Mother places larger cup(s) over smaller cups, hiding the smaller one(s) underneath. Rims of cups are facing downwards.
Unnests	Takes nested cups apart (whether they are inside or over one another).
Pushes tower to infant	Mother pushes tower of two or more stacked cups towards the infant.
Animates	This can involve moving the cups around to gain the infant’s attention. E.g., moving cups around on the table, in the air, shaking, etc.
Peek-a-boo	E.g., hiding one cup inside another then revealing it, often using the bear face cup.
Taps cups together	Tapping two cups together.
Taps cup(s) with finger(s)	Mother taps the cup(s) using her finger(s) or whole hand.
Taps cup(s) on table	Mother taps the cup(s) against the surface of the table.
Taps infant with cup	Mother touches the infant using a cup. May be the infant’s hand, arm, face, etc.
Holds cups in place	This can involve steadying the tower or nest of cups, often when the infant is manipulating it.
Holds cup(s) up to show infant	Mother may hold cup(s) up high or in front of the infant’s face.
Puts cup(s) in front of infant	This can involve placing/pushing cup(s) on the table in front of the infant. This does not count when mother pushes the stacked tower of cups to infant.
Arranging	This can involve turning cups over, lining cups up, changing their position on the table etc.

Due to ending the experiment early because of infant fussiness, two mothers and infants did not complete the stacking rings demonstration. As reported in the previous section, one mother and infants’ data were removed from the stacking rings demonstration because of the pacifier causing distraction from the task, resulting in a lack of engagement with the task. This meant that, for three mothers, data on each variable was averaged across three toys rather than four. Due to difficulty seeing the eyes because of glasses, one mother’s gaze data was excluded across all four toy demonstrations. Another mother’s gaze data was removed from the shape sorter demonstration due to glare from glasses in this particular task.

Finally, one mother's gaze data was excluded from both the cup and ring stacking demonstrations due to failure to remain seated for a large portion of these tasks.

Coders were the first author of this study, as well as trained undergraduates, postgraduates, and interns. To avoid any potential bias, all videos were coded without knowledge of infants' performance on the action prediction task. Videos were also coded with the sound muted to avoid any influence of speech upon action coding. To check for inter-rater agreement, 26% of participants were double coded for the caterpillar demonstration, and 21% of participants were double coded for the stacking cups, stacking rings, and shape sorter demonstrations. For the caterpillar demonstration, intraclass correlation coefficients ranged from .71 (for number of gaze bouts per minute) to .99 (for infant possession and joint contact) indicating moderate to excellent reliability. For the stacking cups demonstration, intraclass correlation coefficients ranged from .67 (for number of action types per possession) to 1 (for mother possession) indicating moderate to excellent reliability. For the stacking rings demonstration, intraclass correlation coefficients ranged from 0.79 (for number of action types per possession) to 0.99 (for both mother and infant possession) indicating good to excellent reliability. For the shape sorter demonstration, we achieved an intraclass correlation of .41 for number of gaze bouts per minute. However, for proportion of time spent gazing at the infant and average length of gaze bout we achieved intraclass coefficients of .78 and .60 respectively, indicating good reliability for our gaze data. For the remaining variables, intraclass coefficients ranged from .66 (for number of exchanges per minute) to 1 (for infant possession) indicating moderate to excellent reliability. For each participant, scores on each variable were averaged across all four toys and z-transformed scores for each variable were calculated. We also created a composite score averaged across all eight z-transformed variables.

Time 2 Procedure: Eye-tracking and Motor Tasks

Setup and equipment

As detailed in the previous chapter, infants returned approximately one week later to complete the eye-tracking and motor tasks. Eye-tracking data were collected using a Tobii Pro X3-120 eye-tracker, with a 120Hz sampling rate. The Tobii default fixation filter was used to define eye fixation: a stable gaze (within 0.75 visual degrees) for a minimum of 200ms. The eye-tracker was attached to a 23" (58.4cm) monitor (1920 x 1080p). The monitor was mounted onto the wall via a moveable arm that was adjusted for optimal height and distance of the infant sitting on the parent's lap (approximately 60cm). A webcam was attached to the monitor so that the mother and infant could be viewed by the experimenter from another room. Tobii Studio (Tobii Technology, Stockholm, Sweden) eye-tracking software was used for calibration and to record eye-gaze data whilst E-prime (Psychology Software Tools, Pittsburgh, PA) was used to present stimuli. Prior to the experiment, all infants underwent a five-point calibration to Tobii Studio's pre-set locations. We required data to be valid for five calibration points for both eyes before beginning the experiment. Throughout the eye-tracking task, parents were asked to hold the infant securely and orient them toward the screen but not to influence the infant's looking to objects on the screen. As detailed in the previous chapter, infants first underwent an observational training phase in which they were assigned to experimental or control conditions. However, this phase of the eye-tracking task is not relevant for the focus of the current study and thus exclusion criteria are only applicable for the familiarisation phase, freeze screen, and test trials described below. This meant that we were able to include more infants than in previous eye-tracking analyses in chapter 3. Nevertheless, we still checked for effects of observational training condition in our analyses (see results section below).

Eye-tracking familiarisation phase

Infants underwent an adaptation of Cannon and Woodward's (2012) paradigm in which they were familiarised with videos of a claw-reach action involving an agent using a hand-operated claw to reach, grasp, and pull a goal-object nearer. A teddy bear and a tennis ball were presented in the top corners of the scene on opposite sides of a wall. Below the bear and ball, two claws were also presented on either side of the wall, which could be used to act upon the corresponding toy. Infants viewed three videos in which the same goal-object (bear or ball) was consistently acted upon with the claw (see figure 10a). The agent's hand entered the bottom centre of the screen and used the left or right claw (counterbalanced) to retrieve the toy (bear or ball) on the corresponding side. The toy that was grasped using the claw was also counterbalanced. These videos were approximately five seconds long with some natural variation. Within this phase, as well as the later training phase (see below), a fixation cross was displayed preceding each trial and trials were progressed manually by the experimenter. Upon an alternate key press, the experimenter could also choose to present an attention-grabbing animation to re-gain the infant's attention before proceeding with the trial.

Eye-tracking freeze screen

In-between familiarisation and test trials, all infants viewed a single freeze screen in which the bear and ball had swapped positions, but the claws remained in the same position. This image was presented for five seconds to allow time for infants to encode the swap by fixating on at least one of the two toys (see data processing section below).

Eye-tracking test trials

With the bear and ball in their new positions, infants then viewed four test trials in which the agent's hand moved onto the centre of the screen and froze, making an incomplete action (see figure 10b). These videos were approximately five seconds long with some natural variation. After completing this task, infants underwent a second eye-tracking task as part of a separate project that is beyond the scope of this thesis.

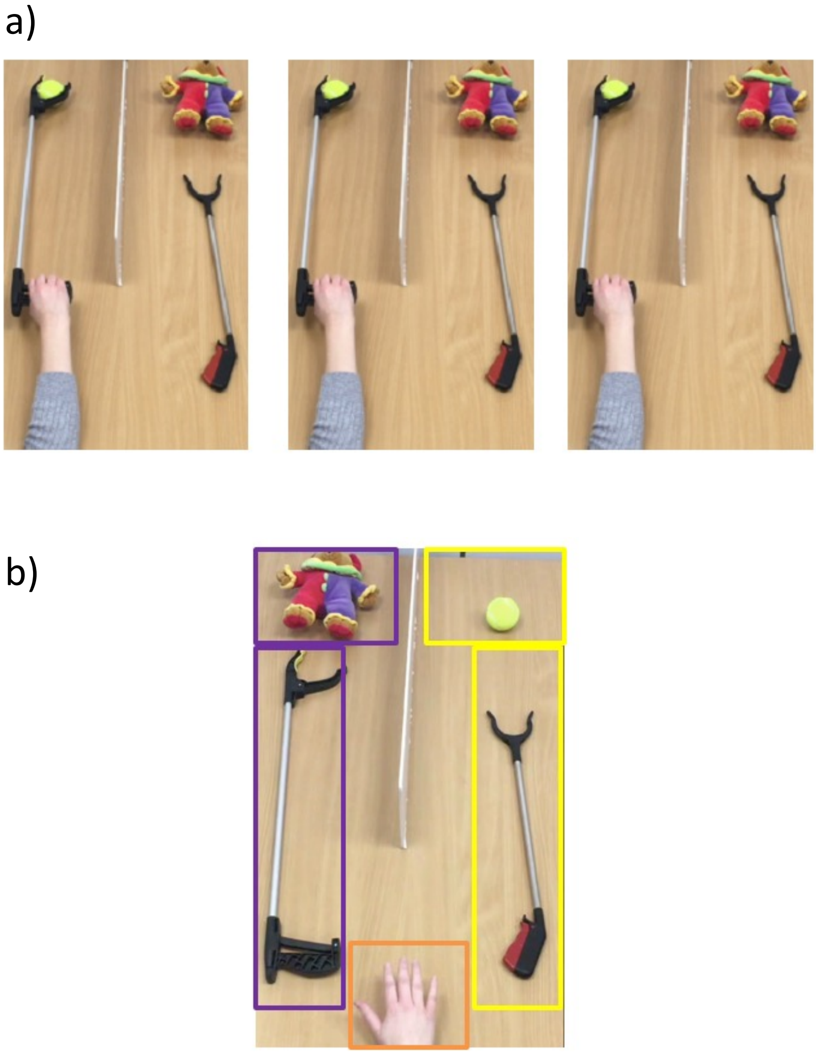

Figure 10. Schematic of eye-tracking paradigm. Still frames depict videos presented. a) Familiarisation: The same goal object (ball or bear) is consistently reached for and drawn close using corresponding claw. b) Test trials: Incomplete reaching action with defined AOIs. In this example, the orange AOI captured fixations to the hand. Purple AOIs captured fixations to the left claw and left toy (inaccurate goal). Yellow AOIs captured fixations to the right claw and right toy (accurate goal). These AOIs were not visible to infants when viewing the videos.

Motor trials

After completing the eye-tracking task, infants completed a motor task, identical to that described in chapter 3, which was inspired by previous work (Gerson et al., 2015b; Sommerville et al., 2008; Sommerville & Woodward, 2005). This consisted of pre-training, training, and post-training trials in which infants had the chance to perform a cloth-reach and a magnet-reach (in a counterbalanced order) to retrieve a magnetic toy train. Only infants' existing motor ability to perform these actions during three pre-training trials was relevant for the current study, as improvement from pre- to post-training trials was assessed in the previous chapter. Infants sat on a parent's lap facing the table. An experimenter sat next to the infant and the task was recorded by a second experimenter for later coding. In each trial, infants were given the opportunity to perform the tool-reach to retrieve the toy train without any guidance from the experimenter (see figure 11). The experimenter set up the task in front of the infant and then looked down at the table. If the infant was distracted or uninterested in the task, the experimenter tapped on the train to draw the infant's attention to the problem but did not provide more specific cues to prompt the infant's actions. The trial ended when the infant obtained the toy or when 30s had elapsed. When the infant obtained the toy, the experimenter praised them by saying "well done" enthusiastically. If the infant did not obtain the toy, the experimenter did not say anything to the infant and proceeded with the next trial.

Figure 11. Experimental setup for motor task. Infants performed tool-reach actions (i.e., a cloth- and magnet-reach) without any guidance from the experimenter. Note that only the cloth-reach example is shown in this figure.

Data Processing

Eye-tracking data

As described in more detail in the previous chapter, infants who did not attend to at least one video of each action type during familiarisation trials for a minimum of 40% of the time were excluded from all subsequent analyses (*N* = 3). 52 infants' data were remaining for assessment of fixation data during the freeze screen.

Using the same criteria as in the previous empirical chapters, eye fixations to both the bear and ball were measured to infer whether infants encoded the swap in positions during the freeze screen. AOIs (matched in size, 170 x 210p) were defined over the bear and ball to ensure that infants fixated on at least one of these objects to encode the swap in locations. If infants did not fixate on either of these objects, we concluded that they did not encode the swap during this phase. If an infant did not encode the swap during the freeze screen, we used the same method to investigate whether the swap was encoded during initial test trials. This is because infants may not have noticed the swap in toy positions during the freeze screen but subsequently had the opportunity to notice the swap during test trials. If the swap was encoded during test trials, subsequent test trials were then coded for predictions. No trials were coded for predictions before the swap had been encoded, as fixations before this point were not likely to be anticipatory of the agent's intention. If we could not confirm that an infant encoded the swap after these investigations, their data were excluded from analyses as fixations recorded during test trials were likely to be arbitrary rather than predictive (*N* = 1).

In test trials, AOIs for the bear and ball (matched in size, 150 x 210p), as well as the left and right claw (matched in size, 90 x 650p) were defined along with an AOI covering the hand (120 x 240p, see figure 10b for a representation of AOIs). Fixations were coded from the exact moment the hand entered the screen on each test trial (after approximately one second, with natural variation between videos). This was precisely defined for each trial. After fixating on the hand, we coded the first object that infants fixated on (hand-to-object prediction). As in previous visual prediction paradigms (Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018), fixations to a toy or claw without first fixating on the hand were not considered predictive anticipations of the familiarised action. When infants made hand-to-object fixations on the familiarised goal object or its newly corresponding claw, this was considered an *accurate prediction.* When infants made hand-to-object fixations on the familiarised claw, or the non-familiarised object (in the familiarised location) this was considered an *inaccurate prediction.* Fixations to claws were included as predictive looks because infants were familiarised with the action trajectory of the hand using the claw to grasp the toy. If infants do not understand the agent's two-step action of using the tool to reach the goal-object, we expect them to look from the hand to the claw that the agent manipulated during familiarisation (i.e., an inaccurate prediction in which infants are only predicting a one-step action). In contrast, infants who understand the agent's two-step action of using the claw to achieve the preferred toy should look to the claw that can reach for the familiarised toy (i.e., an accurate prediction). If infants did not fixate on the hand or did not fixate on another AOI following fixation on the hand, these trials were considered as *no anticipation* and were excluded from analyses. If infants did not make predictive anticipations on any test trial their data were excluded from analyses (*N* = 7). Saccades through AOIs were not coded as predictive looks. All looking time and fixation data were output by E-Prime (Psychology Software Tools, Pittsburgh, PA) and processed using Matlab (MATLAB, 2019). Timestamps of fixation to each AOI determined the order of fixations and thus allowed us to code predictive looks. Each infant's proportion of predictive fixations was calculated separately for accurate and inaccurate predictions, with number of predictive trials as the denominator. Table 10 shows the proportion of infants in each condition who made accurate predictions, inaccurate predictions, and no anticipations within each test trial. Accurate predictions increased from first to last test trial, particularly in trials 3 and 4. In contrast, inaccurate predictions varied across trials, but most were made within the first two test trials. Table 11 represents the various categories of performance patterns shown by individual infants across test trials. That is, the extent to which infants' prediction accuracy improved, declined, or varied from the first to last test trial in which accurate or inaccurate predictions were made. Infants are also categorised into those who made accurate predictions only or inaccurate predictions only. Average numbers of predictive trials indicate the extent to which infants made predictions (i.e., accurate or inaccurate) as opposed to no anticipations. Most infants within the comparison condition made inaccurate predictions only (38%). Within the control condition, most infants made accurate predictions only (37%) or inaccurate predictions only (42%).

Table 10

Proportion of infants in each condition who made accurate predictions, inaccurate predictions, and no anticipations within each test trial.

	Trial 1	**Trial 2**	**Trial 3**	**Trial 4**
Comparison condition (*N* = 24)				
Accurate	17%	17%	38%	29%
Inaccurate	46%	50%	21%	25%
No anticipation	38%	33%	42%	46%
Control condition (*N* = 19)				
Accurate	32%	21%	26%	21%
Inaccurate	32%	32%	16%	42%
No anticipation	37%	47%	58%	37%

Table 11

Patterns of prediction accuracy from first to last predictive test trial, as well as the average number of test trials in which infants made a prediction (i.e., accurate or inaccurate) as opposed to no anticipation.

Accuracy across predictive test trials	**%**	**Average number of predictive test trials**
Comparison condition (*N* = 24)		
Accuracy improved (inaccurate to accurate)	21%	3.00
Accuracy declined (accurate to inaccurate)	13%	2.33
Accuracy varied	17%	3.25
Accurate predictions only	13%	2.33
Inaccurate predictions only	38%	1.78
Control condition (*N* = 19)		
Accuracy improved (inaccurate to accurate)	11%	3.50
Accuracy declined (accurate to inaccurate)	5%	2.00
Accuracy varied	5%	3.00
Accurate predictions only	37%	1.71
Inaccurate predictions only	42%	2.25

Motor data

Across each of the three trials, infants' cloth- and magnet-reach actions were coded for the extent to which infants engaged in well-structured planful actions. An action was coded as planful if the infant maintained visual attention to the toy while using the tool to draw it near in one continuous movement, then obtained or touched the toy within three seconds of completion of the pull (Gerson et al., 2015b; Sommerville & Woodward, 2005). For each tool, every infant's motor efficiency was operationalised by calculating the proportion of planful actions made across the three trials. To check for inter-rater agreement, more than 30% of infants' motor data was double coded by a second coder. There was moderate agreement between coders, $\kappa = .77$, $p < .001$. No infants made planful magnet-reach actions, so only the cloth-reach action was included in analyses. Four infants did not complete the motor task, leaving 39 infants' motor data for inclusion in statistical analyses.

Results

An initial Spearman's correlation assessed whether there was any relation between age at second visit (i.e., when infants completed the eye-tracking and motor tasks) and infants' proportion of accurate predictions during test trials. This revealed no significant correlation between variables, $r(41) = .14$, $p = .38$. As reported in the previous chapter, a Spearman's correlation revealed that infants' age at their second visit was positively correlated with motor ability, *r*(37) = .40, *p* = .01. Therefore, we controlled for age in all subsequent analyses that included infants' motor ability at pre-training as an independent variable.

A preliminary analysis was run to test whether sex, condition (control vs comparison), or motor ability (high vs low) had any significant effect on proportion of accurate predictions in the slightly larger sample of infants in the current study. In a GLZM, proportion of accurate predictions was entered as the ordinal dependent variable. Sex, condition, and motor ability were added as predictor variables. Age at session two was added as a covariate to control for the relation between age and motor ability. This revealed no significant main effects of sex, Wald χ^2 (1) = .12, *p* = .74, β = -.04, condition, Wald χ^2 (1) = .11, *p* = .74, β = -.06, or motor ability, Wald χ^2 (1) = .01, *p* = .91, β = .02, upon proportion of accurate predictions. There was also no significant interaction between condition and motor ability, Wald χ^2 (1) = 1.58, *p* = .21, β = .31.

As we did not obtain adequate inter-rater agreement on the simplification dimension of our global motionese coding (i.e., 22.22%), we decided to explore whether our global rating of simplification correlated with the fine-grained measure of simplification (i.e., number of action types per possession). We also explored whether any other global ratings of motionese correlated with our fine-grained measures, particularly relations between the interactiveness dimension of our global scale and the fine-grained quantifications of interactiveness (i.e., number of gaze bouts per minute, gaze percentage, average length of gaze bout, and number of exchanges per minute). See table 12 for Pearson correlation matrix. This revealed no significant correlation between global simplification and number of action types per possession. However, there was a significant negative correlation between global simplification and percentage of infant possession. There were also no significant correlations between global interactiveness and any fine-grained measures of interactiveness. However, there was a negative correlation between global interactiveness and infant possession. That is, the more time infants spent in possession of toys, the less interactive mothers were with their infants. There was also a positive correlation between global interactiveness and number of action types per possession that was marginally significant. As well as this, there was a positive correlation between global enthusiasm and average length of gaze bout. There was also a positive correlation between global repetitiveness and number of gaze bouts per minute that was marginally significant. Note that caution should be taken in the interpretation of these findings given the large number of tests. We decided not to control for multiple comparisons (e.g., Bonferroni correction) as this would lead to reduced power and thus type II errors.

Table 12

Pearson correlation matrix for global and fine-grained motionese variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. proximity (G)	-																	
2. interactiveness (G)	.11	-																
3. enthusiasm (G)	.17	.20	-															
4. range of motion (G)	-.08	.15	.41**	-														
5. repetitiveness (G)	.14	.12	.41**	-.00	-													
6. punctuation (G)	.07	.43**	.19	.49**	.05	-												
7. simplification (G)	-.05	-.01	.36*	.04	.39**	.16	-											
8. rate (G)	.32*	.25	.26	.44**	.40**	.42**	.10	-										
9. composite (G)	.35*	.52**	.67**	.60**	.54**	.65**	.27	.76**	-									
10. # gaze bouts per minute (FG)	.05	.08	-.12	-.13	.29†	.04	-.12	.05	.07	-								
11. gaze % (FG)	.02	.08	.22	.02	.04	.11	-.07	.02	.14	.28†	-							
12. average length gaze bout (FG)	.03	-.06	.33*	.18	-.15	.09	.03	.02	.13	-.60**	.52**	-						
13. # exchanges per minute (FG)	.02	.24	-.07	-.09	.13	.13	-.20	-.14	.04	.08	.40**	.15	-					
14. # action types per possession (FG)	.08	.27†	.13	-.20	.17	.07	.13	.10	.15	.08	.29	.08	.31*	-				
15. mother possession % (FG)	.11	-.11	-.09	-.14	.02	.08	-.04	.11	-.00	-.04	.30	.18	.29	.64**	-			
16. infant possession % (FG)	.10	-.35*	.02	-.17	.13	-.24	-.36*	-.15	-.16	.12	.01	-.01	.14	.21	-.11	-		
17. joint contact % (FG)	.06	-.13	.02	-.05	.09	.07	-.02	-.02	.01	-.04	-.07	-.07	.01	.23	.52**	.09	-	
18. composite (FG)	.13	.01	.12	-.14	.20	.10	-.14	.01	.10	.22	.70**	.32*	.61**	.62**	.72**	.26	.44**	-

G = global motionese variable, FG = fine-grained motionese variable

$^{\dagger}p < .08$, $^{*}p < .05$, $^{**}p < .01$

Global motionese

Prior to main analyses of our global motionese data, we confirmed that the seven dimensions of motionese (excluding simplification) cohere as a unitary phenomenon. An inter-item reliability analysis revealed a Chronbach's alpha of .72. This means that modifications associated with infant-directed action were intercorrelated. In other words, mothers who showed exaggeration on one dimension also showed exaggeration on other dimensions (also see table 12 for correlations between global variables). A Pearson correlation ensured that infants' age at first visit (i.e., when mothers and infants completed the toy demonstration task) was not correlated with mothers' engagement in any of the seven dimensions of motionese or the global composite score (*ps* > .43). Next, we used a MANOVA to ensure that there was no effect of infants' sex on motionese dimensions or the global composite score, $F(9, 33) = .93$, $p = .51$, Wilk's $\Lambda = .80$, partial $\eta^2 = .20$.

As detailed earlier in the chapter, infants were provided with a pacifier during either the stacking rings or shape sorter demonstrations (counterbalanced) as part of a larger study that is not of interest in the current study. We investigated whether the presence of a pacifier had any effect on mothers' engagement in motionese by calculating average motionese scores for only the stacking rings and shape sorter demonstrations. T-tests ensured that the presence of a pacifier versus no pacifier was not associated with mothers' engagement in motionese during the stacking rings demonstration, $t(38) = 1.22$, $p = .46$, or the shape sorter demonstration, $t(41) = .71$, $p = .60$.

Next, we checked for any effects of each motionese dimension and the observational training condition (control vs comparison) that infants underwent as part of the previous study (see chapter 3) upon infants' proportion of accurate predictions. To do this, we used separate GLZMs for each dimension as a continuous predictor variable, condition as a categorical predictor, and proportion of accurate predictions as the ordinal dependent variable. This revealed that repetitiveness had a marginally significant effect upon proportion of accurate predictions, Wald χ^2 (1) = 2.97, $p = .085$, $\beta = .17$. However, this trend was not significant in a follow-up Spearman's correlation analysis, $t(41) = .24$, $p = .30$. No other models revealed significant effects of, or interactions between, motionese dimensions and observational training condition (*ps* > .15).

Based on previous literature (Fukuyama et al., 2015; van Schaik et al., 2019) we explored whether mothers' engagement in motionese may differ dependent on their infants' motor ability. In a MANCOVA, all seven motionese dimensions as well as the global composite score were entered as dependent variables. Infants' motor ability (high vs low) was entered as the predictor variable. Infants' age at the second visit was added as a covariate to control for the relation between age and motor ability. This revealed no significant effect of infants' motor ability on mothers' engagement in motionese, *F*(8, 29) = .84, *p* = .58, Wilks' Λ = .81, partial η^2 = .19.

The main analysis of our global motionese data examined whether infants who were exposed to a higher level of motionese made a higher proportion of accurate predictions. A Spearman's correlation revealed no significant correlations between proportion of accurate predictions and the global composite score or any individual dimensions of motionese (*ps* > .28).

In line with the previous empirical chapters, we speculated that after making an initial prediction (accurate or inaccurate), infants may not continue to make meaningful predictions on subsequent test trials. This is in line with Cannon and Woodward's (2012) paradigm that used a single test trial to assess accurate visual prediction of an incomplete action. Separate binomial regressions were run for each of the seven dimensions of motionese, as well as for the global composite score. First prediction type (accurate vs inaccurate) was entered as the binary dependent variable in each. Proximity was not a significant predictor of first prediction type, Wald χ^2 (1) = .05, *p* = .82, β = -.16, *SE* = .71, and neither was interactiveness, Wald χ^2 (1) = .52, *p* = .47, β = .40, *SE* = .55, enthusiasm, Wald χ^2 (1) = .01, *p* = .92, β = -.04, *SE* = .39, range of motion, Wald χ^2 (1) = .001, *p* = .97, β = .02, *SE* = .51, repetitiveness, Wald χ^2 (1) = 2.60, *p* = .11, β = .81, *SE* = .50, punctuation, Wald χ^2 (1) = 1.36, *p* = .24, β = .52, *SE* = .45, simplification, Wald χ^2 (1) = .53, *p* = .47, β = .42, *SE* = .57, nor rate, Wald χ^2 (1) = .38, *p* = .54, β = .24, *SE* = .39. Global composite score was also not a significant predictor of first prediction type, Wald χ^2 (1) = .97, *p* = .33, β = .82, *SE* = .83.

Fine-grained data

Prior to main analyses of our fine-grained motionese data, a preliminary Pearson's correlation assessed whether infant's age at first visit was correlated with any fine-grained motionese variables (i.e., gaze bouts per minute, gaze percentage, average length of gaze bout, exchanges per minute, number of action types per possession, percentage of mother possession, percentage of infant possession, and percentage of joint contact) or the fine-grained composite score. There were no significant correlations between infants' age at their first visit and these variables (*ps* > .11). As previously reported in table 12, other interesting correlations between fine-grained variables reached significance, including positive correlations between number of exchanges per minute and number of action types per possession, percentage of mother possession and number of action types per possession, as well as percentage of joint contact and percentage of mother possession.

Next, we used a MANOVA to ensure that there was no effect of infants' sex on all fine-grained motionese measures as well as the composite score, *F*(8, 33) = 1.11, *p* = .38, Wilk's Λ = .79, partial η^2 = .21. Two separate MANOVAs also ensured that the presence of a pacifier versus no pacifier had no effect on fine-grained motionese variables during the ring stacking demonstration, *F*(8, 16) = .67, *p* = .71, Wilk's Λ = .75, partial η^2 = .25, or the shape sorter demonstration, *F*(8, 13) = .96, *p* = .51, Wilk's Λ = .63, partial η^2 = .37.

Next, we checked for any effects of each fine-grained motionese variable and the observational training condition (control vs comparison) that infants underwent as part of the previous study (see chapter 3) upon infants' proportion of accurate predictions. To do this, we used separate GLZMs for each motionese variable as a continuous predictor, condition as a categorical predictor, and proportion of accurate predictions as the ordinal dependent variable. This revealed a significant effect of number of exchanges per minute upon proportion of accurate predictions, Wald χ^2 (1) = 6.50, *p* = .01, β = .16. No other models revealed significant effects of, or interactions between, motionese variables and observational training condition (*ps* > .11).

Next, we explored whether fine-grained motionese may differ dependent on infants' motor ability. We used a MANCOVA with infants' motor ability (high vs low) as the independent variable. All fine-grained motionese measures, as well as the composite score, were entered as dependent variables. Infants' age at the second visit was added as a covariate to control for the relation between age and motor ability. This revealed no significant differences in mothers' engagement in motionese dependent on infants' motor ability, *F*(8, 28) = .83, *p* = .58, Wilk's Λ = .81, partial η^2 = .19.

Our main analysis investigated our hypothesis that mothers' engagement in motionese would facilitate infants' accurate predictions of a novel tool-use action. A Spearman's correlation revealed a significant relation between number of exchanges per minute and proportion of accurate predictions, *r*(41) = .33, *p* = .03. That is, infants whose mothers passed toys to them more frequently made the highest proportion of accurate predictions in the eye-tracking task (see figure 12). This is an interesting finding, as object exchanges (i.e., when a mother relinquishes a toy from her own hand to the infant's hand) were somewhat rare, with average number of exchanges per minute across all four toys ranging from a minimum of 0.6 to a maximum of 1.06 (*M* = 0.41). No other correlations between infants' proportion of accurate predictions and fine-grained motionese variables reached significance (*ps* > .18).

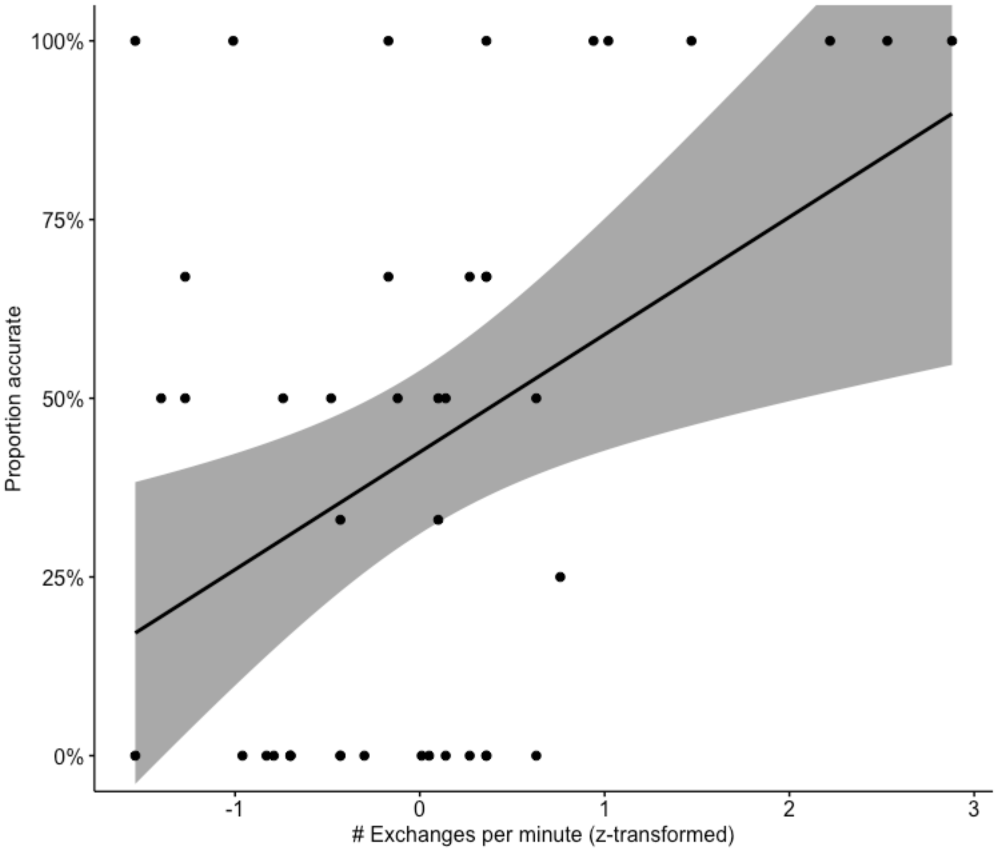

Figure 12. Scatterplot to show number of exchanges per minute by proportion of accurate predictions (with 95% confidence interval). *p* = .03.

In line with the analyses of our global motionese data as well as previous empirical chapters, we also investigated whether there was any effect of mothers' engagement in fine-grained motionese upon infants' first predictions (i.e., the first trial in which each infant made an accurate or inaccurate visual prediction). Separate binomial regressions were run for each fine-grained motionese variable, as well as for the fine-grained composite score. First prediction type (accurate vs inaccurate) was entered as the binary dependent variable in each. Number of gaze bouts per minute was not a significant predictor of first prediction type, Wald χ^2 (1) = .33, *p* = .57, β = .18, *SE* = .32, and neither was gaze percentage, Wald χ^2 (1) = .4, *p* = .53, β = .20, *SE* = .32, average length of gaze bout, Wald χ^2 (1) = .01, *p* = .93, β = .03, *SE* = .32, number of action types per possession, Wald χ^2 (1) = .07, *p* = .79, β = -.09, *SE* = .32, percentage of mother possession, Wald χ^2 (1) = .11, *p* = .75, β = -.10, *SE* = .32, nor percentage of infant possession, Wald χ^2 (1) = .01, *p* = .91, β = .04, *SE* = .32. However, number of exchanges per minute was positively associated with number of accurate first predictions, Wald χ^2 (1) = 4.22, *p* = .04, β = .78, *SE* = .38 (see figure 13). There was also a marginally significant negative association between joint contact and number accurate first predictions, Wald χ^2 (1) = 3.76, *p* = .053, β = -.75, *SE* = .39 (see figure 14). Fine-grained composite score was not a significant predictor of first prediction type, Wald χ^2 (1) = .09, *p* = .77, β = .19, *SE* = .65.

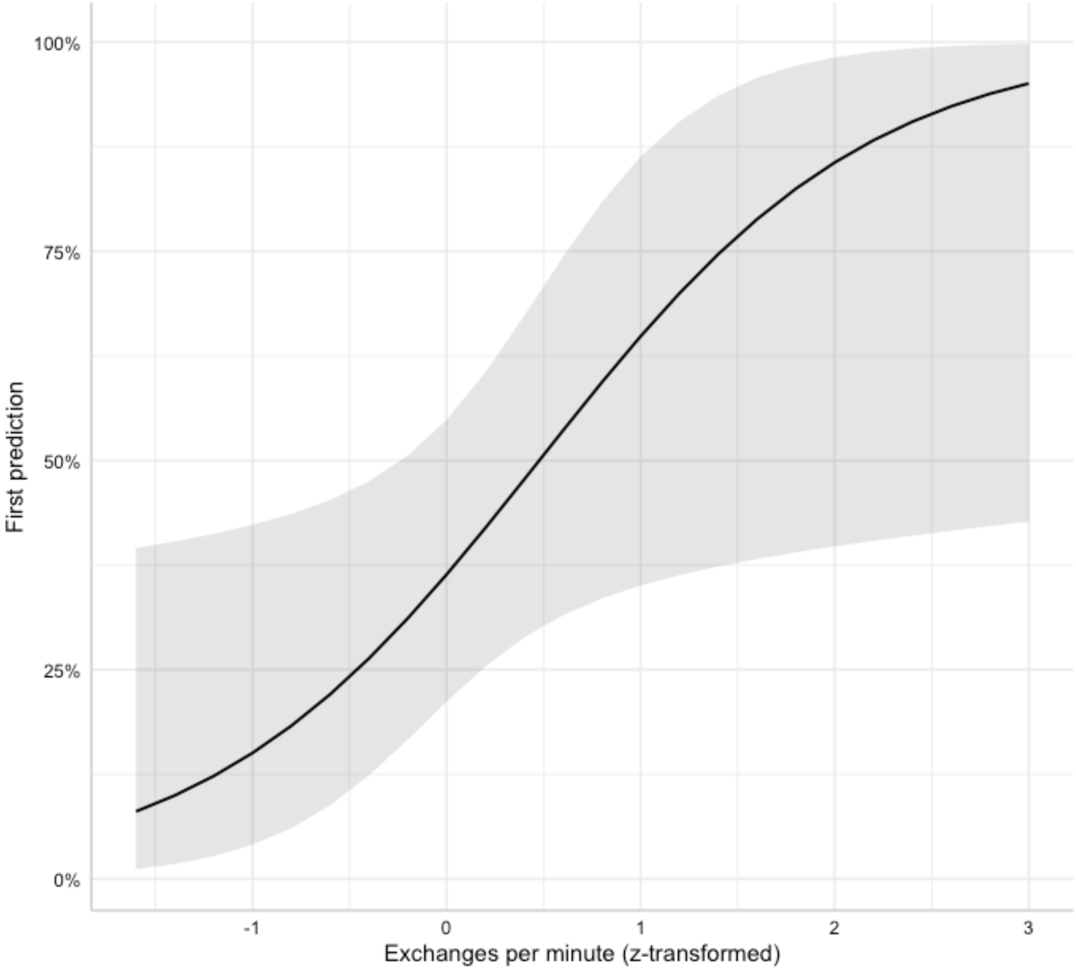

Figure 13. The predicted probability that an infant would make an accurate first prediction based on number of exchanges per minute during toy demonstrations (with 95% confidence interval). *p* = .04.

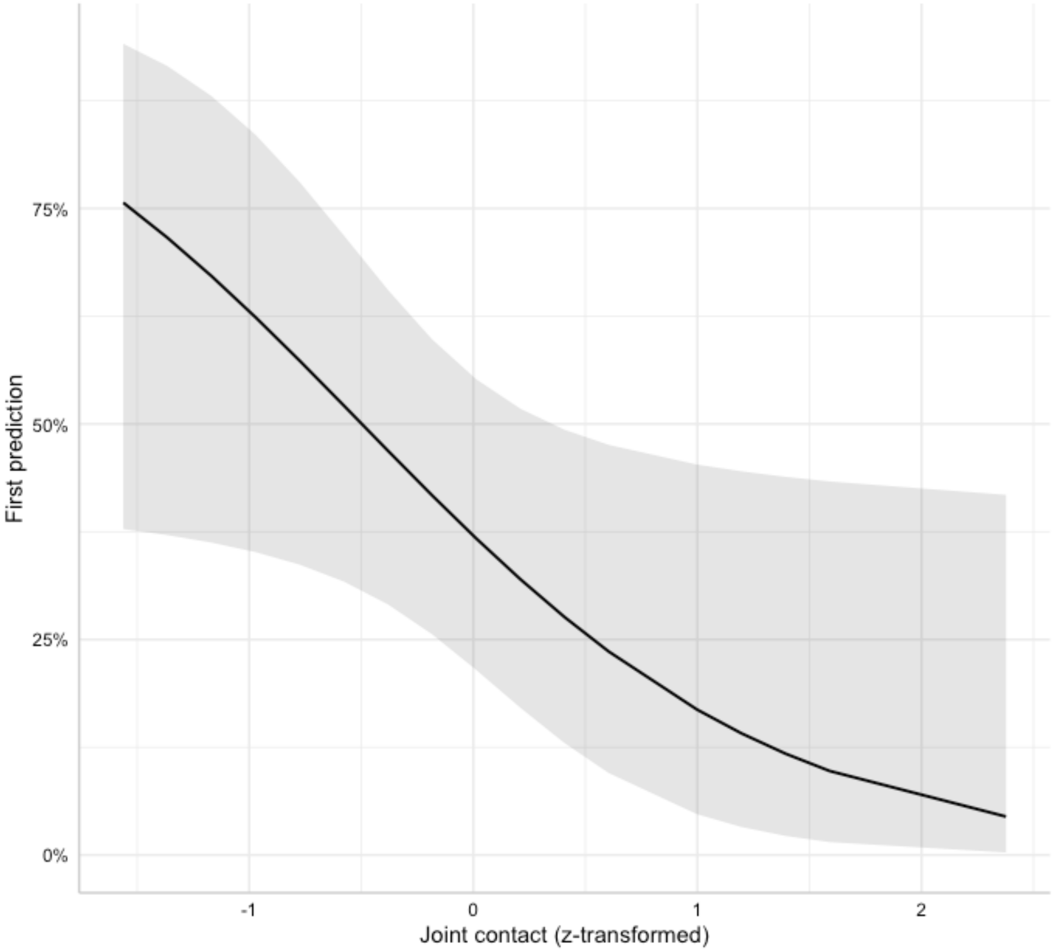

Figure 14. The predicted probability that an infant would make an accurate first prediction based on joint contact during toy demonstrations (with 95% confidence interval). *p* = .053.

Additional analyses of simplification

After quantifying our fine-grained measure of simplification by coding the number of action types per possession, we speculated that this may not be a valid measure of mothers' use of simple actions when demonstrating toys to their infants. As the average number of mother possessions (i.e., the number of times the mother was in contact with the toy or any part of it) across all four toys varied greatly between participants (*min* = 14.75, *max* = 43.75, *M* = 27.55, *SD* = 6.54), our measure of action types per possession captured repetition of action types to some extent. That is, mothers who were in possession of the toy a higher number of times also performed a higher number of action types. Number of action types per possession also positively correlated with number of exchanges per minute (i.e., a fine-grained measure of interactiveness) and with our global measure of interactiveness, although this positive correlation was marginally significant (see table 12 for correlation coefficients). Therefore, action types per possession also seemed to capture how engaged mothers were with the task as well as how interactive they were when demonstrating toys to their infant. As this was not our intention, we decided to re-evaluate how we calculated simplification. Rather than calculating number of action types per possession, we decided to quantify simplification by calculating number of action types per *minute* of possession. We did this by dividing the total number of action types by the total time that each mother was in possession of the toy. This was calculated separately for each toy, then averaged across all four toys. Raw scores were then z-transformed and we ran similar analyses to that which previously included action types per possession as a variable.

First, we used a Pearson correlation to assess any potential relation between our global measure of simplification and the fine-grained quantification of action types per minute. We also explored whether any other global motionese dimension correlated with action types per minute and ensured that infants' age at their first visit did not correlate with action types per minute. There was no significant correlation between action types per minute and global simplification, $r(41) = .21$, $p = .18$, or any other dimension of global motionese ($ps > .23$). There was also no significant correlation between age and action types per minute, $r(41) = .12$, $p = .46$.

Next, we checked for any effects of the observational training condition (control vs comparison) that infants underwent as part of the previous study (see chapter 3), as well as any interactions between training condition and action types per minute upon proportion of accurate predictions. Proportion of accurate predictions was entered as the ordinal dependent variable in a GLZM. Condition was added as a categorical predictor and action types per minute was added as a continuous predictor. This revealed no significant effect of condition, Wald χ^2 (1) = .85, $p = .36$, $\beta = .11$, or action types per minute, Wald χ^2 (1) = .00, $p = .99$, $\beta = -.001$, upon proportion of accurate predictions. There was also no significant interaction between these two variables, Wald χ^2 (1) = .1, $p = .76$, $\beta = -.04$.

Next, we explored whether the number of action types that mothers performed may differ dependent on infants' motor ability. In an ANCOVA, action types per minute was entered as the dependent variable and motor ability (high vs low) was entered as the independent variable. Infants' age at the second visit was added as a covariate to control for the relation between age and motor ability. This revealed no significant differences in number of action types per minute dependent on infants' motor ability, $F(1, 36) = .67$, $p = .42$, partial $\eta^2 = .02$.

To replicate our main analyses, a Spearman's correlation revealed no significant correlation between action types per minute and proportion of accurate predictions, $r(41) = -.03$, $p = .84$. In a binomial logistic regression, first prediction type (accurate vs inaccurate) was entered as the binary dependent variable and action types per minute was entered as the continuous predictor. This was not significant, χ^2 (1) = .01, $p = .94$.

Discussion

Previous literature suggests that when demonstrating toys to an infant, parents tend to modify their actions in an infant-directed manner (i.e., motionese). This is known to increase infants' attention to actions and is thought to help infants to identify individual sub-steps of actions (e.g., Brand et al., 2002, 2007, 2008). Motionese also promotes infants' and toddlers' attempts to imitate actions (e.g., Koterba & Iverson, 2009; Williamson & Brand, 2014). It therefore seems feasible that infant-directed actions could help support infants' understanding of the overarching intentional purpose of a series of actions, such as stacking rings to make a tower. Exposure to infant-directed actions during day-to-day interactions between infants and their caregivers could also contribute to infants' capability to identify the intentional structures of actions that they come across more broadly. In the current study, we investigated the extent to which variability in mothers' infant-directed actions may modulate 10-month-olds' ability to understand, and thus predict, the outcome of a novel intentional action. In a parent-infant interaction task, mothers demonstrated four toys to their infant. In a separate visit, infants completed an eye-tracking paradigm (detailed in the previous chapter) in which we measured their ability to visually predict the outcome of an incomplete action based on a person's prior goal. We coded parent-infant interactions for various dimensions of motionese using global and fine-grained behavioural coding scales adapted from Brand et al. (2002, 2007).

Our main findings revealed that no individual dimensions of global motionese, or global motionese as a whole, were associated with the proportion of accurate predictions that infants made across all four test trials in our eye-tracking task. Analyses of our fine-grained measures of motionese revealed that infants whose mothers passed them toys more frequently (i.e., exchanges) made the highest proportion of accurate predictions across all four test trials. When analysing only the first trial in which infants made any type of prediction (i.e., accurate or inaccurate), infants exposed to more toy exchanges also made the highest number of accurate first predictions. Exchanges between mothers and infants were rare, with some mothers making no exchanges with their infants and others making only a few. Previous research by Brand et al. (2007) found that the number of object exchanges between mothers and infants were higher when demonstrating to older 11- to 13-month-olds as compared with younger 6- to 8-month-olds, possibly due to older infants' capacity to switch attention (Ruff & Rothbart, 1996). It is important to note that exchanges in Brand et al.'s (2007) study were coded differently. That is, exchanges were counted each time possession of a single toy was transferred between a mother and infant. In our study, we could not count exchanges in the same way, as mothers and infants could both be in possession with a part, or multiple parts, of the toy at one time (e.g., multiple stacking rings). We therefore operationalised exchanges as each time the mother relinquished a toy from her own hand to the infant's hand. Our findings suggest that object exchanges are particularly helpful for infants' understanding of novel observed actions. This is likely because exchanges help to mark boundaries between the intentional sub-steps within actions and keep infants engaged with the activity (Brand et al., 2007). Longitudinal work could further investigate whether exchanges between mothers and infants increase with infants' age and how exchanges help promote the development of action understanding.

Unexpectedly, the current findings also demonstrated that less shared manipulation of toys (i.e., joint contact) between mothers and infants was associated with a higher number of accurate first predictions. This seems counter-intuitive as we expected that joint contact with toys, especially when scaffolding infants' actions or performing joint actions, would help to facilitate intentional understanding. Instead, our findings imply that when engaged in joint contact with toys, mothers may have interfered with their infant's manipulation of the toy and potentially the infant's attempt to imitate actions (e.g., shaking and banging behaviours as in Koterba & Iverson, 2009). This ties in with the idea of parental autonomy support, as opposed to controlling behaviour, in contexts of infant exploration. Examples of such supportive behaviour can include waiting for the infant to require assistance before intervening, providing informative feedback and encouragement, as well as assisting the infant appropriately given their abilities (e.g., Grolnick et al., 1984, 2002; Whipple et al., 2011). Based upon our findings, future research could better distinguish between shared actions (e.g., assisting the infant in posting a shape into a shape sorter) and parental interference (e.g., inhibiting the infant's exploratory behaviour) when coding infant-directed action.

Based on previous literature which found that parents modulate their action demonstrations dependent on the infants' motor skills (e.g., Fukuyama et al., 2015; van Schaik et al., 2019), we used a motor task to assess infants' ability to perform tool-use actions with objects that were different to those demonstrated in the mother-infant interaction task. However, we found no association between infants' proficiency at performing tool-use actions and mothers' use of motionese. Although van Schaik et al. (2019) took mothers' evaluations of their infants' general motor proficiency into account, they also assessed the relation between infants' successful attempts at producing the demonstrated action and mothers' subsequent action modulations. Fukuyama et al. (2015) similarly assessed infants' attempts to perform the demonstrated cup stacking action. It may be that infants' attempts at performing task-relevant manipulations of demonstrated actions (e.g., putting one cup inside another) in the current study is related to mothers' in-the-moment modulations of motionese during our interaction task, rather than infants' broader motor ability.

From a methodological standpoint, the current study used both a global and fine-grained coding scheme adapted from previous motionese literature. Although the fine-grained coding scheme attempted to better capture the global measures of interactiveness and simplification, these global and fine-grained measures did not significantly correlate with one another. This leads us to question the extent to which these coding schemes measure the same or different features of mother-infant interactions. Although the global coding scale of motionese includes joint attention as an attribute of interactiveness, the fine-grained method of quantifying interactiveness by coding mothers' gaze to their infant's face did not reciprocally capture infants' visual attention to their mother's face or to the toy. This type of attention has been found to facilitate infants' intentional understanding in a separate strand of research in which shared attention to an object of mutual interest has been measured. This is known as triadic engagement (Brandone et al., 2019). As described in more detail in the first chapter of this thesis, Brandone and colleagues found that 8- to 9-month-olds who engaged in more bouts of triadic engagement (i.e., attentional shifts between a toy, an experimenter's face, and the same toy) demonstrated better understanding of an intentional action in an eye-tracking paradigm. In addition to this, the tendency to engage in triadic interaction at 6- to 7-months of age predicted intentional action understanding three months later. As joint attention is defined as a global feature of interactiveness within motionese, gaze shifts to the parents' face as well as to the toy being demonstrated could also be coded as part of fine-grained motionese in future work. This could also enable researchers to measure the extent to which exchanges and joint contact modulate infants' attention.

Whilst the previous chapters in this thesis focused on the role of comparison as a cognitive factor that can facilitate action understanding, the current study focused on motionese as a separate social factor that facilitates action understanding. As infants also underwent a comparison training intervention in the current study, we assessed whether there was any interaction between comparison and motionese upon infants' ability to make accurate predictions in our eye-tracking paradigm. However, our findings were not significant. Nevertheless, in-the-moment comparisons between actions may still occur within parent-infant interactions. For example, an infant may compare or align their parent's actions of stacking a ring with their own attempts to place a ring on the tower. It is possible that both comparison and motionese play a joint role in contributing to the development of action understanding by creating a collaborative context in which infants can compare their own goal-directed actions with another's.

To our knowledge, this is the first study to provide evidence that modulations in infant-directed action promote infants' understanding of others' intentional actions. More exchanges and less joint contact of toys between mothers and infants were most beneficial for facilitating action understanding, likely due to the effectiveness of exchanges in attracting infants' attention to intentional sub-units of actions (Brand et al., 2007). However, future work should also assess how these dimensions may modulate infants' attention and thus how differences in infant attention contribute to action understanding (e.g., Brandone et al., 2019).

Chapter 5

Adult mirror responses to perceptually similar yet semantically distinct actions

Abstract

When observing another person perform a goal-directed action, the objects and tools that are being manipulated inform our prediction of how the action will unfold. This is because we hold learned knowledge of conventional actions that are usually performed with certain items. In study 4, we replicated an EEG paradigm previously used with 10-month-old infants (Ní Choisdealbha, 2015) to investigate whether the adult sensorimotor system shows a differential response to already learned actions that are perceptually similar at onset yet differ in the way that they unfold (i.e., the action is performed in a conventional or unconventional manner). In a within-subjects design, adult participants (*N* = 20) observed actors grasping and eating from spoons using identical grasps that differed in orientation relative to the spoon. These differing grasps led to outcomes that are motorically congruent or incongruent with standard spoon use. Event-related desynchronisation in the alpha (8 – 12Hz) and beta (15 – 25Hz) frequency bands over central and occipital regions was measured and compared for congruent and incongruent grasping and eating actions. Our main findings were similar to those seen in infants with more proficient self-feeding skills, such that adults showed greater sensorimotor activation during incongruent actions at the stage that the spoon reached the mouth. This finding was specific to the beta, rather than the alpha, bands demonstrating a response that is specific to motor, rather than attentional, processes. The contrast between incongruent and congruent actions at the grasp stage of the action did not reach significance. We conclude that the sensorimotor system responds to the stage of the action that is most motorically unusual and difficult to simulate. These results are in line with predictive coding accounts of mirror system function such that existing representations of the action require updating and this induces greater sensorimotor activation.

Introduction

In the previous empirical chapters, we investigated the role of observational comparison, motor ability, and infant-directed actions upon infants' ability to predict the end-goal of a novel action. Beyond goal-directedness alone, other literature suggests that the semantic, or meaningful, features of objects and tools also plays a role in an individual's prediction of how an action will unfold. That is, the type of object being acted upon may be associated with learned knowledge that specifies what to do with the object (e.g., the semantic representation that food is meant to be eaten). Similarly, the type of tool being used to perform an action is often associated with a target object or category of target object (e.g., cutlery for use with food), as well as specific actions related to the tool's conventional use (e.g., how to grasp and manipulate cutlery). Therefore, an individual holds a semantic representation of the tool's function as well as how to manipulate the tool to achieve its function (e.g., Creem & Proffitt, 2001; De Bellis et al., 2016; van Elk et al., 2010b). This means that the objects or tools that are present inform an individual's expectation of the correct action to be performed (e.g., Gerson et al., 2017). For an action to be semantically correct, it is important that there is contingency between the action and the target or goal. For example, putting a spoon in one's mouth would not be eating if there were no food on the spoon. The way in which an object is grasped informs an individual of the motor commands needed to successfully complete a functional action. For example, the orientation of a tool relative to the grasp acts as a predictor of how an action will unfold. Therefore, semantic information that is present at the onset of an action can inform action prediction within the sensorimotor system.

Within the first year of life infants begin to process every-day actions, such as spoon-feeding, semantically. They are able to predict that a tool will be brought to the correct target, such as bringing a spoon to the mouth or a phone to the ear (e.g., Gredebäck & Melinder, 2010; Hunnius & Bekkering, 2010). Infants also show differential neural responses to outcomes that are congruent or incongruent with the initiation of the action (e.g., Kaduk et al., 2016; Reid et al., 2009; Stapel et al., 2010), as well as when actions are performed in an unexpected manner (e.g., Langeloh et al., 2018). In addition to the semantic relations between tools and target objects, infants are sensitive to the types of grasps used to obtain an object. For example, Daum et al. (2009) presented 6- and 9-month-olds with stimuli in which a person reached for an occluded object. When the previously occluded object became visible, 6- and 9-month-olds looked longer when the outcome of the grasp was incongruent with the shape of the hand during the reach (i.e., the size of the object was smaller or larger than expected based on hand aperture during the reach), relative to when the reach and grasp were congruent (i.e., hand aperture was appropriate for the size of the target object). Therefore, infants as young as 6-months of age are able to predict the size of an unseen object based on the characteristics of the grasp. Such sensitivity to the characteristics of a grasp leading to a particular outcome is related to infants' ability to execute grasping actions (Loucks & Sommerville, 2012).

As described in the general introduction of this thesis, the execution and observation of actions elicits mirrored neurocognitive representations in the sensorimotor system. EEG studies have found mu (sensorimotor alpha) and beta desynchronisation patterns during the observation of actions that are similar to those involved in the execution of the same action.

Observed actions that are more motorically familiar elicit greater sensorimotor activation (i.e., lower alpha and beta power relative to baseline) in both adults (e.g., Cannon et al., 2014; Gerson et al., 2017) and infants (e.g., Cannon et al., 2016; van Elk et al., 2008). In adults, motor mirroring is recorded over fronto-central and central electrodes and occurs in the 8 to 12Hz frequency bands for mu as well as between 15 to 25Hz frequency bands for beta (e.g., Cannon et al., 2014; Gerson et al., 2017; Hari, 2006; Hari et al., 1998; Ménoret et al., 2015; Quandt et al., 2012). During infancy, mu rhythm manifests at a lower 6 to 9Hz frequency (e.g., Cannon et al., 2016; Gerson et al., 2015a; van Elk et al., 2008). Infant beta is less well characterised, with fewer studies including the beta band (cf., van Elk et al., 2008; Virji-Babul et al., 2012).

Whilst it is known that the motor system is activated in response to actions that are motorically familiar (e.g., Calvo-Merino et al., 2006; Fox et al., 2016; Gerson et al., 2015a; Stapel et al., 2016), other literature suggests that the motor system is also activated in response to actions that violate previous semantic knowledge (e.g., Cross et al., 2012; Gerson et al., 2017). According to the predictive coding account of mirror system function (Kilner et al., 2007), the motor system incorporates semantic and motor processes when forming predictions about an unfolding action. The observation of the first step of an action (e.g., grasping a spoon containing food) provides us with an assumed goal and thus an expectation of the action outcome (e.g., eating from the spoon). In addition to this, we also make a prediction of the kinematics we expect to observe based on our representation of the motor commands involved in the action. However, when an unexpected action is carried out, the difference between the observed and predicted action results in a prediction error, and thus the need to update our representation of the action in order to reduce this error. Prediction error is minimised when the observed action is similar to the predicted action. This is consistent with the aforementioned literature that has found increased sensorimotor activation in infants' response to actions that are unexpected (Kaduk et al., 2016; Stapel et al., 2010). Comparable findings have also been borne out within the adult literature (e.g., Koelewijn et al., 2008). Therefore, the sensorimotor system does not simply mirror an observed action but is sensitive to the semantic correctness or expectedness of the action due to the type of tool or object that is being acted upon. To some extent, the predictive coding account can also explain why the mirror system responds to actions that are incomplete or occluded (e.g., Southgate et al., 2010; Umiltà et al., 2001) by inferring the end-goal of an action based on presentation of the action onset. In addition to this, sensorimotor system activation has been found in response to simply viewing a picture of a tool that is associated with a certain function (Proverbio, 2012). Previous literature has also found sensorimotor activation during the preparatory stages of goal-directed actions (Ní Choisdealbha, 2015; Southgate et al., 2010). For example, Southgate et al. (2010) presented 9-month-olds with an incomplete action that involved a hand disappearing behind an occluder. The authors found greater sensorimotor activation before a hand with a grasping posture disappeared behind the occluder in comparison with a back of hand posture. This is because infants interpreted the grasping hand as goal-directed and thus predicted an action outcome. These findings support Daum et al.'s (2009) aforementioned looking time findings, as infants in both studies predicted a particular outcome based on the characteristics of a grasp posture. Related findings come from a study by Ní Choisdealbha (2015) in which 10-month-olds observed complete actions involving actors grasping a spoon and eating from it. Once the grasp was accomplished, sensorimotor activation decreased, followed by an increase in sensorimotor activation as infants predicted the next step of the action (i.e., eating from the spoon). These findings suggest that the sensorimotor system encodes an action in terms of its individual sub-steps, rather than as a single continuous movement. However, infants in Ní Choisdealbha's (2015) study were not all proficient at using a spoon to eat, leading us to question whether the more sophisticated adult sensorimotor system would respond differently.

There is literature to suggest that semantic knowledge and sensorimotor system function interact (e.g., Creem & Proffitt, 2001; De Bellis et al., 2016; Koelewijn et al., 2008; van Elk et al., 2010b). For example, Creem and Proffitt (2001) found that semantic processing of a tool's function informs an individual's execution of an appropriate grasp, such as picking up a toothbrush by its handle rather than by its bristles. When participants also performed an unrelated semantic distractor task, such as naming an object that is semantically associated with another (e.g., trumpet – piano), this affected their ability to execute a functional grasp. In contrast, the performance of an unrelated visuo-spatial task did not affect participants' grasps. This means that interference of appropriate grasping was specific to semantic rather than visuomotor processes. Therefore, semantic representations of objects inform the visuomotor system with regards to how to grasp a tool in a way that is appropriate for its use. van Elk et al.'s (2010b) EEG study provides further support for the neural processes underlying the interaction between semantic and sensorimotor processes. Sensorimotor activation was measured during the performance of actions involving differing grips of every-day objects that resulted in semantically correct or incorrect end postures when the object was brought to the correct goal location. Findings revealed greater beta desynchronisation during the execution and maintenance of semantically incorrect end postures relative to correct end postures. Incorrect postures also resulted in a stronger subsequent beta-rebound during maintenance of the posture. Therefore, object knowledge seems to be organised around postural representations within the sensorimotor system. Similar findings in beta desynchronisation have also been found in response to simply observing incorrect actions (Koelewijn et al., 2008).

A recent study by Gerson et al. (2017) investigated the extent to which the sensorimotor system responds to observed actions that vary independently in terms of motoric and semantic familiarity. Adult participants' experience with motor skills and semantic knowledge was manipulated in a week-long training paradigm. Motor training with a novel action (e.g., using chopsticks) increased sensorimotor activation (i.e., beta desynchronisation) when observing the same action relative to observing the novel action before training. On the other hand, sensorimotor activity decreased (i.e., beta synchronisation) when a semantically unfamiliar action (e.g., using pliers to pick up food) was given a context to make the observed action more familiar (e.g., *this* person used pliers to eat). The authors concluded that motor familiarity and semantic familiarity with an action are processed simultaneously yet independently within the sensorimotor system. In line with the predictive coding account (Kilner et al., 2007), Gerson and colleagues' findings suggest that semantic knowledge contributes to the updating of predictions as an action unfolds. That is, providing participants with new semantic knowledge during training allowed them to update their motor representation of the action, hence the decrease in sensorimotor activity after training due to minimising prediction error.

In another recent study that focused on semantic knowledge of tools, De Bellis et al. (2016) presented participants with pairs of tools that had different functions but were associated with similar functional hand grips. For example, using nail clippers and a small stapler are associated with the same functional manipulations. This was compared with scenarios in which the pairs of tools shared the same goal but required different functional hand grips (e.g., a basic corkscrew vs a winged corkscrew). When asked whether tools matched with one another in terms of functional goals, participants were slower to make such judgements when tool pairs had different functional goals but were associated with the same hand grip. The authors speculated that this may be due to competing motor representations for each action when attempting to process non-motor related semantic features of objects.

In an attempt to investigate how the sensorimotor system responds to grasping actions that differ in semantics and thus provide the observer with an expectation of how the ensuing action will unfold, the current EEG study presented adult participants with videos of actors grasping and eating from spoons. In a within-subjects design, identical grasping actions were performed in each condition. However, the orientation of the spoon in each grasp differed between conditions and thus was semantically congruent or incongruent with the upcoming functional action, and required predictions about what would happen next. That is, differing orientations led to differing kinematic outcomes of eating from a spoon which are either motorically typical of conventional spoon use or motorically atypical.

The paradigm used in the current study is a direct replication of that used in Ní Choisdealbha's (2015) study, which investigated the extent to which the infant sensorimotor system responds to typical and atypical spoon use actions. In a within-subjects design, 10-month-olds observed videos of actors grasping a spoon using a radial grip (i.e., the functional end of the spoon was near the actor's index finger/thumb; see figure 15a and b) that led to an action congruent with standard spoon use. In other scenarios, actors grasped the spoon using an ulnar grip (i.e., the functional end of the spoon was near the little finger; see figure 15c and d) leading to an action that was incongruent with standard spoon use. Mu rhythm (6 to 9Hz in infants) was measured during the actor's grasp of the spoon, as well when the spoon was brought to his/her mouth. This allowed the author to investigate the extent to which infants are able to differentiate between typical and atypical eating actions, as well as whether they are able to make predictions about the ensuing action from the point that the hand grasps the spoon. In a similar context to which we measured infants' own motor skill in performing tool-use actions in the previous chapter, Ní Choisdealbha (2015) also measured infants' own motor planning skills during spoon use. Efficient motor planning factors included picking up the spoon with few touches, maintaining correct orientation of the spoon from pick up to conclusion, as well as successfully eating from the spoon. Across all infants, no effect of action congruence on sensorimotor activation was found. However, infants who were more motorically skilled at performing self-feeding showed a greater difference in sensorimotor activation during the observation of incongruent relative to congruent eating actions only at the point that the spoon reached the actor's mouth. Regardless of motor ability, infants did not show any difference in sensorimotor activation during the execution of the grasp between conditions. Therefore, at only 10-months of age, infants do not yet process the semantics of a grasp in relation to the ensuing action.

At this age, infants' own skills at using a spoon to self-feed were still emerging, with many of the most planful infants needing to adjust their grasp or swap hands in order to eat from the spoon successfully. Due to a lack of experience with self-feeding, infants of this age may not have formed semantic associations between the orientation of a grasp and subsequent success in using a spoon to eat. Other literature has found that, although planning and execution of spoon use begins to develop in the first year of life, infants cannot consistently execute efficient grasping and self-feeding with a spoon until around 14-months of age (Kaur et al., 2020; McCarty et al., 2001).

In order to address why infants did not show any difference in sensorimotor activation during the grasping stage of the action, we first need to establish whether adults who are proficient at performing these actions themselves show differential sensorimotor responses when observing actions that are perceptually similar yet semantically distinct. Once this is established, this will allow us to more systematically investigate the developmental emergence of semantic representations of actions within the sensorimotor system. In our replication of Ní Choisdealbha's (2015) paradigm, we expect to find greater sensorimotor activation (mu and beta) in response to incongruent actions relative to congruent actions. We also hypothesise that the sensorimotor system will differentiate between grasps that are perceptually similar yet differ in semantics. That is, we expect to find greater sensorimotor activation during observation of a grasp that is incongruent with standard spoon use, relative to a grasp that is congruent, based on the orientation of the spoon. This would support a predictive account of mirror system function that incorporates semantic representations of tool-use actions before the action outcome is apparent perceptually (Kilner et al., 2007). Further to this, we expect to see more sensorimotor activation over the central region as compared with the occipital region, demonstrating a motor specific response. These results in an experienced adult sample would add further credibility to Ní Choisdealbha's (2015) findings in infants with emerging self-feeding skills and may pave the way for future research in developmental populations with more advanced self-feeding skills. Whilst the other empirical chapters in this thesis investigate the factors influencing infants' prediction of unfamiliar tool-use actions that they have never seen before; this research will shed light on how the correctness of an already learned tool-use action alters an individual's prediction of an ensuing action.

Method

Participants

Based on an a priori power calculation with an effect size of .25, alpha level of .05, and beta power of .8, we recruited a sample of 24 female undergraduate students at Cardiff University to participate in this study. However, data from the first four participants were excluded due to data collection errors, leaving a final sample of 20 participants (M_{age} = 19.25, *age range* = 18.36 – 20.33). An all-female participant sample was recruited because the same participants also underwent a separate unrelated EEG study that specifically required females. There were no expected carryover effects from the other EEG paradigm. The current study took approximately six minutes to complete, whereas the unrelated study took approximately 30 minutes to complete. To control for any potential effects of fatigue, the order in which participants completed the EEG tasks was counterbalanced. We also ensured that data were qualitatively similar between the orders in which participants completed the two studies.

Stimuli

Stimuli were identical to those used in Ní Choisdealbha (2015). Video stimuli involved actors picking up and eating from spoons. Actors were seated at a table with a spoon resting on a plate in front of them. Three different actors were used. For each actor, a total of eight scenarios were filmed. The handle of the spoon was placed on the right or left side of the bowl.

The actor then used their right or left hand to pick up and eat from the spoon. For visual variety, the spoon contained either vanilla yoghurt or chocolate mousse. These actions were timed using a metronome to maintain consistency across actors and scenarios. Videos were each approximately 6 seconds long with some natural variation. The time at which the actor reached and grasped the spoon, followed by bringing the spoon to the mouth and eating from it was similar across videos and, importantly, did not differ between conditions. Stimuli did not contain any audio and were edited so that the actor's face was only visible from the nose down.

The critical manipulation involved differing the orientation of the spoon*.* We also counterbalanced the hand the actor used to grasp the spoon (i.e., right hand vs left hand). In half of all scenarios, the spoon was grasped using a radial grip. That is, the forefinger and thumb were nearest the bowl of the spoon, as in a normal (congruent) means of holding a spoon (see figure 15a). The spoon was then brought to the mouth and inserted from the same side of the body as the hand used, resulting in a straightforward eating action (see figure 15b). In other cases, the spoon was grasped using an ulnar grip with the little finger nearest the functional end (see figure 15c). The spoon was then brought across the midline of the body and the bowl of the spoon was inserted at the opposite side of the mouth from the arm used. This resulted in an unusual (incongruent) endpoint that would be motorically difficult to achieve (see figure 15d). The shape and motion of the hand was identical for the reach, grasp, and lifting of the spoon in both left and right-handed grasps and between conditions. At the onset of the video, other than the radial and ulnar sides of the actor's hand to the functional end of the spoon, there were no cues as to how the action would proceed. This only became apparent from the moment the actor moved their arm to either side of the midline. Between the congruent and incongruent conditions, actions were perceptually similar at the grasp time point, whereas actions were perceptually distinct at the point the spoon reached the mouth. During baseline trials, static abstract images were displayed. These were used as a non-biological contrast in which no sensorimotor activity was expected, as well as to help maintain participants' attention to the task. These baseline trials were displayed for 6 seconds to match the duration of the videos and are similar to those used in previous action observation studies (Marshall et al., 2011; Saby et al., 2012).

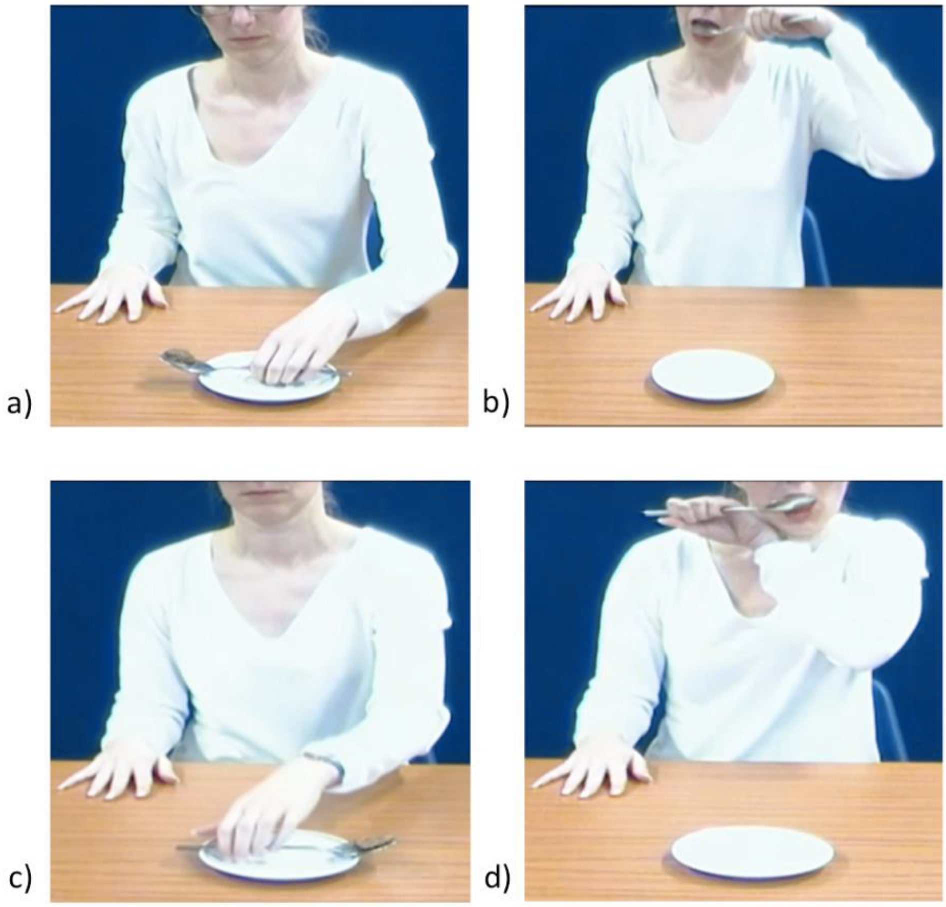

Figure 15. Still frames from stimulus videos. *a* and *b* are frames from a single stimulus video showing a standard radial grasp of the spoon (*a*) which congruently leads to a standard eating action (*b*). *c* and *d* are frames from a single stimulus video showing an incongruent ulnar grasp of the spoon (*c*) which leads to an unusual eating action (*d*).

Procedure

EEG recording

Participants were given standard instructions for EEG research, such as to refrain from movement and try to reduce blinking. EEG data were obtained using a 32-electrode EEG system arranged according to the 10-20 system. The signal was referenced to the vertex during recording, amplified using an actiCHamp amplifier and digitised at 500Hz. Following application of the EEG cap and electrodes to the participant's head, electrodes were checked for impedance ranges. In case of electrodes with poor impedance (above 30kΩ) attempts to improve it were made (i.e., by improving adherence of the electrode to the scalp or adding additional electrolyte gel to the associated well).

Stimuli were presented using E-Prime (Psychology Software Tools, Pittsburgh, PA) and each participant's EEG was recorded using a BrainVision actiCHamp Plus system (Brain

Products, GmbH, Gilching, Germany). The ongoing EEG was monitored using this program from the control room. Presentation order was determined using pseudorandomization software (Mix, van Casteren & Davis, 2006). Congruent and incongruent actions, as well as left- and right-handed actions, could not be viewed twice in a row. No constraints were used for the presentation of baseline images. Participants were shown an equal number of 16 congruent and 16 incongruent stimuli, as well as 32 baseline trials. The total duration of the task was 6 minutes and 24 seconds.

Catch trials

In order to maintain participants' attention to stimuli, eight trials were followed by catch trial questions, distributed pseudo-randomly across all trials. On these catch trials, the question "Did they grasp the spoon in a usual or unusual way?" appeared on screen. The optional answers (usual vs unusual) were presented either side of the screen and participants responded using a corresponding (left or right) button press, respectively.

EEG Data Processing and Analysis

Grasping and eating can be considered as conceptually different goal-directed actions that have distinct differences in terms of predictions to be made. Therefore, responses to the grasp time point and the mouth time point were considered separately. During presentation of each stimulus, the EEG data file was tagged with event marks indicating when the actor's hand first made contact with the spoon in each video (i.e., grasp events), as well as when the spoon first made contact with the actor's lips (i.e., mouth events). These were generally in the same time periods across videos but were determined separately for each video so as to be precise. The average timing of these event marks across stimuli were used as event marks in baseline trials to include comparable timings despite the lack of distinct sub-actions within baseline videos. That is, baseline trials also contained two event marks; one that was comparable in time with hand contact and another that was comparable with mouth contact.

EEG data were processed in MATLAB (version R2019b, TheMathworks, Inc.) using the FieldTrip toolbox (Donders Institute for Brain, Cognition and Behaviour, www.ru.nl/neuroimaging/fieldtrip). Each participant's EEG recording was bandpass filtered with a highpass threshold of 50Hz and a lowpass threshold of 1Hz. A detrending procedure was applied. Each participant's EEG recording was then segmented in line with event files. This data was visually inspected, and any excessive EEG artefacts were removed before applying ICA to remove eye-movement and heart-beat components. A fast Fourier transformation was performed on the data with a 500ms sliding Hanning taper and spectral smoothing of 3Hz. This provided mean power values for each frequency between 1 and 50Hz represented in the EEG data at each electrode for each trial. These methods were the same as those used by Ní Choisdealbha (2015) and follow standard procedures for event-related desynchronisation studies (e.g., Cannon et al., 2014; Gerson et al., 2017). This process was performed once for the grasp events and once for mouth events.

Based on Ní Choisdealbha's (2015) work as well as previous literature that has investigated alpha (e.g., Fox et al., 2016) and beta (e.g., Hari et al., 2006) desynchronisation, the average of these values was obtained for each participant at central electrodes C3, Cz and C4 for each of the alpha frequency bands (8 – 12Hz), as well as each of the beta frequency bands (15 - 25Hz), before and after the grasp and mouth event. Research suggests that sensorimotor desynchronisation (in both alpha and beta frequency bands) is specific to central scalp locations as opposed to non-central locations, such as occipital regions (e.g., Fox et al., 2016; Hari & Salmelin, 1997; Ritter et al., 2009). To confirm this distinction, we also obtained averages for each participant at occipital electrodes O1, Oz and O2 for each of the alpha frequency bands (8 – 12Hz), as well as the beta frequency bands (15 – 25Hz), before and after the grasp or mouth event in each of the three conditions. Average frequency data were analysed 850ms pre and post grasp, as well as 850ms pre and post mouth. 850ms was chosen as a defining event either side to ensure that there was no overlap in the EEG recorded during grasping and eating. A natural log transformation was then performed on the data. This is in line with standard practice for event-related desynchronisation research. Negative values indicate less power and hence more sensorimotor activation. Positive values indicate more power and hence less sensorimotor activation.

Results

In our first analysis, we investigated whether there was any difference in power between the congruent and incongruent conditions relative to baseline across all time points. We used two separate regression models for alpha and beta to investigate the main effects of condition (congruent and incongruent relative to baseline), scalp region (occipital relative to central), and site (left and right electrodes relative to midline). Each model also assessed interactions between condition and region, as well as region and site.

Across central and occipital regions in the alpha band, there was significantly less power (i.e., more sensorimotor activation) in the congruent condition as well as the incongruent condition relative to baseline. Across conditions, there was significantly less power in the central region relative to occipital. The analysis also revealed less power over the midline relative to the right electrode. However, there was no significant effect of the left electrode relative to midline. There were no significant interactions between variables (see table 13). Figure 16 shows alpha power by condition and region collapsed across electrode sites.

Table 13

Results of regression analysis for alpha by condition, region, and site.

	β coefficient	*SE*	t-value	p-value
Main effects				
Intercept	-.74	.12	-6.33	<.001
condition [congruent] (vs baseline)	-.21	.05	-3.94	<.001
condition [incongruent] (vs baseline)	-.34	.05	-6.24	<.001
region [occipital] (vs central)	.86	.07	12.38	<.001
site [left electrode] (vs midline)	.06	.05	1.19	.23
site [right electrode] (vs midline)	.15	.05	2.79	<.005
Interaction effects				
condition [congruent] * region [occipital]	.13	.08	1.70	.09
condition [incongruent] * region [occipital]	.12	.08	1.57	.12
region [occipital] * site [left electrode]	.07	.08	.91	.36
region [occipital] * site [right electrode]	-.04	.08	-.48	.63

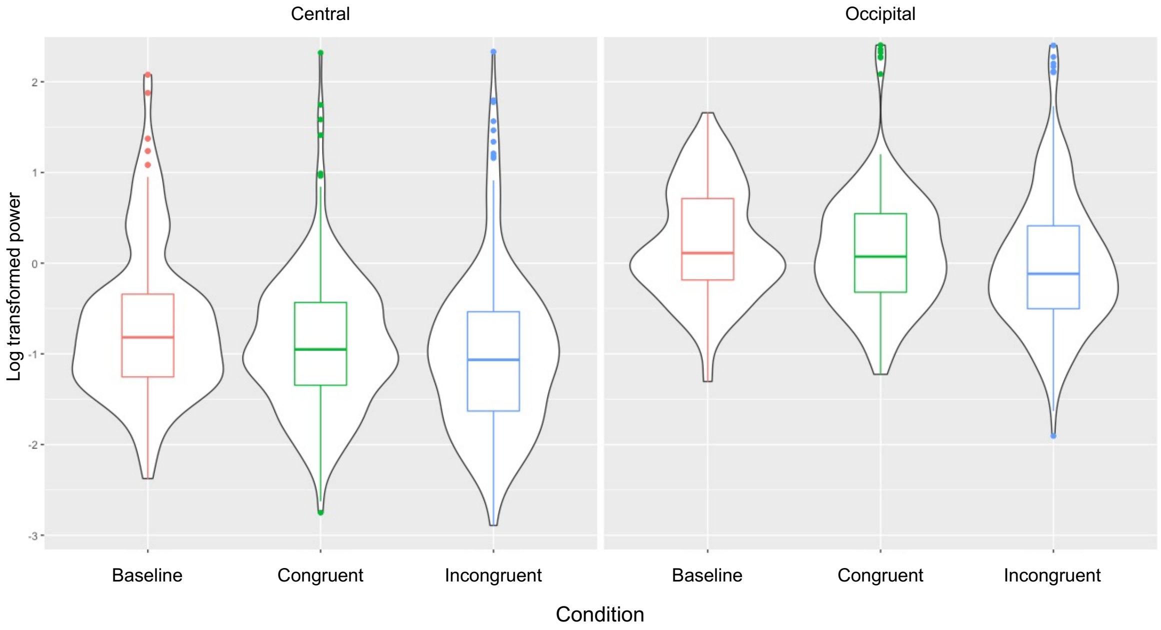

Figure 16. Alpha power by condition and region (collapsed across electrode sites).

Similar results were found in the beta band, with significantly less power in both the congruent and incongruent conditions relative to baseline. There was also less power over the central region relative to the occipital region. A significant interaction between condition and region is followed up below. There was also less power over the midline relative to both right and left electrode sites, indicating that sensorimotor activation is greater along the midline than laterally (see table 14). There were also significant interactions between condition and region, as well as region and site, which are followed up below.

Table 14

Results of regression analysis for beta by condition, region, and site.

	β coefficient	*SE*	t-value	p-value
Main effects				
Intercept	-1.99	.13	-14.79	<.001
condition [congruent] (vs baseline)	-.12	.05	-2.47	.01
condition [incongruent] (vs baseline)	-.21	.05	-4.33	<.001
region [occipital] (vs central)	.93	.06	15.01	<.001
site [left electrode] (vs midline)	.30	.05	6.38	<.001
site [right electrode] (vs midline)	.28	.05	5.84	<.001
Interaction effects				
condition [congruent] * region [occipital]	.20	.07	2.96	.003
condition [incongruent] * region [occipital]	.17	.07	2.53	.01
region [occipital] * site [left electrode]	-.17	.07	-2.47	.01
region [occipital] * site [right electrode]	-.16	.07	-2.41	.02

Follow-up simple effects tests (tukey adjusted) for the condition and region interaction revealed a significant contrast between the incongruent and baseline conditions over the central region, with less power in the incongruent condition than the baseline condition (β = -.21; *SE* = .07; *t* = -2.93; *p* = .003). There were no other significant contrasts between conditions over the central region (*ps* > .09) or the occipital region (*ps* > .09). These results show that there is less power in the central region in response to the incongruent condition relative to baseline and thus align with our hypothesis that the incongruent action is the most challenging to simulate. Figure 17 shows beta power by condition and region collapsed across electrode sites. Simple effects tests (tukey adjusted) to follow up the region and site interaction revealed significant contrasts between electrode sites over the central region, with less power over the midline than the left electrode site (β = -.30; *SE* = .07; *t* = 4.35, *p* < .001) and less power over the midline than the right electrode site (β = -.28; *SE* = .07, *t* = 3.98, *p* < .001). Over the occipital region, there was less power over the midline than the left electrode site (β = -.15; *SE* = .07, *t* = 2.07, *p* = .04). No other contrasts reached significance (*ps* > .09).

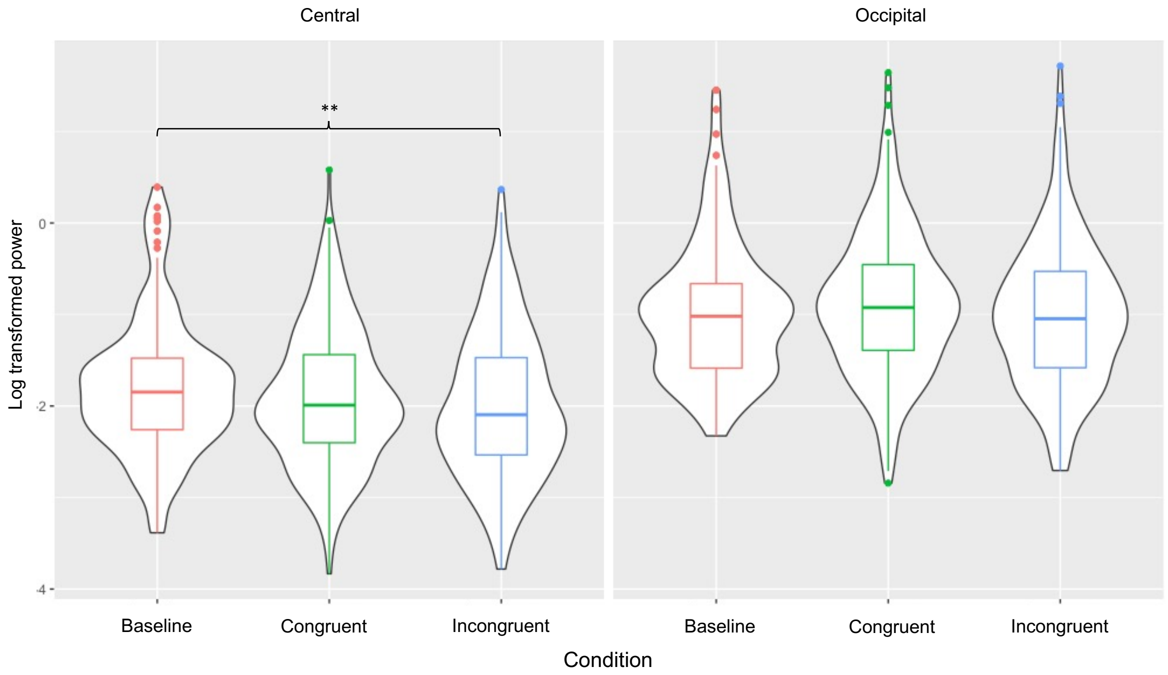

Figure 17. Beta power by condition and region (collapsed across electrode sites). ***p* < .01.

In our next analysis, we investigated our main hypothesis that there will be more sensorimotor activation in response to incongruent relative to congruent actions; thus, we directly contrasted these two conditions. We also assessed whether the stage of the action (i.e., grasp or mouth) elicited a differential response due to a predictive process. We expected that the incongruent grasp may elicit more sensorimotor activation than a congruent grasp due to prediction of an atypical ensuing action. The incongruent mouth stage was also expected to elicit more sensorimotor activation than the congruent mouth stage, as this is perceptually and motorically unusual. In line with previous literature (Ní Choisdealbha, 2015; Southgate et al., 2010) that found more sensorimotor activation before an action stage followed by a rebound upon its completion, we also explored whether there was less power before each action stage relative to after. To address these hypotheses, we used a regression model to investigate the main effects of condition (congruent relative to incongruent), action (mouth relative to grasping), time (post-action relative to pre), region (occipital relative to central), and the interactions between these variables. We also controlled for the effects of, and interaction between, region and site that was revealed in our previous analysis.

In the alpha band, this analysis revealed that there was significantly less power in the central region relative to occipital. There was significantly less power in response to the mouth stage of the action relative to the grasp. There was no significant effect of incongruent condition relative to congruent. However, there was a significant interaction between condition and action, as well as between region and action. The three-way interaction between region, action, and time also reached significance. There were no other significant two-way or three-way interactions (see table 15).

Table 15

Results of regression analysis for alpha by condition, action, time, and region. The interaction between region and site was controlled for.

	β coefficient	*SE*	t-value	p-value
Main effects				
Intercept	-.57	.13	-4.58	<.001
condition [incongruent] (vs congruent)	.01	.09	.12	.90
region [occipital] (vs central)	.72	.10	7.15	<.001
action [mouth] (vs grasping)	-.66	.09	-7.28	<.001
time [post-action] (vs pre-action)	-.08	.09	-.94	.35
site [left electrode] (vs midline)	.03	.06	.54	.59
site [right electrode] (vs midline)	.09	.06	1.70	.09
Interaction effects				
condition [incongruent] * region [occipital]	.12	.13	.96	.34
condition [incongruent] * action [mouth]	-.33	.13	-2.55	.01
region [occipital] * action [mouth]	.42	.13	3.27	.001
condition [incongruent] * time [post]	.09	.13	.68	.50
region [occipital] * site [left]	.09	.08	1.19	.24
region [occipital] * site [right]	.01	.08	.08	.94
condition [incongruent] * region [occipital] * action [mouth]	-.18	.18	-1.02	.31
condition [incongruent] * region [occipital] * time [post]	-.29	.18	-1.59	.112
condition [incongruent] * action [mouth] * time [post]	-.06	.18	-.34	.74
region [occipital] * action [mouth] * time [post]	-.44	.18	-2.44	.02
condition [incongruent] * region [occipital] * action [mouth] * time [post]	.41	.26	1.62	.11

Follow-up simple effects tests (tukey adjusted) for the condition and action interaction revealed significantly less power for the incongruent mouth stage of the action than the congruent mouth stage (β = -0.30; *SE* = .08; *t* = -3.91; *p* < .001). The contrast between incongruent and congruent conditions for the grasp stage of the action did not reach significance (β = .04; *SE* = .08; *t* = .57; *p* = .57). Figure 18 shows alpha power by condition and action stage collapsed across regions.

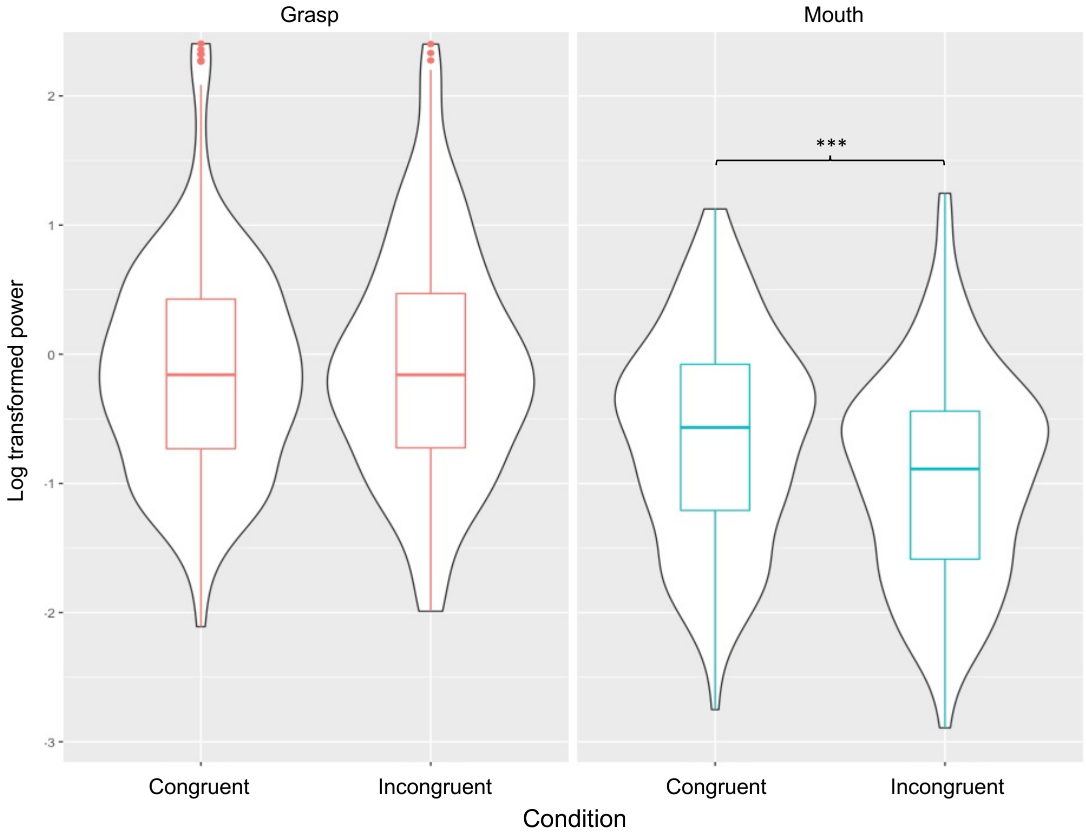

Figure 18. Alpha power by condition and action (collapsed across regions). ***$p < .001$.

Simple effects tests (tukey adjusted) to follow up the region and action interaction found less power over the central region than the occipital region for both the grasp stage (β = -.90; *SE* = .06; *t* = 14.53; *p* < .001) and the mouth stage (β = -1.11; *SE* = .06; *t* = 17.89; *p* < .001), further supporting our finding that sensorimotor activation is specific to central regions. Figure 19 shows alpha power by region and action stage collapsed across conditions.

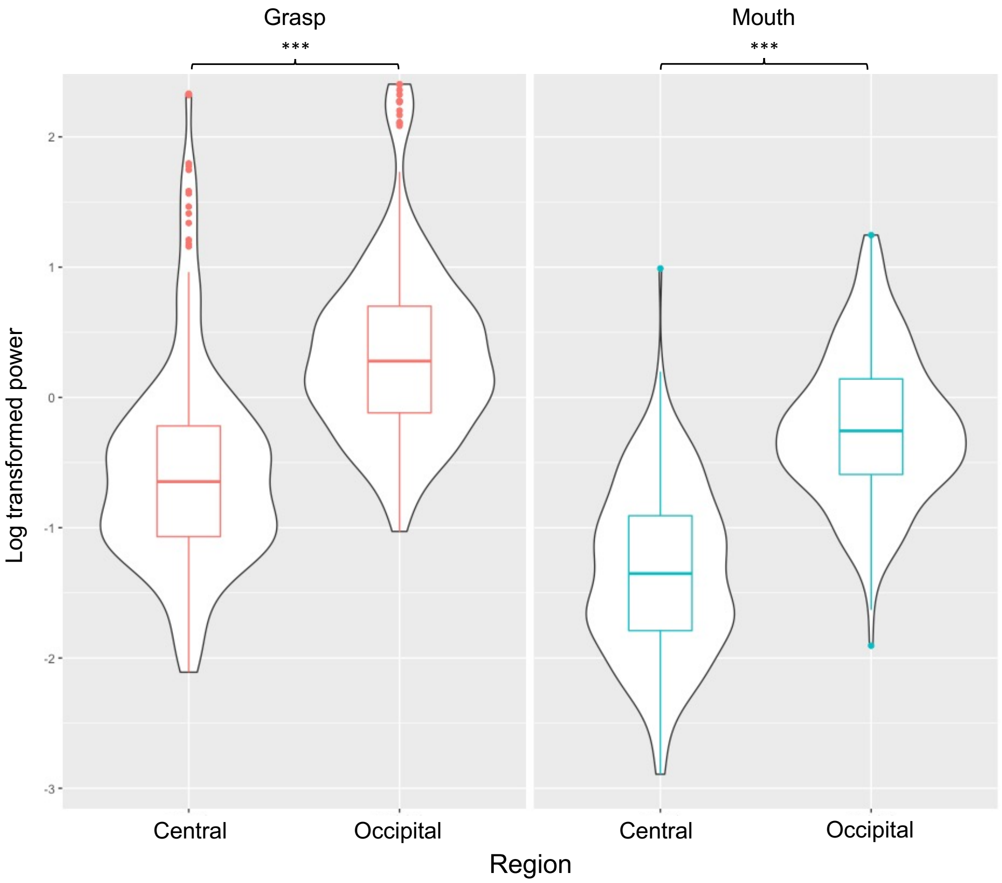

Figure 19. Alpha power by region and action (collapsed across conditions). ****p* < .001.

Simple effects tests (tukey adjusted) to follow up the three-way interaction between region, action, and time found significant contrasts that further support our finding that the mouth stage of the action elicits more sensorimotor activation than the grasp stage of the action both before and after the action has been carried out (see table 16). For both central and occipital regions, time contrasts between the same action did not reach significance. That is, there was no significant difference in power between pre and post grasp, or pre and post mouth. Therefore, our findings do not support previous literature which suggested more sensorimotor activation during the preparatory stages of action sub-steps (Ní Choisdealbha, 2015; Southgate et al., 2010). Figure 20 shows alpha power by time, region, and action stage collapsed across conditions.

Table 16

Simple effects contrasts for alpha by region, time, and action.

	β coefficient	*SE*	t-value	p-value
Region = central				
pre mouth (vs pre grasp)	-.82	.09	-9.35	<.001
post grasp (vs pre grasp)	-.04	.09	-.48	.63
post mouth (vs pre mouth)	.01	.09	.06	.95
post mouth (vs post grasp)	-.77	.09	-8.82	<.001
Region = occipital				
pre mouth (vs pre grasp)	-.49	.09	-5.59	<.001
post grasp (vs pre grasp)	.11	.09	1.29	.20
post mouth (vs pre mouth)	-.07	.09	-.85	.40
post mouth (vs post grasp)	-.68	.09	-7.72	<.001

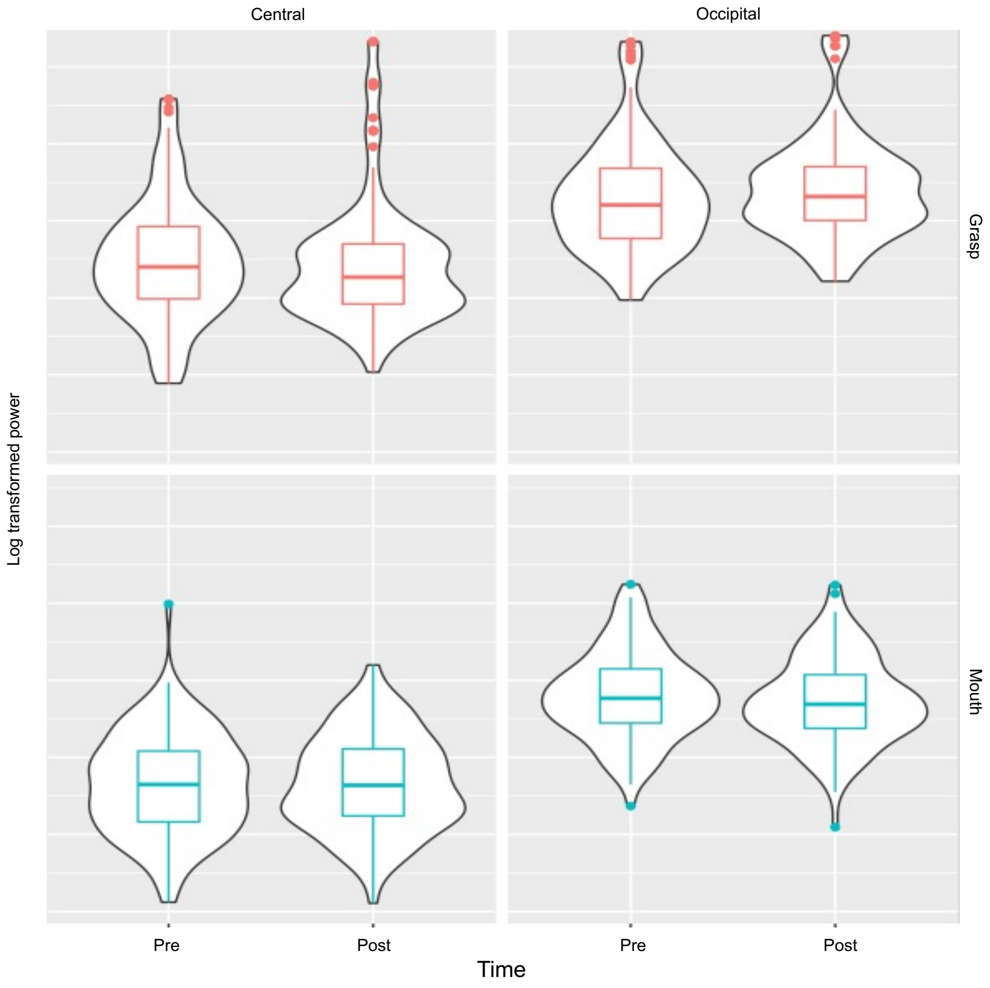

Figure 20. Alpha power by time, region, and action (collapsed across conditions).

In the beta band, our regression model revealed significantly less power over the central region relative to occipital (β = -1.10; *SE* = 0.10; *t* = 11.57; *p* < .001), as well as less power in response to the mouth stage of the action relative to the grasp (β = -.57; *SE* = .08; *t* = -6.75; *p* < .001). No other significant main effects (*ps* > .64), two-way interactions (*ps* > .05), or three-way interactions (*ps* > .15) were revealed. As we found no effects of time or site in this model, we further simplified the model by removing these two variables (see table 17). This again revealed less power over the central region relative to occipital and less power in response to the mouth stage of the action relative to the grasp. There was also a significant interaction between condition and action stage. The two-way interactions between condition and region as well as region and action stage did not reach significance. However, the three-way interaction between these variables was marginally significant.

Table 17

Results of regression analysis for beta power by condition, region, and action.

	β coefficient	*SE*	t-value	p-value
Main effects				
Intercept	-1.65	.13	-12.28	<.001
condition [incongruent] (vs congruent)	.03	.06	.45	.65
region [occipital] (vs central)	.97	.06	15.77	<.001
action [mouth] (vs grasping)	-.53	.06	-8.60	<.001
Interaction effects				
condition [incongruent] * region [occipital]	.14	.09	-1.65	.10
condition [incongruent] * action [mouth]	-.23	.09	-2.69	.007
region [occipital] * action [mouth]	.09	.09	1.02	.31
condition [incongruent] * region [occipital] * action [mouth]	.23	.12	1.86	.06

Like our findings in the alpha band, simple effects tests to follow up the condition and action interaction found significantly less power for the incongruent mouth stage of the action than the congruent mouth stage (β = -.16; *SE* = .08; *t* = -2.00; *p* = .05). The contrast between incongruent and congruent conditions for the grasp stage of the action did not reach significance (β = -.04; *SE* = .08; *t* = -.53; *p* = .59). Figure 21 shows beta power by condition and action collapsed across regions.

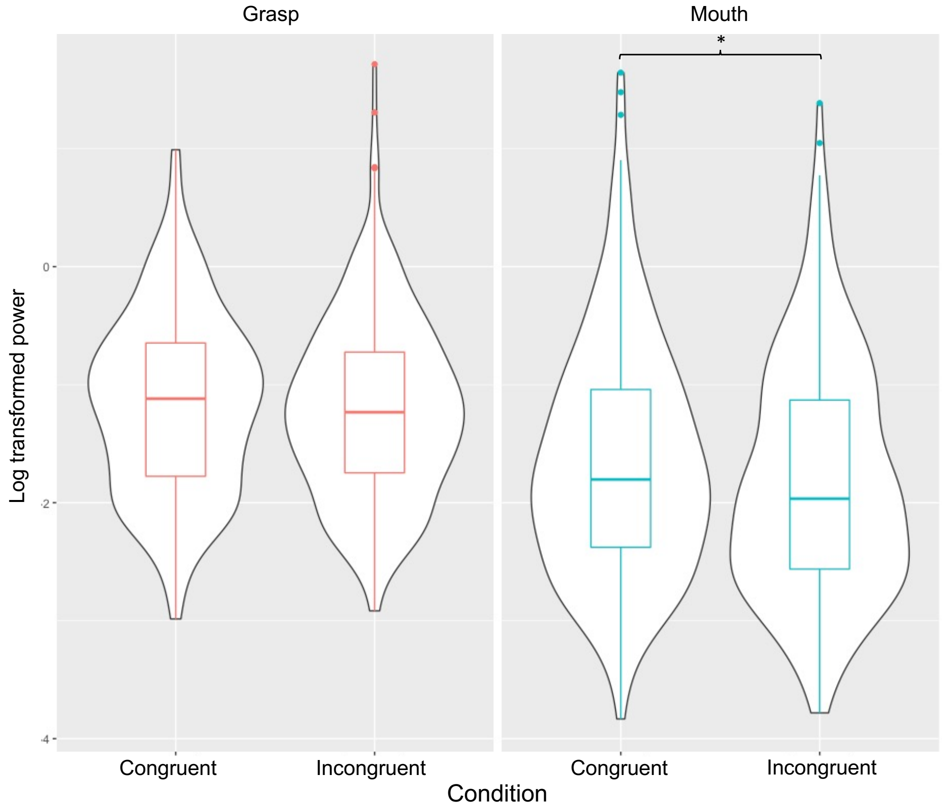

Figure 21. Beta power by condition and action (collapsed across regions). **p* = .05

Simple effects (tukey adjusted) to follow up the three-way interaction between action stage, condition, and region found significant contrasts that support our finding that there is less power in the central region than the occipital region in response to the observation of both congruent and incongruent actions. The contrasts also confirm that the difference between congruent and incongruent conditions is found in response to the mouth, rather than the grasp, stage of the action (see table 18). Figure 22 shows beta power by condition, action stage, and region.

Table 18

Simple effects contrasts for beta by action, condition, and region.

	β coefficient	*SE*	t-value	p-value
Action = grasp				
incongruent occipital (vs congruent occipital)	-.12	.10	-1.22	.22
congruent central (vs congruent occipital)	-.97	.09	-10.27	<.001
incongruent central (vs incongruent occipital)	-.83	.09	-8.75	<.001
incongruent central (vs congruent central)	.03	.09	.29	.77
Action = mouth				
incongruent occipital (vs congruent occipital)	-.12	.10	-1.26	.21
congruent central (vs congruent occipital)	-1.10	.09	-11.20	<.001
incongruent central (vs incongruent occipital)	-1.15	.09	-12.10	<.001
incongruent central (vs congruent central)	-.21	.09	-2.17	.03

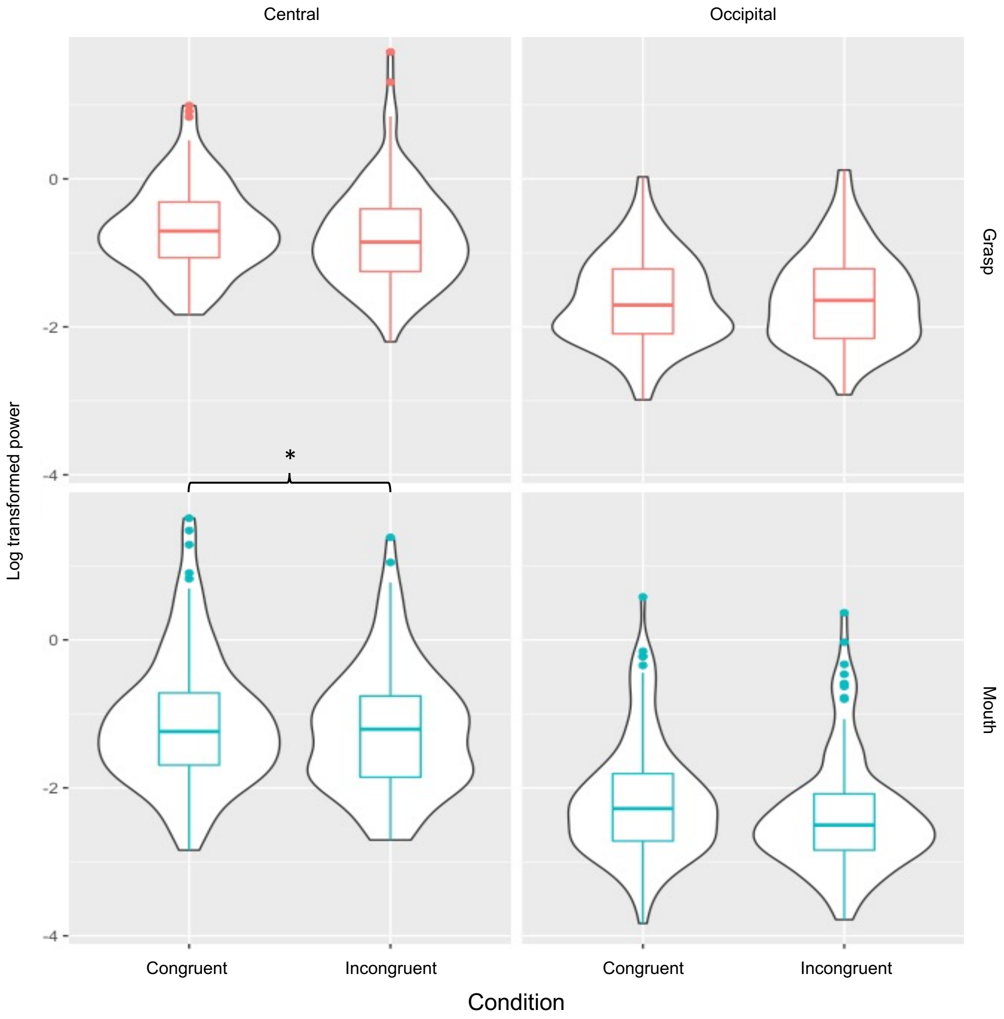

Figure 22. Beta power by condition, action, and region.

Discussion

The current study aimed to investigate whether the adult mirror system differentially responds to motorically identical yet semantically distinct tool-use actions. Due to ambiguous findings regarding the neural processes that underly semantic processing within the developing sensorimotor system, we replicated Ní Choisdealbha's (2015) EEG paradigm using an adult sample to verify how the fully developed sensorimotor system processes identical grasps that differ in their meaning. That is, the hand acted similarly in both conditions up until the point immediately after a spoon grasp was enacted. The differing orientation of the grasp led to an action outcome that was either typical or atypical of conventional spoon use, and thus motorically familiar or unfamiliar, respectively.

Our main findings revealed greater sensorimotor activation in response to atypical actions at the point that the spoon reached the actor's mouth, rather than at the grasp stage of the action. This was found over both central and occipital regions within the alpha band. In contrast, this finding was only present over the central region within the beta band, demonstrating a motor specific response. These results support our hypothesis that motorically experienced adults can differentiate between typical and atypical actions. It is worth noting that the inclusion of catch trials to maintain participants' attention to the task may have primed participants' motor responses in some way. Nevertheless, it is likely that participants would be aware of the difference between actions regardless of this. We also expected that the adult sensorimotor system would detect subtle semantic differences in grasping actions that relate to typical versus atypical ensuing actions. However, we did not find support for this, as there was no difference in sensorimotor activation between conditions at the grasp stage of the action. Instead, our findings are similar to those seen in infants with more proficient self-feeding skills (Ní Choisdealbha, 2015), demonstrating that the sensorimotor system responds to the stage of the action that is most motorically unusual and difficult to simulate. This is in line with the view that actions that are less motorically familiar elicit greater sensorimotor activation (e.g., Cross et al., 2012; Gerson et al., 2017; Koelewijn et al., 2008).

Sensorimotor activation was also specific to the observation of actions (both typical and atypical) compared with a baseline that did not involve observing an action (i.e., viewing an abstract image). We also found greater sensorimotor activation over the central region compared with the occipital region. This demonstrates a motor specific response that supports previous literature (e.g., Fox et al., 2016; Quandt et al., 2012). However, our main finding that sensorimotor activation was greatest during the mouth stage of atypical actions was specific to the beta band over the central region. This may be because alpha desynchronisation is a result of attentional processes (e.g., Snyder & Foxe, 2010), whereas the beta band has mirroring properties that are less sensitive to attentional confounds (e.g., Hari et al., 1998).

With regards to the extent that the mirror system is involved in action prediction, our results do not provide support for the notion that there is greater sensorimotor activation during the preparatory stage of an action, such as before a grasp is complete or before the spoon reaches the mouth (Ní Choisdealbha, 2015; Southgate et al., 2010). Nevertheless, the difference in sensorimotor activation between atypical and typical actions during the mouth stage of the action could be explained in terms of prediction error (Kilner et al., 2007). That is, the observed action results in an unexpected outcome (i.e., bringing the spoon across the body to eat) which results in greater sensorimotor activation due to the need to update the existing representation of the action and thus reduce this error.

An event related component that is thought to be a measure of semantic incongruency is called the N400. Much like the way that greater sensorimotor activation is found when there is greater error between the predicted and observed action, N400 amplitude indicates the level of difficulty in integrating a stimulus into the semantic context (Bach et al., 2009; Sitnikova et al., 2003). This occurs around 400ms after the presentation of a meaningful stimulus and is known to reflect activation of semantic knowledge (e.g., Bach et al., 2014; Decroix et al., 2020; van Elk et al., 2010a). When processing concrete picture stimuli, or during retrieval of visual semantic information, this negative waveform is found over anterior regions (van Elk et al., 2010a; West & Holcomb, 2002). A study by van Elk et al. (2010a) found a stronger N400 when participants were preparing to actively perform meaningful goal-directed actions (e.g., bringing a cup to the mouth) compared with meaningless actions (e.g., bringing a cup to the eye). Further to this, recent findings by Decroix et al. (2020) suggest that the N400 is involved in the semantic integration of motor components (i.e., the correctness of an object grip) as well as goal-related information (i.e., whether the grip enables functional use of the object). The N400 response in Decroix et al.’s (2020) study was strongest when participants observed picture stimuli of incorrect grips that would restrict functional use of the object (e.g., a power grip of an upside-down pencil) than any other combination of grip or object orientation pairing. Despite differences in study design, Decroix et al.’s (2020) findings for grip and object pairings that would result in meaningless, restricted use of an object (i.e., strongest N400 for the most atypical grip and orientation) were in the opposite direction to van Elk et al.’s (2010a) findings (i.e., strongest N400 when performing meaningful object to location pairings). This difference may be due to Decroix et al.’s (2020) focus on differing object grips in contrast with van Elk et al.’s (2010a) focus on performing correct versus incorrect functional actions. A study by Sitnikova et al. (2003) also found a stronger N400 effect when tools were incongruent with the functional action performed (e.g., using a rolling pin instead of a razor to shave) relative to when tools were congruent with actions. Interestingly, N400 amplitude was recorded shortly after objects became visible in the video stimuli, suggesting that the N400 is involved in the processing of object identification and scene comprehension. Based on the role of the N400 during preparatory stages of actions (van Elk et al., 2020b; Decroix et al., 2020) and object identification (Sitnikova et al., 2003), it is questionable as to what extent the N400 may be characteristic of predictive processes. To our knowledge, research is yet to investigate whether the N400 differentiates between grasps of an object that differ in orientation, yet always result in a successful functional goal, as in the current paradigm. Although the sensorimotor system does not differentiate between grasps that are perceptually similar, grasp actions that are incongruent with standard spoon use (compared with congruent grasps) may instead be characterised by a greater N400 amplitude (based on the findings of Decroix et al., 2020). In a developmental population, such differentiation in N400 amplitude may only be found for infants who are proficient at producing the action (Ní Choisdealbha, 2015). Across both adult and developmental populations, N400 amplitude during the grasp may relate to greater sensorimotor activation after the grasp has been enacted, as the ensuing eating action becomes motorically and perceptually peculiar. In this way, N400 at the onset of an action may be predictive of later sensorimotor activation as the action unfolds.

As mentioned in the introduction of this chapter, previous findings suggest that infants look longer to action outcomes that are incongruent with the characteristics of a grasp (Daum et al., 2009) and sensitivity to the characteristics of a grasp leading to a particular outcome is related to infants' ability to execute grasping actions (Loucks & Sommerville, 2012). Therefore, using the current paradigm with a developmental sample, we might expect infants to look longer to incongruent action outcomes as opposed to congruent outcomes; particularly those who are proficient at performing the observed action (Ní Choisdealbha, 2015). Longer looking may also be associated with greater sensorimotor activation (Southgate et al., 2010) as both of these represent the updating of existing action representations.

To conclude, the current study supports previous findings such that there is greater activation in the sensorimotor system during the observation of actions that are motorically unusual or unfamiliar (Cross et al., 2012; Gerson et al., 2017; Koelewijn et al., 2008). However, our results were not in line with our hypothesis that the sensorimotor system would detect subtle semantic differences at the onset of an action in which the orientation of a grasp provides the observer with an expectation of how the action will unfold. Therefore, at least in the current paradigm, semantic knowledge and sensorimotor system function do not interact. We have discussed literature to suggest that the N400 ERP is sensitive to the contextual semantics of grasping tools in different orientations as well whether actions are typical or atypical of conventional tool-use. The extent to which the N400 is involved in predictive processes at the onset of actions could be an interesting future direction for research concerning the online prediction of tool-use actions. Future work could also consider the extent to which longer looking may be associated with greater sensorimotor activation, both of which are characteristic of prediction error.

Chapter 6

General Discussion

Anticipating the outcome of actions performed by others around us is a core part of social functioning. Predicting what our peers are about to do can enable us to adapt our own actions in various social contexts. For example, playing basketball, cooking with a friend, or driving a car all require in-the-moment predictions of others' intentional goals. When cooking breakfast, your friend may place a frying pan on the hob and then open the refrigerator. In knowing that your friend's aim is to fry eggs, you will likely predict that she will take the eggs from the fridge. You could then adapt your own actions in order to collaborate in the activity, such as passing her the spatula. Such understanding of others' intentions begins to emerge within the first year of life and is a precursor to a more complex understanding that other people's thoughts, beliefs, emotions, and desires are different to one's own (Wellman et al., 2004, 2008; Woodward, 1998). This is known as theory of mind, which is classically thought to emerge at around 4 years of age (but see Scott & Baillargeon, 2017 for a review of evidence for earlier emergence). For example, you may anticipate that your friend is going to have the false belief that the eggs are in the refrigerator, and so you can intervene and save her time by telling her that they are in the cupboard. Such understanding of other people's beliefs is critical for healthy social functioning, with impairments in theory of mind understanding being characteristic of disorders such as autism spectrum disorder or schizophrenia (Baron-Cohen, 1997; Sprong et al., 2007). Longitudinal evidence suggests that the understanding of other people's intentional actions during infancy predicts later theory of mind development (Aschersleben et al., 2008; Sodian et al., 2016; Wellman et al., 2008). The study of how intentional understanding typically develops during infancy is a new and emerging area of research yet is important for contributing to our knowledge of the origins of theory of mind.

A fundamental theory known as the active experience or 'humans-first' account suggests that in early development, infants begin to understand others' actions as they begin to produce actions themselves. Infants struggle to understand and predict actions that they have never performed themselves (e.g., Cannon & Woodward, 2012; Krogh-Jespersen & Woodward, 2018; Woodward, 1998). However, adults, older children, and in some cases even infants are eventually able to understand actions that they cannot self-produce (e.g., Gerson et al., 2017; Hunnius & Bekkering, 2010). To some extent, the all agents or cue-based account of action understanding can explain how an individual is able to understand an intentional action that they have never produced. This theory suggests that, if cues to intention are present, infants can understand the intention of any human or non-human agent (e.g., Csibra et al., 1999; Gergely et al., 1995). As early as we can measure, both self-produced experience and cues to intention are important for intention understanding. For example, infants as young as 3-months of age can learn about action intentions through early motor interventions in which they are given first-hand experience manipulating objects (e.g., Gerson & Woodward, 2014b; Sommerville et al., 2005). At the same age, infants can interpret a self-moving box as intentional if it exhibits cues to agency (e.g., Luo, 2011; Luo & Baillargeon, 2005). Therefore, rather than viewing the all agents and cue-based theories as separate entities, it seems feasible to consider how both theories come together to explain the development of intentional action understanding. Comparison between one's own familiar actions and another's less familiar actions is one way in which the gap between these two theories could be bridged (e.g., Gerson & Woodward, 2012; Gerson & Woodward, 2014a). The first two empirical studies in this thesis attempted to shed light on the process by which comparison between familiar and less familiar actions can occur. Further to this, we also investigated how infant-directed action can facilitate action understanding by drawing infants' attention to the purposeful sub-steps of actions. Finally, we considered how already learned actions that are associated with certain objects and tools can inform an individual's prediction of how an intentional action will unfold.

In study 1, we investigated how observational comparison between a familiar hand-reach action and tool-reach actions (i.e., a cloth-reach and a magnet-reach) that share the same goal facilitate 12-month-olds' understanding of a never-before-seen action with a different goal. This involved measuring infants' visual predictions of a complex claw-reach action that they cannot typically understand and predict on their own. This way, we could measure whether infants are able to generalise intentional understanding initially formed via comparison in order to understand and predict novel actions that they observe. In an eye-tracking paradigm, infants who had the opportunity to compare actions in which the same goal was reached for via different means (i.e., a hand-reach, cloth-reach, and magnet-reach) accurately predicted the outcome of the claw-reach action more often than infants in a control condition who observed the same tool reaching for different goals (i.e., differing colours of toy cars). However, this difference between conditions was marginally significant. Nevertheless, infants in the comparison condition systematically made accurate predictions that were significantly different from chance, whereas infants in the control condition performed at chance level. These results support our hypothesis that by comparing different ways of achieving the same goal, infants can generalise their intentional understanding to predict the outcome of a new action without knowing how to perform it themselves. Therefore, our first empirical study suggests that comparison is a cognitive factor that could bridge the gap between infants' understanding of actions that they can self-produce and more complex or unfamiliar actions.

Following on from this, study 2 investigated whether we would see similar results in a sample of younger 10-month-olds who may rely more-so on visual cues to intention than their own motor knowledge due to having less advanced motor skills. Younger infants tend to struggle more with tool-reach actions and were expected to be less efficient at performing and understanding them than 12-month-olds. Therefore, we hypothesised that we may see a similar pattern of results, but with a larger difference between comparison and control conditions in 10-month-old infants. Infants in study 2 underwent the same comparison and control training conditions as in study 1. However, we made small changes to this study, such that during observational training infants in the control condition viewed the same actions as those seen in the comparison condition, except they viewed the familiar hand-reach after observing the less familiar cloth- and magnet-reach actions. This was so that the hand-reach did not serve as an initial basis for comparison with the tool-reach actions, yet the two conditions were better matched in terms of perceptual variability. In addition to this, we attempted to gain lower attrition rates by manually progressing each trial upon a key press once the infant was attending to the screen. This seemed to work in our favour, as only 12.5% of the recruited sample for study 2 were excluded due to failure to make any type of prediction, compared with 34.4% in study 1. We also took variability in motor skill into account by measuring infants' ability to perform the same tool-use actions that they had seen during observational training. This involved the performance of a cloth-reach action (i.e., pulling on a cloth to draw close and retrieve a toy that was situated on the end of it) and magnet-reach action (i.e., connecting a magnet to a toy and drawing it close to retrieve it) before and after receiving motor training with the experimenter. Surprisingly, infants showed the reverse trend to that seen in study 1, such that infants in the control condition made a higher proportion of accurate predictions than those in the comparison condition. However, this difference between conditions was only marginally significant. In addition to this, infants in the control condition who were less efficient at performing actions in the motor task made a significantly higher proportion of accurate predictions than infants in the same condition who were more efficient in the motor task. These findings do not support previous literature that suggests a role for existing motor skill in facilitating understanding and prediction of other's actions (e.g., Gerson et al., 2015b; Gerson & Woodward, 2014b, 2014c; Sommerville et al., 2005, 2008). In addition to this, across both conditions, infants who viewed the cloth-reach first made more accurate visual predictions. Although we expected that infants would not be able to learn from our new control paradigm in study 2, it seems that infants were still able to learn from this intervention. We speculate that infants in the control condition were able to compare between tool-reach actions and the familiar hand-reach even when viewing the hand-reach last. Also, viewing the cloth-reach first may be an easier, more familiar, basis for comparison with the magnet-reach action than when viewing these tool-use actions in the opposite order. This finding was true regardless of whether infants had viewed the hand-reach action first or last and is also consistent with the fact that no infants were able to perform the magnet-reach action in the motor task. Overall, it seems that the control condition in study 2 may have been more beneficial for intentional learning than the comparison condition, especially for infants with low motor ability who rely more so on cues to compare than their own motor repertoires when making sense of observed intentional actions. To some extent this could support cue-based theory, such that infants can understand the intentional goals of actions if cues to intention are present (e.g., Csibra et al., 1999; Gergely et al., 1995; Luo, 2011), with comparable action structures (i.e., same goal) being the cue for intentional learning in this case. When bringing together our findings from studies 1 and 2, it is unclear as to what order of familiar and less familiar actions is most optimal for facilitating the development of intentional action understanding without further research. Nevertheless, our findings still suggest that observational comparison between actions helps to contribute to infants' broader understanding of intentional actions.

In line with both active experience and cue-based theory, our findings open the door for future work to further investigate the interplay between existing motor knowledge and cues to compare between observed actions. This could be theoretically interesting, as other work has identified an overlap between infants' own proficiency at self-producing actions and their ability to interpret cues to intention when observing novel actions (e.g., Brand et al., 2015; Gredebäck & Melinder, 2010). For example, Brand et al. (2015) found that infants as young as 6-months of age who had mastered crawling could interpret a circle's movements as goal-directed, whereas non-crawling infants did not. This stimulus involved adding agentive cues to a non-human object, such that a circle was self-propelled and took various routes (sometimes avoiding obstacles) to achieve the same goal (i.e., equifinality). Like the link between infants' own locomotive ability and their representation of a circle as an intentional agent, potentially due to its comparable ability to locomote, we expected that infants who were efficient at performing tool-use actions would extract intentional relations when comparing between tool-use actions. However, our findings in study 2 raise the question of whether motor ability has a varying effect on action understanding.

In study 3, the thesis shifts in focus to how social factors may help to facilitate infants' understanding of others' intentional actions. Specifically, we investigated whether mothers' use of infant-directed actions (i.e., motionese) are associated with infants' ability to accurately predict the outcome of a novel intentional action. Study 3 utilised the eye-tracking data that we had collected from study 2 along with behavioural data from mother-infant interaction sessions in which mothers demonstrated toys to their 10-month-old infants. We later coded demonstration videos on both a global scale (Brand et al., 2002) and a fine-grained scale of motionese (Brand et al., 2007). As study 3 did not have the same exclusion criteria as study

2 (i.e., did not require attention to training trials), a higher number of infants were retained in study 3 (N = 43) as compared with study 2 (N = 39). Our findings revealed that infants who were exposed to more frequent exchanges and less joint contact of toys within fine-grained motionese made more accurate visual predictions in eye-tracking test trials. It is likely that exchanges help to mark boundaries between the intentional sub-steps within actions and keep infants engaged with the activity (Brand et al., 2007). Although we did not expect that joint contact of toys would be negatively associated with accurate first predictions, we speculate that this finding may be due to mother's interference with their infant's exploration of the demonstrated toy, as opposed to assisting or scaffolding the infant's actions. Together, these findings suggest that more exchanges and less interference from mothers contributes to infants' broader understanding of never-before-seen actions. As mothers and infants engaged in play that mimicked how they might act at home, the findings presented in this thesis provide a snapshot of how mothers use infant-directed actions when interacting with their infant on a daily basis. Through experience with identifying intentional sub-steps of actions in infant-directed action contexts, our results suggest that infants can generalise such intentional understanding to predict the outcome a never-before-seen action. To our knowledge, this is the first study to provide evidence that infant-directed action contributes to infants' broader understanding of other people's intentional actions. By enabling infants to parse the intentional sub-steps of actions that they cannot necessarily produce, infant-directed action is another factor that could help facilitate infants' comparison between intentional actions that they can and cannot produce. In particular, exchanges could help to draw infants' attention to actions demonstrated before and after the exchange, leaving an interesting question as to what types of actions are most beneficial for infants' learning.

As mentioned above, for efficiency, studies 2 and 3 analysed eye-tracking data that were collected from the same sample of 10-month-old infants. These participants' data were also used as part of a larger study that utilised the same mother-infant interaction data to investigate a separate research topic (see chapter 4 for more information). Use of the same eye-tracking data across two studies may potentially raise concern over questionable research practices such as overlapping publication (van Raaij, 2018). Further to this, the studies in this thesis used small sample sizes, which are characteristic of infant research, and consequently have low statistical power that leads to increased risk of false positives and false negatives (Davis-Kean & Ellis, 2019; Oakes, 2017). After having initially collected data in order to test a particular hypothesis, re-analysis of the data to investigate separate research questions is sometimes considered suspicious. Such practice can also be referred to as 'repurposing' the data for a problem other than its original purpose (Wang et al., 2021). However, as part of a wide research project designed with multiple purposes, we had formulated separate hypotheses with a plan to collect data from the same sample during the initial stages of our experimental design. Thus, we argue that findings regarding separate avenues of investigation that utilise the same infant sample, both within and outside of the scope of this thesis, are complementary rather than overlapping. With distinct objectives explored in studies 2 and 3, we have drawn separate conclusions from the findings in each study, making contribution to separate theories. We also used similar methods of analysing the eye-tracking data across both studies, which is one means of avoiding potential problems with data reuse (van Raaij, 2018).

Finally, study 4 moved beyond goal-directedness to investigate how the meaningful, or semantic, features of objects and tools play a role in an individual's prediction of how an action will unfold. For example, food is meant to be eaten and a spoon is used with food. Dependent on a tool's conventional use, individuals also hold representations of specific actions that are associated with tools, such as how to grasp and use a spoon. Using an adult sample, we replicated an EEG paradigm previously used with 10-month-old infants (Ní Choisdealbha, 2015) to investigate whether the adult sensorimotor system shows a differential response to grasping actions that are carried out in a conventional versus unconventional manner. Participants watched videos in which an actor grasped a spoon, yet the orientation of the spoon differed, meaning that the semantics of the unfolding action were either congruent or incongruent with standard spoon use. To investigate mirrored neurocognitive representations within the sensorimotor system, we measured alpha and beta desynchronisation (i.e., sensorimotor activation) over central and occipital regions and found greater sensorimotor activation during incongruent actions at the stage that the spoon reached the mouth. There was no difference in sensorimotor activation between congruent and incongruent actions at the onset of the action, which involved the actor grasping the spoon in a manner that would lead to a conventional or unconventional eating action, respectively. However, increased sensorimotor activation in response to the part of the action that is the most motorically unusual supports the notion of prediction error (Kilner et al., 2007). This means that the unexpected action outcome (i.e., using a spoon in an unusual manner) results in greater sensorimotor activation due to the need to update the existing representation of the action and reduce prediction error. These findings were similar to those seen in infants who were more proficient at self-feeding (Ní Choisdealbha, 2015) and thus further verify that subtle semantic differences do not necessarily lead to differences in sensorimotor activation when making a prediction at action onset. Rather, the sensorimotor system responds to outcomes that are incongruent with the initial prediction in line with the predictive coding account of mirror system function. These findings support the active experience account such that knowledge of the motor commands required to perform familiar actions using certain tools and objects informs an individual's prediction of how another person will act.

Taken together, the findings in this thesis provide further insight into the way in which certain comparison, infant-directed action, and semantic knowledge contribute to the development of infants' and adults' intentional action understanding. Our results suggest that comparison is a cognitive process that helps to facilitate to the development intentional action understanding, however it is not yet clear exactly *how* this process occurs. It is also questionable as to what extent motor experience plays a varying role in combination with agentive cues during this process. For example, we do not yet know whether infants with more advanced motor skills are more sensitive to intentional cues, or whether infants with less motor experience rely on cues to intention as their only means of determining the purpose of an action. In addition to comparison, the development of intentional action understanding can be facilitated in social contexts in which an adult demonstrates an object in an exaggerated manner. This thesis provides the first evidence to suggest a link between infant-directed action and infants' ability to infer the intentional structure of novel actions. Within a social context, object exchanges between an adult seem particularly helpful for facilitating action understanding, likely by marking the beginning and end of an action as well as the sub-units within an action sequence (e.g., stacking a tower of rings). In later development, when an individual has gained experience and knowledge of actions associated with certain tools and objects (e.g., cutlery for use with food), the presence of certain objects can inform one's prediction of how a peer will act. During in-the-moment action prediction, the sensorimotor system adapts by minimising the error between predicted and observed actions, thus constantly informing intentional knowledge beyond infancy and into adulthood. The following sections of this chapter will discuss where our findings sit within the broader theoretical contexts of comparison, infant-directed action, and semantics.

The role of comparison in action understanding

This first two empirical studies presented in this thesis aimed to investigate how observational comparison between familiar and less familiar actions could facilitate infants' understanding of actions beyond those that they can already produce. For example, an infant may compare his or her own action of eating food by hand with an older sibling's action of eating food using cutlery. A new field of research has demonstrated that comparison can help infants to understand the intentional structure of a claw-reach action that they have previously compared with a familiar hand-reach (Gerson & Woodward, 2012, 2014a). We hypothesised that such experience comparing between different actions in which the same goal is achieved would also contribute to infants' broader ability to anticipate the goals of actions that they have never seen before. The first two empirical studies in this thesis provided infants with the opportunity to compare between a hand-reach action and tool-reach actions (i.e., a cloth-reach and magnet-reach) that shared the same goal. We then assessed whether infants in the comparison condition could generalise their knowledge from training to identify the intentional goal of a novel claw-reach action. Whilst previous comparison literature used goal imitation procedures as a measure of intentional goal recognition, we used an eye-tracking paradigm adapted from Cannon and Woodward (2012) to measure in-the-moment prediction of an unfolding tool-use action. This enabled us to precisely investigate when infants attend to certain features of the observed action and thus measure visual anticipations of the outcome of an incomplete action.

The comparison training condition in the current work was inspired from Gerson and Woodward's (2014a) experiment in which a matched goal label was used as a consistent cue to facilitate in-the-moment comparison between familiar and less familiar actions that occurred simultaneously. In the first two studies presented in this thesis, we presented infants with a series of multiple observed instances of actions that share the same goal object as a consistent cue to compare. That is, rather than comparing their own familiar hand-reach with a less familiar claw-reach action as in Gerson and Woodward's (2014a) study, infants observed another person's hand-reach action followed by two examples of tool-reach actions (i.e., a cloth-reach and a magnet-reach) that shared the same goal, and each example was repeated three times. Due to our marginal findings in study 1, followed by the reverse trend of what we expected in study 2, we can question whether additional cues to intention are needed for infants to successfully compare between the hand and tool-reach examples within our paradigm as well as more broadly. For example, adding matched goal labels to our comparison training condition might better promote comparison between actions (e.g., Gentner et al., 2011; Gerson & Woodward, 2014a; Loewenstein & Gentner, 2005), especially seeing as younger infants have a limited working memory and may rely more-so on cues to intention than older infants who can draw upon their more experienced motor repertoires when identifying intentional goals (e.g., Biro & Leslie, 2007; Brand et al., 2015). Infants may also find it easier to make comparisons between actions that are presented simultaneously (e.g., Christie & Gentner, 2010; Gentner et al., 2011; Gerson & Woodward, 2014a), rather than actions that are presented serially as in our paradigm. Also, clear action outcomes in which an object is brought to a target, such as bringing food to the mouth (e.g., Ní Choisdealbha, 2015) or placing a toy in a container (e.g., Falck-Ytter et al., 2006) using different tools may help to highlight the purpose of intentional actions. Whilst our stimuli simply involved a reach to obtain an object, other literature has involved subsequent actions such as lifting the object (e.g., Biro & Leslie, 2007) or moving an object from one location to another (e.g., Király et al., 2003; Jovanovic et al., 2007). In our paradigm, the addition of a consistent goal label (e.g., "car") or a consistent sound effect (e.g., Gerson & Woodward, 2014a; Kim et al., 2015) may be particularly helpful for infants to infer the same relational goal structure between the hand

and tool-use actions that are serially presented. In addition to this, an action outcome could be added (e.g., placing the object in a container) to give the grasping action an added purpose. This way, infants would not only have a single consistent cue for comparison, but three matched cues, being a consistent goal, label, and outcome. This raises interesting questions for future work to investigate what types of cues to intention best facilitate infants' ability to successfully compare between actions that they can and cannot produce.

In considering potential limitations of our findings from studies 1 and 2, it is clear that the generalisation of intention understanding from comparison training to a novel claw-reach action is a complex task for infants to achieve. After comparing a familiar hand reach action with cloth- and magnet-reach actions, generalisation to the claw-reach is arguably too cognitively demanding; especially for 10-month-old infants. That is, the claw-reach requires a squeeze of the handle in order to manipulate the pincer end and grasp the object, as opposed to simply reaching and drawing a goal-object close using a simple two-step action. Therefore, it is questionable whether infants might make more accurate predictions if we were to replace the claw with a less complex tool, like a cane (Sommerville et al., 2008). One way to assess whether infants attempt to interpret the claw-reach in terms of its separate components (i.e., squeezing a handle to manipulate the pincer) rather than a simpler two-step action (i.e., grasping the claw to draw an object close) would be to assess looking behaviour during familiarisation trials. This way, we could investigate how infants process the complete action while it is in motion. For example, AOIs covering the handle of the claw, the pincer end, as well as the goal objects could enable us to determine how much infants are attending to various parts of the visual scene during different events (e.g., grasping vs grabbing). This could help to inform us of how infants might anticipate the means of achieving a goal after the hand remains static during test trials.

Infants' ability to extract relational information during the process of observational comparison could also be assessed using neurophysiological methods, such as EEG. As discussed earlier in the thesis, a body of literature suggests that both observation and performance of the same action elicit a similar pattern of activation in the sensorimotor system (e.g., Fox et al., 2016; Gerson et al., 2017; Hari, 2006; Järveläinen, et al., 2004). In accordance with the active experience account, both developmental literature (e.g., Cannon et al., 2016; Gerson et al., 2015a; van Elk et al., 2008) and findings with adults (Aglioti et al., 2008; Calvo-Merino et al., 2006) suggest that such a mirrored sensorimotor response is sometimes stronger when an individual holds an established representation of an action. In the adult sensorimotor system, greater sensorimotor activation has been found when dancers observe a move that is within their own motor repertoire compared with a move that they have not performed (Calvo-Merino et al., 2008). Also, infants do not show increased sensorimotor activation during the observation of another infant crawling until they can crawl themselves (van Elk et al., 2010). Mirrored sensorimotor activation in response to reaching and grasping actions becomes stronger with experience (Cannon et al., 2016) and greater sensorimotor activation is present during ongoing prediction of an intentional hand-reach (i.e., a grasping posture) as opposed to a non-intentional hand-reach (i.e., a back-of-hand posture, Southgate et al., 2010). The observational training and eye-tracking paradigms presented in this thesis are well suited for the addition of neurophysiological measures. By measuring sensorimotor activation during a paradigm similar to that used in studies 1 and 2, we hypothesise that infants would show greater sensorimotor activation during the observation of actions that they know how to perform (i.e., a reach and grasp by hand). As well as this, greater sensorimotor activation during observation of the cloth and magnet-reach actions may be representative of the degree to which infants have successfully compared between these actions and the familiar hand-reach. Infants may also show greater sensorimotor activation during accurate prediction of the novel claw-reach action in test trials relative to infants who make inaccurate predictions. As individual differences in infants' ability to perform actions has been associated with greater sensorimotor activation (Cannon et al., 2016; Gerson et al., 2015a; van Elk et al., 2008), we also hypothesise that there would be a relation between sensorimotor activation and motor ability when making accurate predictions. The addition of a neurophysiological measure to our paradigm would enable us to shed light on the neurological process by which infants can build upon what they have previously produced to compare between, and predict the outcome of, actions that are outside of their own motor repertoires. This could help us to clarify the extent to which comparison facilitates action understanding above and beyond active experience alone. Such neurophysiological research could also resolve some of the open questions that the research in this thesis has raised. For example, sensorimotor activation is a sensitive measure that could help determine the order in which more or less familiar actions should be presented to infants to promote successful comparison; as well as the extent to which comparison is helpful for infants with varying levels of motor skill. We may see greater sensorimotor activation when an infant observes a more familiar action (e.g., a cloth-reach) followed by a less familiar action (e.g., a magnet-reach) than vice versa, depending on the optimal order for facilitating identification of the same intentional structure. This brings interesting questions for future research to provide evidence for a sensorimotor role within comparison.

An alternative way to interpret the findings presented in the first two empirical chapters of this thesis, especially the unexpected results of study 2, is that rather than learning about the goal of the claw-reach action infants are learning more about the means of using the tool to achieve the goal. Through a similar lens to that of habituation literature (Oakes, 2010), infants' attention may be more easily drawn to the aspect of a visual scene that is repeatedly changing, rather than aspects of the scene that remain the same. In other words, during our observational training paradigm, infants may attend more to the changing tool and thus may learn more about the different ways to achieve a goal rather than the importance of achieving the same goal in each example. The investigation of looking behaviour during familiarisation could also help to establish whether our training paradigm is unintentionally teaching infants more about the means of achieving goals. That is, if infants pay more attention to the claws rather than the goals during familiarisation, then this could imply that they have learned that the means of obtaining the goal is more important. This may be a more convincing argument to explain why infants in the comparison condition, as well as those with high motor ability, are making more location-based predictions in study 2.

Although studies presented in this thesis are cross-sectional, we also recognise the advantage that longitudinal methods could have in capturing the emergence of intentional action understanding within the first year of life (e.g., Brandone et al., 2019). Longitudinal research could help us to establish whether factors such as comparison, infant-directed action, and semantic knowledge play a causal role in facilitating the development of infants' action understanding. As infants' intentional understanding naturally becomes more advanced, we can also assess the way in which factors such as comparison, motor skill, or infant-directed action help promote this understanding independently as well as bi-directionally. In early development, in-the-moment comparison between a familiar hand-reach and a less familiar tool-reach may be most beneficial for facilitating infants' sensitivity to intentional goal structures, particularly when engaging in reaching actions themselves. However, in more difficult scenarios in which infants observe someone else perform goal-directed reaching actions, the use of consistent cues (such as labels) can help infants to make sense of the intentional relations involved in the action (e.g., Gerson & Woodward, 2014a). As motor skill and working memory develops with age, infants could go beyond in-the-moment comparisons to serial comparisons, and eventually develop the ability to compare between actions experienced across time.

The relation between infant-directed action and infants' action understanding

Further to the role of comparison in the development of intentional understanding, we hypothesised that another factor that could facilitate action understanding beyond active experience alone is infant-directed action. Study 3 investigated whether there is a relation between mothers' use of infant-directed action and infants' understanding of actions that they cannot perform. For example, an infant may struggle to place a ring on a ring stack but with the help of an adults' demonstration can identify the goal of stacking all of the rings. Previous literature suggests that one purpose of infant-directed action is to help infants to identify intentional sub-units of actions (e.g., Brand et al., 2002, 2007, 2008). There is also evidence that infant-directed action facilitates infants' subsequent success at producing a demonstrated action as well as their memory of the action. In addition to this, adults modulate their action demonstrations dependent on their evaluation of the infant's motor proficiency (van Schaik et al., 2019). We questioned whether exposure to infant-directed action, and thus experience with identifying sub-units of various actions, may facilitate infants' broader ability to understand the intentional structure of novel actions. In study 3, we measured various aspects of mothers' infant-directed actions in a play session that captured how mothers naturally demonstrate toys to their infants on a daily basis. We hypothesised that more exposure to infant-directed action during toy demonstrations would be associated with more accurate predictions of the novel claw-reach action in our eye-tracking paradigm. Data for study 3 was collected from the same sample of infants as in study 2, except accurate prediction data was analysed in association with mothers' engagement in motionese during toy demonstration tasks. The findings in study 3 confirm that experience with infant-directed actions, particularly frequent exchanges of toys, are related to infants' visual prediction of the novel claw-reach action. Our findings imply that exchanges are beneficial for infants' intentional learning because they help to mark boundaries between intentional sub-steps within actions (Brand et al., 2007). However, our research did not specifically investigate what types of actions were performed before and after exchanges. For example, we may question whether a parent passes a cup to an infant before (or after) stacking it on a tower. Thus, an exchange may act as a means of drawing the infant's attention to the action that was about to be (or had just been) performed. Alternatively, exchanges may act as a form of scaffolding after an infant's attempt at performing an action, such as nesting or stacking a cup. Based on our findings, future research could more specifically investigate how exchanges are helpful for infants' learning from collaborative activities. Further to this, our findings demonstrated that less joint contact (i.e., shared manipulation) of toys between mothers and infants was associated with a higher number of accurate first predictions. This was somewhat unexpected, as we hypothesised that joint action with toys would help to facilitate intentional understanding. Beyond the specifics of exaggerated infant-directed actions, our findings may tap into broader aspects of caregiver support. As discussed in chapter 4, joint contact may have been characteristic of mothers' interference with their infant's manipulation of the toy. Therefore, our findings imply that beyond motionese, intrusiveness versus responsiveness to infants' exploration of objects is important to distinguish when defining what type of caregiver input contributes to the development of intentional action understanding (e.g., Grolnick et al., 1984, 2002; Whipple et al., 2011). An interaction may score highly in terms of motionese but may actually be highly intrusive. For example, a caregiver may dictate an entire interaction with exaggerated or dramatic object demonstrations but may not be providing the type of input that the infant can learn from. Therefore, future research could further investigate what types of input are most beneficial for infants' intentional learning. Our intuition is that collaborative support within shared actions, such as following an infant's lead or scaffolding an infant's handling of an object may be most beneficial for facilitating learning.

To establish a relation between infant-directed action and action understanding in the current research, we based our measures of motionese off original behavioural coding schemes that were the first to be developed for measuring infant-directed action (Brand et al., 2002, 2007). However, more recent advances in technology offer new ways to quantify this phenomenon by measuring features of movement involved in actions. For example, van Schaik et al. (2019) used optical motion tracking to measure modulations of infant-directed actions in terms of 3D distance covered, velocity, proximity, and repetitions. Motion tracking usually involves the use of markers or sensors that are placed on the demonstrator's wrist and index finger (or any other part of the body that is necessary, van Schaik & Dominici, 2020) to track their movements. New software could also enable us to apply posthoc motion tracking to video data without the need for markers (Mathis et al., 2018; van Schaik & Dominici, 2020). Arguably, the coding of proximity and repetition using motion tracking (van Schaik et al., 2019) provide similar information to that quantified by human raters when coding motionese (Brand et al., 2002, 2007). However, the measurement of these features using motion tracking may be a more accurate and efficient process which does not require the need to double code a subset of the data for reliability. There may be some level of overlap between these methods in the measurement of repetition. However, repeated motion is arguably different to repeated action types and a human rater may be needed to differentiate these. For example, repeatedly stacking a cup could involve the same or a different movement each time a cup is added to the tower. Nevertheless, features of movement that can be measured using motion tracking provide us with additional information about the motion involved in infant-directed action at a more intricate level than is possible with human raters. Further to the dimensions of motionese measured using human raters, motion tracking allows researchers to investigate different features of infant-directed action that may relate to infants' action understanding and prediction.

The role of semantic knowledge in action prediction

In the final empirical chapter of this thesis, we investigated the extent to which the sensorimotor system is involved in the prediction of ongoing actions. In line with the active experience account of action understanding and prediction, we were particularly interested in whether the sensorimotor system differentiates between known representations of tool-use actions that are carried out in a conventional way (i.e., semantically congruent with existing motor representations) and those that are carried out in an unconventional way (i.e., semantically incongruent with existing motor representations). As outlined earlier in this chapter, we found that the adult sensorimotor system is not differentially activated at the onset of incongruent versus congruent actions but responds to the part of the action that is most motorically unusual. Study 4 was a direct replication of Ní Choisdealbha's (2015) paradigm that had originally been used to investigate the extent to which the infant sensorimotor system responds to typical and atypical spoon use actions. With a sample of 10-month-old infants, Ní Choisdealbha (2015) found no effect of action congruence on sensorimotor activation. Therefore, we used the same paradigm to investigate whether the experienced sensorimotor system responds differently in an adult sample. Our findings with experienced adults are similar to those seen only in infants who were proficient at using a spoon to eat (Ní Choisdealbha, 2015) and are also in line with the predictive coding account of mirror system function (Kilner et al., 2007).

Despite the significant difference between conditions, one potential limitation with this study is that over repeated trials, participants may have learned that the actor always achieves their goal of eating from the spoon regardless of whether the means to achieve this is carried out in a conventional or unconventional manner. Across trials, participants could have learned about the new semantic context, meaning that eating from a spoon in an unconventional manner became more familiar or 'normal' to participants. This idea is in line with the way in which participants in Gerson et al.'s (2017) study learned about a new semantic context (e.g., *this* person uses pliers to eat) that resulted in changes in sensorimotor activity over time. That is, sensorimotor beta activity reduced (i.e., beta synchronisation) as the new semantic context became less unusual through learning. Therefore, investigating potential changes in sensorimotor activity due to learning throughout trials is an important direction for future research.

In future work, the role of semantics in action prediction could also be investigated using eye-tracking with a developmental or adult sample. Based on the work presented in this thesis as well as previous literature supporting the active experience account (e.g., Cannon & Woodward, 2012; Gredebäck et al., 2009; Krogh-Jespersen & Woodward, 2018), we hypothesise that as infants become more familiar with both producing and recognising a learned tool-use action, such as eating from a spoon, they will better understand the goals of the same action being performed by someone else (but see Hunnius & Bekkering, 2010; Southgate et al., 2008 for counterevidence). Previous literature has assessed infants' ability to grasp and use tools, such as a spoon, in a planful manner (e.g., McCarty et al., 1999, 2001; Ní Choisdealbha, 2015). We might expect infants (particularly those who are planful at using tools) to look longer to an action outcome that is motorically or semantically incongruent with the predicted outcome (Daum et al., 2009). Increased sensorimotor activation may also be associated with longer looking due to prediction error (Kaduk et al., 2016; Kilner et al., 2007;

Stapel et al., 2010). Using an adult sample, research could also investigate whether changing an individual's semantic representation of an action that was previously unusual (e.g., using pliers to eat in Gerson et al., 2017) may influence visual predictions of an action outcome. In other words, learning a new way of manipulating tools or objects may change an individual's expectation of the correct action to be performed and thus their speed to predict the outcome of the action.

The potential for a combined role of comparison, motionese, and semantics in the development of action understanding

The studies in this thesis have investigated the roles of comparison, motionese and semantics as separate factors that complement active experience in the development of action understanding. However, it is likely that these three factors, amongst others, play a combined role in facilitating the development of intentional action understanding across different social contexts. Social interactions between infants and adults or older children provide ideal situations in which cues for comparison are likely present (Gerson & Woodward, 2012, 2014a). For example, an infant might compare their own action of hitting a drum by hand with an adult's action of hitting the drum using sticks. In an interplay between comparison and infant-directed action, the adult may also exaggerate their action of hitting the drum with sticks in an infant-directed manner, which may involve lifting the sticks up high before hitting the drum. Tying in with the classic Vygotskian theory of scaffolding (Vygotsky, 1978; Wood et al., 1976), the adult might then pass the drumsticks to the infant (i.e., an object exchange), so that the infant can attempt to perform the demonstrated action (Brand et al., 2002, 2007, 2008; Koterba & Iverson, 2009; van Schaik et al., 2019). Much like the way that comparison can support infants' learning about new intentional actions, infant-directed action could also help infants learn about actions that they are not yet capable of performing. For example, an adult may help an infant to identify the individual sub-units involved in completing an overarching action, such as putting jigsaw pieces together to complete a puzzle. A young infant may not be capable of fitting the pieces together, but with the help of the adult's demonstration, can understand the goal of completing the puzzle. This idea is supported by van Schaik et al.'s (2019) finding that mothers change their use of motionese depending on their evaluation of their infant's motor skill.

In addition to comparison and infant-directed action, the information present from objects in a scene (i.e., semantics) can also inform an infant's expectations about the types of actions to be performed (Creem & Profitt, 2001; Ní Choisdealbha, 2015; van Elk et al., 2010b). For example, if an infant understands that jigsaw pieces are meant to be fitted together and observes an older child tip jigsaw pieces out of a box, the infant will likely anticipate that the child will begin fitting the pieces together to complete the puzzle. To give another example, an infant may know that a toothbrush is for cleaning teeth and that toothpaste is meant to be applied to the toothbrush. Therefore, when observing an adult grasp a tube of toothpaste, the infant may anticipate that the individual will apply the toothpaste to a toothbrush then bring the toothbrush to their mouth.

Using some of the latest methodological advancements, research could begin to investigate the way in which motionese, semantics, and comparison play an interrelated role in facilitating infants' prediction of ongoing actions. For example, motion tracking could be used in combination with head-mounted eye-tracking to measure infants' real-time prediction of actions when interacting with their parents, as in a recent study by Monroy et al. (2021). Contexts that create ideal opportunity for comparison between the infant's and the adult's actions could also be identified and coded within interactions. For example, an infant may tap a drum by hand whilst the adult taps the drum with drumsticks, either at the same time or successively. Further to this, actions associated with certain objects will likely inform the infant's real-time prediction of how the parent will act. For example, when provided with a telephone, the infant will likely visually anticipate that their parent will bring the phone to their ear (e.g., Hunnius & Bekkering, 2010). In contrast, predictive anticipations may not occur when the parent is demonstrating an unfamiliar object. By defining and coding the occurrence of social contexts for comparison, using objects that are associated with certain actions, and measuring motionese within naturalistic interactions, future studies can begin to answer the question of whether these three areas play a combined role in facilitating action understanding.

Other factors that contribute to action understanding

Beyond the specifics of motionese, caregiver input can also contribute to infants' intentional learning more broadly. With an emphasis on motor experience within the action understanding literature (e.g., Gerson & Woodward, 2014b, c; Krogh-Jespersen & Woodward, 2018; Sommerville et al., 2005; Woodward, 1998), there has been little focus on the social context in which intentional learning takes place. As mentioned above, social interactions provide ideal contexts for comparison to occur. In addition to comparison, parent-infant interaction can also provide opportunity for infants to detect statistical regularities within actions (i.e., statistical learning), supporting actions with speech narrations (i.e., acoustic packaging), as well as the adult and infant's shared attention upon an object of mutual interest (i.e., triadic engagement).

Statistical learning

Statistical learning refers to infants' ability to detect regularities in their environment and has been established within the visual domain (e.g., Fiser & Aslin, 2002), within language (e.g., Romberg & Saffran, 2010), and in action (e.g., Monroy, et al., 2017a, 2017b). As mentioned in the first section of this chapter, statistical regularities allow infants to learn about the structure of continuous natural actions, such as the order in which they tend to follow one another. For example, bathing an infant involves a series of actions including filling the bath with water, undressing the infant, placing the infant in the bath, washing the infant with soap, and rinsing off the soap. When familiar with a process that results in an overarching action, infants can predict the individual sub-steps of the action (Monroy et al., 2017a). By identifying units of a continuous action, infants can establish representations of actions that facilitate their understanding of other people's intentional action sequences (Baldwin et al., 2001). Infants may learn to detect statistical regularity in actions that they cannot necessarily produce through their experience engaging in collaborative social activities with caregivers. The more interactive input that an infant receives from a caregiver may feed into better intention understanding. Variability in the types of actions that infants are exposed to within social contexts could also lead to variation in statistical learning, such as how easy an action sequence is to interpret or parse, as well as how successfully the action sequence attracts infants' attention.

Statistical learning literature originally used looking time paradigms in which infants exhibited longer looking to sequences that differ from an expected structure. For example, in a study by Baldwin et al. (2001), 10- to 11-month-olds were familiarised to natural actions, such as picking up a towel off the floor and hanging it on a rack. In test trials, infants watched interrupted videos in which unnatural pauses were inserted during the agent's motion, and intact videos in which pauses occurred at natural points within the action sequence. Infants looked longer to the interrupted videos, implying that these violated infants' expectations of the typical structure of the action. When watching continuous streams of natural actions during everyday experience, infants learn where typical boundaries or pauses lie and are thus surprised when pauses occur at unnatural points.

Whereas Baldwin et al. (2001) focused on the process of segmentation in facilitating infants' understanding of action structures, subsequent research (Monroy et al., 2017b; Stahl et al., 2014) speaks to infants' ability to recognise the order of an action sequence by tracking statistical regularities (e.g., the probability of an action occurring). In Stahl et al. (2014), 7- to 9-month-olds observed an animated starfish perform a series of whole-body actions, such as 'clap', 'turn', 'twist', and many more. Infants detected violations within the action structure by relying on statistical regularities. For example, if the starfish usually performed 'bow', 'jumping jack', 'scrunch' in this sequential order, infants looked longer when observing an unusual sequence (e.g., 'bow', 'clap', 'scrunch'). Whereas Stahl et al.'s (2014) looking-time measure of infants' statistical learning was reactive, Monroy et al. (2017b) investigated 19-month-olds' prediction of an unfolding action sequence using eye-tracking. Toddlers in a human agent condition observed a person's hand act on a multi-object toy that afforded different actions, such as spinning, pushing, squeezing, and so on. Toddlers in a ghost condition observed these objects move on their own with a spotlight that illuminated each event as it occurred. Toddlers in the human agent condition made accurate predictive looks to upcoming actions based on statistical regularities, whereas toddlers in the ghost condition did not. Accurate predictions in the human agent condition were also related to toddlers' ability to reproduce the action sequence, supporting the active experience, or humans-first, account. These findings suggest that statistical learning is an observational factor by which infants can come to understand action sequences.

To further understand the results of Monroy et al. (2017b), Monroy and colleagues subsequently tested whether 7- to 11-month-olds' ability to produce actions interacts with their ability to learn statistical regularities within an action sequence (Monroy et al., 2017a). This study thus considered how cognitive and motor factors may be integrated. Infants observed a sequence involving two action pairs, whilst the rest of the sequence was random. One pair was performed using a whole-hand grasp (e.g., 'bend' followed by 'push'), whilst the other used a pincer grasp (e.g., 'slide' followed by 'open'). Infants were separated into two groups according to which of the grasps they could efficiently perform. Across both groups, infants made more accurate than inaccurate predictive looks to upcoming actions. Although statistical learning did not differ dependent on infants' grasping ability, infants who were more efficient at producing a pincer grasp made a quicker increase in accurate predictions for the pincer grasp pair, rather than the whole-hand pair. In contrast, infants who were more efficient at a whole-hand grasp showed the reverse effect in their pattern of predictions. These findings provide support for the role of active experience in facilitating learning of statistical regularities, as infants could better predict actions that were the most motorically familiar. However, active experience is not a critical precursor for action prediction, as infants could still predict actions that they were not familiar with, just at a slower rate. In sum, it seems that infants can use a combination of prior motor knowledge and statistical learning to predict a sequence of familiar actions, whereas they rely more so on statistical learning to make sense of unfamiliar actions.

As described earlier in this chapter, the observation of familiar actions is coupled with mirrored activation within the sensorimotor system (e.g., Cannon et al., 2016). Monroy et al. (2019) provide neurophysiological evidence that knowledge from statistical learning can be used to make predictions within the sensorimotor system. Over three days, 18-month-olds were trained with videos of novel action sequences involving statistical regularities that were predictable (deterministic) or unpredictable (random). At test, infants were shown a new sequence that featured the same statistical structure as the training phase. Greater sensorimotor activity preceded deterministic actions, not random actions. This means that infants were predicting specific actions based on statistical likelihood. This is the first evidence revealing that infants can gain new knowledge through statistical learning and use this knowledge to make action predictions within the sensorimotor system. Taken together, these findings suggest that, although statistical learning is an important factor for action understanding, it is also not entirely separate from motor factors. The infant sensorimotor system may use knowledge from comparison in a similar way. To our knowledge there is no neurophysiological evidence for a sensorimotor role within comparison, leaving an interesting open question for future research.

Acoustic packaging

An additional social factor that contributes to the development of action understanding is acoustic packaging. This relatively new area of research has identified that adults use speech narrations to facilitate infants' identification of important segments within an action sequence (Brand & Tapscott, 2007; Meyer et al., 2011). In other words, acoustic packaging is used in combination with demonstrated action to facilitate infants' learning about action structures. A narration before and after an action, as well as before and after sub-components, can help promote infants' understanding of both the completion of, and the process of achieving, goal-directed actions. For example, when demonstrating a ring stacker to an infant, a parent may begin the demonstration with a goal setting phrase, such as "here we go." As the parent places a ring on the stack, they couple this with a description, such as "put the red one on." The next component is coupled with another description, "put the yellow one on," and so on. Upon completion of the ring stack (the ultimate goal) a final phrase is used, such as "we did it!" (Meyer et al. 2011). Whilst some action units may be coupled with a narration, others are not. The presence of narration facilitates infants' identification of actions that occur together. Brand and Tapscott (2007) demonstrated infants' identification of packaged and non-packaged actions, and thus action boundaries, by familiarising 7.5- to 11.5-month-olds with video sequences of a woman performing a task with a bottle (e.g., looking at the bottle, poking finger in the bottle, tilting the bottle). Two of three units of these sequences were coupled with a narration, so that pairs of action units could be packaged together. For example, during the looking and poking units, the infants heard "Wow! Do you see what she's doing? She's blixing!" Infants then simultaneously viewed the pair that had been packaged and a non-packaged pair of videos side-by-side in test trials. Infants older than 9.5-months showed a novelty preference for non-packaged pairs, indicating that they had processed the packaged pair as a unit and could discriminate between the pairs of actions. This is similar to the way in which consistent verbal labels can facilitate comparison between actions. For example, in Gerson and Woodward's (2014a) study, labelling a goal object prompted 10-month-olds to view two different actions as alike and thus make accurate goal imitations.

Meyer et al. (2011) also investigated natural variation (as opposed to experimental manipulation) in the way that mothers use speech to highlight action boundaries when demonstrating toys to younger (6- to 6.5-month-old) and older (9.5- to 13-month-old) infants. They investigated the degree to which action and speech onsets and offsets are temporally aligned, as well as the extent to which action descriptions are used in comparison with non-action descriptions. Across both age groups, the authors found that mothers pair speech with onsets and offsets of action units, rather than speaking more generally when performing actions. Specifically, individual units of actions (e.g., nesting a cup inside another) were temporally closer to onsets of action descriptions (e.g., "blue goes inside") than non-action descriptions (e.g., "look!").

From this research, it seems feasible that infant-directed speech and action work in parallel as a multimodal input to help infants identify segments within an action stream. Acoustic packaging could thus help facilitate infants' action understanding. Through experience with segmenting continuous actions via acoustic packaging, infants may be able to generalise their knowledge of action structures to understand novel actions. This could promote understanding and prediction of ongoing actions.

Triadic engagement

Another social factor that not only captures parents' behaviour, but additionally examines the joint role of infants' social behaviour with the parent, is triadic engagement. Between 9- and 12-months of age, infants become more efficient and organised within their triadic interactions, in which their own and another's attention is focused on an object of mutual interest. This is also referred to as joint engagement or joint attention, in which attention is shared but not always directed at a particular object (Brandone, 2015; Brandone et al., 2019; Carpenter et al., 1998). At this age, infants spend more time in triadic engagement than in dyadic face-to-face play (Bakeman & Adamson, 1984; Brandone, 2015; Brandone et al., 2019). This coincides with a more robust understanding of intentional action knowledge (e.g., Cannon & Woodward, 2012). When infants understand that other people are intentional agents, forms of shared attention can arise when infants interact socially with others. Such interactions might involve building a tower out of blocks, putting toys away, pointing and naming, playing pretend games like eating and drinking, and so on (Tomasello et al., 2005). Barresi and Moore (1996) describe how triadic interactions can help infants to develop a representation of their own and another's actions as intentional. When playing with an object of mutual interest, infants' and adults' actions and intentions are aligned with one another. For example, when building a block tower together, an infant and adult will likely recognise that they are both attending to the same thing. In relation to my discussion of comparisons above, this enables infants to actively compare their own actions with their interactive partner's, contributing to infants' shared representations of their own and another's intentions (e.g., Gerson & Woodward, 2012, 2014a). Therefore, triadic engagement creates a social environment that enables motor and cognitive factors to interact.

Infants less than a year old are more passive within their triadic interactions compared with 12- to 15-month-olds, who tend to be more coordinated. For example, older infants are more likely to direct adult behaviour and help the adult in their role rather than follow the adult's lead (Bakeman & Adamson, 1984; Brune & Woodward, 2007). Recent research by Brandone et al. (2019) provides evidence that triadic engagement promotes young infants' understanding of intentional actions. In the first study, 8- to 9-month-olds' action understanding was assessed via visual prediction towards a goal in an eye-tracking paradigm. Infants viewed a failed reaching action in which an agent reached over a barrier to retrieve a ball (the goal), but narrowly missed it. Infants also underwent a semi-structured play session with an experimenter. Infants' bouts of triadic engagement during the play session were coded as gaze alternations between a toy, the experimenter's face, and the same toy. Infants who engaged in more bouts of triadic, rather than dyadic, engagement made more predictive looks to the goal in the eye-tracking paradigm, meaning they better understood the intentional outcome of the reaching action. Brandone et al. (2019) then conducted a second cross-lagged longitudinal study to assess the directionality of the association between triadic engagement and intentional action understanding. 6- to 7-month-olds underwent the eye-tracking task as in the first study, followed by the play session. They then completed the same tasks again three months later. Tendency to engage in triadic, rather than dyadic, interaction at 6- to 7-months of age predicted intentional action understanding three months later. In contrast, action understanding at 6- to 7-months did not predict later triadic engagement, confirming that triadic interaction supports intentional action understanding, rather than the reverse. The authors concluded that this is likely due to infants' accumulated experience with triadic interaction.

In Brandone et al.'s (2019) study, an experimenter led semi-structured play sessions in which the experimenter remained neutral and only responded following infants' looks to her face. This way, the experimenter remained consistent in her interactive style for all infants. Whilst this only allowed for individual differences in infants' behaviour to drive any variation in accurate predictive looks, differences in parents' levels of interactiveness may provide varying opportunities for infants to engage in triadic interaction. This relates to individual differences in the extent to which parents use infant-directed actions, as discussed earlier in this chapter. Vaughan et al. (2003) support this idea, finding that parents' scaffolding behaviours (e.g., directing, showing, demonstrating) in joint engagement with their 9-month-olds were correlated with infants' initiations of joint attention at 12-months (e.g., initiating eye contact while manipulating a toy, pointing to an object etc.). Shared goals and attention are likely to occur within infant-directed action when parents are attracting infants' attention to an object of mutual interest. In other words, infant-directed action is likely present within contexts of triadic interaction and thus these are two factors that may work in parallel to support infants' action understanding. By coding parents' use of motionese as well as infants' bouts of triadic engagement, future research could assess how these factors may be related, or jointly contribute, to infants' action understanding. The extent to which each of these factors are present may affect the degree to which infants are presented with contexts that encourage comparison between their own and another's actions. Research has yet to explore such questions.

Conclusion

The aim of this thesis was to examine how comparison, motor skill, infant-directed action, and semantic knowledge facilitate the development of intentional understanding. We investigated how infants use comparison and motor skill to their advantage when interpreting actions within the first year of life; how exposure to infant-directed action can help infants to understand the intentional goals of new actions; and the extent to which the sensorimotor system is differentially activated dependent on the observation of actions that support or contradict existing representations. Our findings suggest that observational comparison between actions that infants can and cannot produce can support infants' understanding of novel actions by 12-months of age, and also potentially earlier at around 10-months of age. In a parent-infant interaction context, certain features of infant-directed actions can help infants to identify intentional sub-units of actions and generalise this ability to interpret the goals of new actions. After learning about conventional ways to manipulate tools and objects (e.g., the use of cutlery with food) through active experience, an individual can make predictions about actions to be performed based on the objects that are present. When the outcome of an action is unexpected, the representation of an intentional action is updated within the sensorimotor system, informing new semantic knowledge of actions.

The findings in this thesis provide information about factors that complement active experience to accelerate infants' learning of action understanding to varying extents, in varying contexts, and at different developmental stages in the first year of life. Comparison is one cognitive factor that is useful in contexts where infants can compare their own familiar actions with another person's, whether this is actively (e.g., Gerson & Woodward, 2012) or through observation alone (e.g., Gerson & Woodward, 2014a). Many infants experience adults' use of infant-directed action day-to-day, and our findings are the first to provide evidence that certain features of infant-directed action are associated with infants' more general understanding of other's intentional goals. Whilst the interpretation of novel actions is important for the development of intentional action understanding, learned knowledge of actions associated with certain tools and objects (i.e., semantics) also play a role in action prediction. Our findings confirm that the adult sensorimotor system is differentially modulated based on the observation of actions that are congruent or incongruent with an individual's expectation. These cognitive, social and semantic factors are part of a broader framework of multiple factors that contribute to the development of intentional action understanding. Such factors likely come into play during different contexts and at different stages of development alongside others, such as statistical learning (e.g., Monroy et al., 2019) or acoustic packaging (e.g., Meyer et al., 2011).

There are still many open questions as to how infants generalise beyond what they already know to understand new intentional actions in early development. For example, does the observation of a familiar action followed by less familiar actions best facilitate comparison or vice versa? Are some cues to compare between actions (e.g., matched goals, vocal cues, or subsequent actions) more effective than others? Our research has identified an association between motionese and broader intentional understanding, however we do not yet know exactly *how* motionese helps infants to understand the goals of novel actions. Could this be due to attentional modulations that occur alongside infant-directed action (e.g., triadic engagement)? Targeting such questions will help researchers gain a better understanding of how infants develop an understanding of actions that they have never performed. Cognitive and social facilitators of intention understanding likely continue to support more complex mentalising beyond the first year of life and throughout later development, contributing to children's understanding of other's mental states and theory of mind.

To conclude, this thesis has investigated the role of comparison, motor skill, infant-directed action, and learned knowledge of actions associated with objects in facilitating development of intentional action understanding. These factors collectively facilitate and advance infants' ability to understand and predict actions within the first year of life and beyond. Understanding other people's intentions is crucial for healthy social functioning, as it can help us to navigate collaborative tasks and activities in which we predict how another person will act, such as during team sports, cooking with a partner, or when driving a car. A lack of intentional understanding can lead to issues with theory of mind in later development, which can have negative implications for social functioning. As our knowledge of the origins of infant action understanding evolves through research, we can attempt to mitigate such issues.

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

Aschersleben, G., Hofer, T., & Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Developmental Science*, *11*(6), 862-868. https://doi.org/10.1111/j.1467-7687.2008.00736.x

Bach, P., Gunter, T. C., Knoblich, G., Prinz, W., & Friederici, A. D. (2009). N400-like negativities in action perception reflect the activation of two components of an action representation. *Social Neuroscience*, *4*(3), 212-232. https://doi.org/10.1080/17470910802362546

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development, 55,* 1278-1289. https://doi.org/10.2307/1129997

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

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods, 7*(2), 127-150. https://doi.org/10.1177/1094428104263672

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press. https://doi.org/10.7551/mitpress/4635.001.0001

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences, 19*(1), 107-122. https://doi.org/10.1017/S0140525X00041790

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

Brandone, A. C. (2015). Infants' social and motor experience and the emerging understanding of intentional actions. *Developmental Psychology, 51*(4), 512-523. http://doi.org/10.1037/a0038844

Brandone, A. C., Stout, W., & Moty, K. (2019). Triadic interactions support infants' emerging understanding of intentional actions. *Developmental Science, 23*(2)*,* e12880. https://doi.org/10.1111/desc.12880

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action. *Developmental Science, 5*(1), 72-83. https://doi.org/10.1111/1467-7687.00211

Brand, R. J., Escobar, K., Baranès, A., & Albu, A. (2015). Crawling predicts infants' understanding of agents' navigation of obstacles. *Infancy, 20*(4), 405-415. https://10.1111/infa.12084

Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science, 11*(6), 853-861. https://doi.org/10.1111/j.1467-7687.2008.00734.x

Brand, R. J., Shallcross, W. L., Sabatos, M. G., & Massie, K. P. (2007). Fine-grained analysis of motionese: eye-gaze, object exchanges, and action units in infant-versus adult-directed action. *Infancy, 11*(2), 203-214. https://doi.org/10.1111/j.1532-7078.2007.tb00223.x

Brand, R. J., & Tapscott, S. (2007). Acoustic packaging of action sequences by infants. *Infancy, 11*(3), 321-332. https://doi.org/10.1111/j.1532-7078.2007.tb00230.x

Brugger, A., Lariviere, L. A., Mumme, D. L., & Bushnell, E. W. (2007). Doing the right thing: Infants' selection of actions to imitate from observed event sequences. *Child Development, 78*(3), 806-824. https://doi.org/10.1111/j.1467-8624.2007.01034.x

Brune, C. W., & Woodward, A. L. (2007). Social cognition and social responsiveness in 10-month-old infants. *Journal of Cognition and Development, 8*(2), 133-158. https://doi.org/10.1080/15248370701202331

Calvo-Merino, B., 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., Simpson, E. A., Fox, N. A., Vanderwert, R. E., Woodward, A. L., & Ferrari, P. F. (2016). Relations between infants' emerging reach-grasp competence and event-related desynchronization in EEG. *Developmental Science, 19*(1), 50-62. https://doi.org/10.1111/desc.12295

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science, 15*(2), 292-298. https://doi.org/10.1111/j.1467-7687.2011.01127.x

Cannon, E. N., Yoo, K. H., Vanderwert, R. E., Ferrari, P. F., Woodward, A. L., & Fox, N. A. (2014). Action experience, more than observation, influences mu rhythm desynchronization. *Plos one*, *9*(3), e92002. https://doi.org/10.1371/journal.pone.0092002

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science, 1*(2), 380-392. https://doi.org/10.1111/j.1756-8765.2009.01026.x

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development,* i-174. https://doi.org/10.2307/1166214

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, *11*(3), 356-373. https://doi.org/10.1080/15248371003700015

Corbetta, D., Guan, Y., & Williams, J. L. (2012). Infant eye-tracking in the context of goal-directed actions. *Infancy*, *17*(1), 102-125. https://doi.org/10.1111/j.1532-7078.2011.00093.x

Creem, S. H., & Proffitt, D. R. (2001). Grasping objects by their handles: A necessary interaction between cognition and action. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 218-228. https://doi.org/10.1O37//0O96-1523.27.1.218

Cross, E. S., Liepelt, R., de C. Hamilton, A. F., Parkinson, J., Ramsey, R., Stadler, W., & Prinz, W. (2012). Robotic movement preferentially engages the action observation network. *Human Brain Mapping*, *33*(9), 2238-2254. https://doi.org/10.1002/hbm.21361

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., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition*, *72*(3), 237-267. https://doi.org/10.1016/S0010-0277(99)00039-6

Daum, M. M., Vuori, M. T., Prinz, W., & Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6-and 9-month-old infants. *Developmental Science*, *12*(6), 854-862. https://doi.org/10.1111/j.1467-7687.2009.00831.x

Davis-Kean, P.E., & Ellis, A. (2019). An overview of issues in infant and developmental research for the creation of robust and replicable science. *Infant Behavior and Development, 57,* 101339. https://doi.org/10.1016/j.infbeh.2019.101339

De Bellis, F., Ferrara, A., Errico, D., Panico, F., Sagliano, L., Conson, M., & Trojano, L. (2016). Observing functional actions affects semantic processing of tools: Evidence of a motor-to-semantic priming. *Experimental Brain Research*, *234*(1), 1-11. https://doi.org/10.1007/s00221-015-4432-4

Decroix, J., Roger, C., & Kalénine, S. (2020). Neural dynamics of grip and goal integration during the processing of others' actions with objects: An ERP study. *Scientific Reports*, *10*(1), 1-11. https://doi.org/10.1038/s41598-020-61963-7

de Moor, C. L., & Gerson, S. A. (2020). Getting a grip on early intention understanding: The role of motor, cognitive, and social factors. *Progress in Brain Research*, *254*, 113-140. https://doi.org/10.1016/bs.pbr.2020.05.005

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

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279-293. https://doi.org/10.1016/0163-6383(87)90017-8

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99*(24), 15822-15826. https://doi.org/10.1073/pnas.232472899

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., Ferrari, P. F., & van IJzendoorn, M. H. (2016). Assessing human mirror activity with EEG mu rhythm: A meta-analysis. *Psychological Bulletin*, *142*(3), 291-313. https://doi.org/10.1037/bul0000031

Fukuyama, H., Qin, S., Kanakogi, Y., Nagai, Y., Asada, M, & Myowa-Yamakoshi, M. (2015). Infant's action skill dynamically modulates parental action demonstration in the dyadic interaction. *Developmental Science, 18*(6), 1006-1013. https://doi.org/10.1111/desc.12270

Ganglmayer, K., Attig, M., Daum, M. M., & Paulus, M. (2019). Infants' perception of goal-directed actions: A multi-lab replication reveals that infants anticipate paths and not goals. *Infant Behavior and Development, 57,* 101340. https://doi.org/10.1016/j.infbeh.2019.101340

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development*, *82*(4), 1173-1188. https://doi.org/10.1111/j.1467-8624.2011.01599.x

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65*(2-3), 263-297. https://doi.org/10.1016/S0010-0277(98)00002-X

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287-292. https://doi.org/10.1016/S1364-6613(03)00128-1

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

Gerson, S. A., Bekkering, H., & Hunnius, S. (2015a). Short-term motor training, but not observational training, alters neurocognitive mechanisms of action processing in infancy. *Journal of Cognitive Neuroscience, 27*(6), 1207-1214. https://doi.org/10.1162/jocn_a_00774

Gerson, S. A., Mahajan, N., Sommerville, J. A., Matz, L., & Woodward, A. L. (2015b). Shifting goals: Effects of active and observational experience on infants' understanding of higher order goals. *Frontiers in Psychology, 6*, 310. https://doi.org/10.3389/fpsyg.2015.00310

Gerson, S. A., Meyer, M., Hunnius, S., & Bekkering, H. (2017). Unravelling the contributions of motor experience and conceptual knowledge in action perception: A training study. *Scientific Reports, 7*, 46761. https://doi.org/10.1038/srep46761

Gerson, S. A., & Woodward, A. L. (2012). A claw is like my hand: Comparison supports goal analysis in infants. *Cognition, 122*(2), 181-192. https://doi.org/10.1016/j.cognition.2011.10.014

Gerson, S. A., & Woodward, A. L. (2013). The goal trumps the means: Highlighting goals is more beneficial than highlighting means in means-end training. *Infancy, 18*(2), 289-302. https://doi.org/10.1111/j.1532-7078.2012.00112.x

Gerson, S. A., & Woodward, A. L. (2014a). Labels facilitate infants' comparison of action goals. *Journal of Cognition and Development, 15*(2), 197-212. https://doi.org/10.1080/15248372.2013.777842

Gerson, S. A., & Woodward, A. L. (2014b). Learning from their own actions: The unique effect of producing actions on infants' action understanding. *Child Development, 85*(1), 264-277. https://doi.org/10.1111/cdev.12115

Gerson, S. A., & Woodward, A. L. (2014c). The joint role of trained, untrained, and observed actions at the origins of goal recognition. *Infant Behavior and Development, 37*(1), 94-104. https://doi.org/10.1016/j.infbeh.2013.12.013

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

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., & von Hofsten, C. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Developmental Psychology, 45*(4), 1190-1194. https://doi.org/10.1037/a0015667

Grolnick, W., Frodi, A., & Bridges, L. (1984). Maternal control style and the mastery motivation of one-year-olds. *Infant Mental Health Journal, 5*(2), 72-82. https://doi.org/10.1002/1097-0355(198422)5:2<72::AID-IMHJ2280050203>3.0.CO;2-O

Grolnick, W. S., Gurland, S. T., DeCourcey, W., & Jacob, K. (2002). Antecedents and consequences of mothers' autonomy support: An experimental investigation. *Developmental Psychology, 38*(1), 143–155. https://doi.org/10.1037/0012-1649.38.1.143

Guajardo, J. J., & Woodward, A. L. (2004). Is agency skin deep? Surface attributes influence infants' sensitivity to goal-directed action. *Infancy, 6*(3), 361-384. https://doi.org/10.1207/s15327078in0603_3

Hari, R. (2006). Action–perception connection and the cortical mu rhythm. *Progress in Brain Research*, *159*, 253-260. https://doi.org/10.1016/S0079-6123(06)59017-X

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Sciences*, *95*(25), 15061-15065. https://doi.org/10.1073/pnas.95.25.15061

Hari, R., & Salmelin, R. (1997). Human cortical oscillations: a neuromagnetic view through the skull. *Trends in Neurosciences*, *20*(1), 44-49. https://doi.org/10.1016/S0166-2236(96)10065-5

Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behavior and Development, 28*(4), 466-480. https://doi.org/10.1016/j.infbeh.2005.04.002

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: A study of infants' visual anticipations during action observation. *Developmental Psychology, 46*(2), 446-454. https://doi.org/10.1037/a0016543

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1644), 20130490. https://doi.org/10.1098/rstb.2013.0490

Järveläinen, J., Schürmann, M., & Hari, R. (2004). Activation of the human primary motor cortex during observation of tool use. *Neuroimage, 23*(1), 187-192. https://doi.org/10.1016/j.neuroimage.2004.06.010

Jovanovic, B., Király, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2007). The role of effects for infants' perception of action goals. *Psychologia*, *50*(4), 273-290. https://doi.org/10.2117/psysoc.2007.273

Kaduk, K., Bakker, M., Juvrud, J., Gredebäck, G., Westermann, G., Lunn, J., & Reid, V. M. (2016). Semantic processing of actions at 9 months is linked to language proficiency at 9 and 18 months. *Journal of Experimental Child Psychology*, *151*, 96-108. https://doi.org/10.1016/j.jecp.2016.02.003

Kaur, M., Detherage, A., & Needham, A. W. (2020). Unconventional tool use in infants: Using a familiar tool in a novel way in the second year of life. *Cognitive Development*, *54*, 100881. https://doi.org/10.1016/j.cogdev.2020.100881

Kayhan, E., Heil, L., Kwisthout, J., van Rooij, I., Hunnius, S., & Bekkering, H. (2019). Young children integrate current observations, priors and agent information to predict others' actions. *PLoS One, 14*(5), e0200976. https://doi.org/10.1371/journal.pone.0200976

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

Kim, Z., Óturai, G., Király, I., & Knopf, M. (2015). The role of objects and effects in action imitation: Comparing the imitation of object-related actions vs. gestures in 18-month-old infants. *Infant Behavior and Development*, *41*, 43-51. https://doi.org/10.1016/j.infbeh.2015.07.002

Király, I., Jovanovic, B., Prinz, W., Aschersleben, G., & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness and Cognition*, *12*(4), 752-769.

Koelewijn, T., van Schie, H. T., Bekkering, H., Oostenveld, R., & Jensen, O. (2008). Motor-cortical beta oscillations are modulated by correctness of observed action. *Neuroimage*, *40*(2), 767-775. https://doi.org/10.1016/S1053-8100(03)00084-9

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development, 32*(4), 437-444. https://doi.org/10.1016/j.infbeh.2009.07.003

Krogh-Jespersen, S., & Woodward, A. L. (2014). Making smart social judgements takes time: Infants' recruitment of goal information when generating action predictions. *PloS one, 9*(5), e98085. https://doi.org/10.1371/journal.pone.0098085

Krogh-Jespersen, S., & Woodward, A. L. (2018). Reaching the goal: Active experience facilitates 8-month-old infants' prospective analysis of goal-based actions. *Journal of Experimental Child Psychology, 171,* 31-45. https://doi.org/10.1016/j.jecp.2018.01.014

Langeloh, M., Buttelmann, D., Matthes, D., Grassmann, S., Pauen, S., & Hoehl, S. (2018). Reduced mu power in response to unusual actions is context-dependent in 1-year-olds. *Frontiers in psychology*, *9*, 36. https://doi.org/10.3389/fpsyg.2018.00036

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development, 7*(2), 173-187. https://doi.org/10.1207/s15327647jcd0702_2

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition, 108*(3), 732-739. https://doi.org/10.1016/j.cognition.2008.06.013

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315-353. https://doi.org/10.1016/j.cogpsych.2004.09.004

Loucks, J., & Sommerville, J. A. (2012). The role of motor experience in understanding action function: The case of the precision grasp. *Child Development*, *83*(3), 801-809. https://doi.org/10.1111/j.1467-8624.2012.01735.x

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science, 14*(2), 453-460. https://doi.org/10.1111/j.1467-7687.2010.00995.x

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

Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation and execution in 14-month-old infants: An event-related EEG desynchronization study. *Developmental science*, *14*(3), 474-480. https://doi.org/10.1111/j.1467-7687.2010.00991.x

McCarty, M. E., Clifton, R. K., & Collard, R. R. (1999). Problem solving in infancy: The emergence of an action plan. *Developmental Psychology*, *35*(4), 1091-1101. https://doi.org/10.1037/0012-1649.35.4.1091

McCarty, M. E., Clifton, R. K., & Collard, R. R. (2001). The beginnings of tool use by infants and toddlers. *Infancy*, *2*(2), 233-256. https://doi.org/10.1207/S15327078IN0202_8

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental Science, 10*(1), 126-134. https://doi.org/10.1111/j.1467-7687.2007.00574.x

Ménoret, M., Bourguignon, M., & Hari, R. (2015). Modulation of rolandic beta-band oscillations during motor simulation of joint actions. *PloS one*, *10*(7), e0131655. https://doi.org/10.1371/journal.pone.0131655

Meyer, M., Hard, B., Brand, R. J., McGarvey, M., & Baldwin, D. A. (2011). Acoustic packaging: Maternal speech and action synchrony. *IEEE Transactions on Autonomous Mental Development, 3*(2), 154-162. https://doi.org/10.1109/TAMD.2010.2103941

Monroy, C., Chen, C. H., Houston, D., & Yu, C. (2021). Action prediction during real-time parent-infant interactions. *Developmental Science*, *24*(3), e13042. https://doi.org/10.1111/desc.13042

Monroy, C. D., Gerson, S. A., & Hunnius, S. (2017a). Infants' motor proficiency and statistical learning for actions. *Frontiers in Psychology, 8,* 2174. https://doi.org/10.3389/fpsyg.2017.02174

Monroy, C. D., Gerson, S. A., & Hunnius, S. (2017b). 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

Monroy, C. D., Meyer, M., Schröer, L., Gerson, S. A., & Hunnius, S. (2019). The infant motor system predicts actions based on visual statistical learning. *Neuroimage, 185,* 947-954. https://doi.org/10.1016/j.neuroimage.2017.12.016

Nagai, Y., & Rohlfing, K. J. (2008, August). Parental action modification highlighting the goal versus the means. In *2008 7th IEEE International Conference on Development and Learning* (pp. 1-6). IEEE. https://doi.org/10.1109/DEVLRN.2008.4640796

Ní Choisdealbha, Á. (2015). *Semantic and motor processes in infant perception of object-directed and tool-mediated action* [Doctoral thesis, Lancaster University]. Lancaster EPrints. https://eprints.lancs.ac.uk/id/eprint/79179/1/2016NiChoisdealbhaPhD.pdf

Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development, 11*(3), 255-268. https://doi.org/10.1080/15248371003699977

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy, 22*(4), 436-469. https://doi.org/10.1111/infa.12186

Open Science Framework (2019). *Findlay, C. L., & Gerson, S. A.* [pre-registration]. https://osf.io/cp9zj/

Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese tone and stress languages. *Infant Behavior and Development, 14*(4), 415-440. https://doi.org/10.1016/0163-6383(91)90031-M

Paulus, M., Hunnius, S., & Bekkering, H. (2011). Can 14- to 20-month-old children learn that a tool serves multiple purposes? A developmental study on children's action goal prediction. *Vision Research, 51*(8), 955-960. https://doi.org/10.1016/j.visres.2010.12.012

Proverbio, A. M. (2012). Tool perception suppresses 10–12 Hz μ rhythm of EEG over the somatosensory area. *Biological Psychology*, *91*(1), 1-7. https://doi.org/10.1016/j.biopsycho.2012.04.003

Quandt, L. C., Marshall, P. J., Shipley, T. F., Beilock, S. L., & Goldin-Meadow, S. (2012). Sensitivity of alpha and beta oscillations to sensorimotor characteristics of action: An EEG study of action production and gesture observation. *Neuropsychologia*, *50*(12), 2745-2751. https://doi.org/10.1016/j.neuropsychologia.2012.08.005

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: Evidence for semantic processing systems. *Developmental psychology*, *45*(3), 620-629. https://doi.org/10.1037/a0015209

Ritter, P., Moosmann, M., & Villringer, A. (2009). Rolandic alpha and beta EEG rhythms' strengths are inversely related to fMRI-BOLD signal in primary somatosensory and motor cortex. *Human Brain Mapping*, *30*(4), 1168-1187. https://doi.org/10.1002/hbm.20585

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 906-914. https://doi.org/10.1002/wcs.78

Ruff, H. A., & Rothbart, M. K. (2001). *Attention in early development: Themes and variations.* Oxford University Press.

Rutherford, M. D., & Przednowek, M. (2012). Fathers show modifications of infant-directed action similar to that of mothers. *Journal of Experimental Child Psychology, 111*(3), 367-378. https://doi.org/10.1016/j.jecp.2011.10.012

Saby, J. N., Marshall, P. J., & Meltzoff, A. N. (2012). Neural correlates of being imitated: An EEG study in preverbal infants. *Social Neuroscience*, *7*(6), 650-661. https://doi.org/10.1080/17470919.2012.691429

Santos, D. C., Gabbard, C., & Goncalves, V. M. (2001). Motor development during the first year: A comparative study. *The Journal of Genetic Psychology*, *162*(2), 143-153. https://doi.org/10.1080/00221320109597956

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, *21*(4), 237-249. https://doi.org/10.1016/j.tics.2017.01.012

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real–world events: An electrophysiological investigation. *Psychophysiology*, *40*(1), 160-164. https://doi.org/10.1111/1469-8986.00016

Snyder, A. C., & Foxe, J. J. (2010). Anticipatory attentional suppression of visual features indexed by oscillatory alpha-band power increases: A high-density electrical mapping study. *Journal of Neuroscience*, *30*(11), 4024-4032. https://doi.org/10.1523/JNEUROSCI.5684-09.2010

Sodian, B., Licata, M., Kristen-Antonow, S., Paulus, M., Killen, M., & Woodward, A. (2016). Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: A longitudinal investigation. *Child Development*, *87*(4), 1221-1232. https://doi.org/10.1111/cdev.12533

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), 1249-1256. https://doi.org/10.1037/a0012296

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition, 95*(1), 1-30. https://doi.org/10.1016/j.cognition.2003.12.004

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. https://doi.org/10.1016/j.cognition.2004.07.004

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., Karoui, I. E., & Csibra, G. (2010). Motor system activation reveals infants' on-line prediction of others' goals. *Psychological Science, 21*(3), 355-359. https://doi.org/10.1177/0956797610362058

Sprong, M., Schothorst, P., Vos, E., Hox, J., & Van Engeland, H. (2007). Theory of mind in schizophrenia: meta-analysis. *The British Journal of Psychiatry*, *191*(1), 5-13. https://doi.org/10.1192/bjp.bp.107.035899

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development, 85*(5), 1821-1826. https://doi.org/10.1111/cdev.12247

Stapel, J. C., Hunnius, S., Meyer, M., & Bekkering, H. (2016). Motor system contribution to action prediction: Temporal accuracy depends on motor experience. *Cognition*, *148*, 71-78. https://doi.org/10.1016/j.cognition.2015.12.007

Stapel, J. C., Hunnius, S., van Elk, M., & Bekkering, H. (2010). Motor activation during observation of unusual versus ordinary actions in infancy. *Social Neuroscience, 5*(5-6), 451-460. https://doi.org/10.1080/17470919.2010.490667

Tomasello, M., Carpenter, M., Call, J, Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675-691. https://doi.org/10.1017/S0140525X05000129

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, *31*(1), 155-165. https://doi.org/10.1016/S0896-6273(01)00337-3

van Elk, M., van Schie, H. T., & Bekkering, H. (2010a). The N400-concreteness effect reflects the retrieval of semantic information during the preparation of meaningful actions. *Biological Psychology*, *85*(1), 134-142. https://doi.org/10.1016/j.biopsycho.2010.06.004

van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage, 43*(4), 808-814. https://doi.org/10.1016/j.neuroimage.2008.07.057

van Elk, M., van Schie, H. T., van den Heuvel, R., & Bekkering, H. (2010b). Semantics in the motor system: motor-cortical beta oscillations reflect semantic knowledge of end-postures for object use. *Frontiers in Human Neuroscience*, *4*, 8. https://doi.org/10.3389/neuro.09.008.2010

van Raaij, E. M. (2018). Déjà lu: On the limits of data reuse across multiple publications. *Journal of Purchasing and Supply Management, 24*(3), 183-191. https://doi.org/10.1016/j.pursup.2018.06.002

van Schaik, J. E., & Dominici, N. (2020). Motion tracking in developmental research: Methods, considerations, and applications. *Progress in Brain Research*, *254*, 89-111. https://doi.org/10.1016/bs.pbr.2020.06.007

van Schaik, J. E., Meyer, M., van Ham, C. R., & Hunnius, S. (2019). Motion tracking of parents' infant- versus adult-directed actions reveals general and action-specific modulations. *Developmental Science, 23,* e12869. https://doi.org/10.1111/desc.12869

Vaughan, A., Mundy, P., Block, J., Burnette, C., Delgado, C., Gomez, Y., Meyer, J., Neal, A. R., & Pomares, Y. (2003). Child, caregiver, and temperament contributions to infant joint attention. *Infancy, 4*(4), 603-616. https://doi.org/10.1207/S15327078IN0404_11

Virji-Babul, N., Rose, A., Moiseeva, N., & Makan, N. (2012). Neural correlates of action understanding in infants: influence of motor experience. *Brain and Behavior*, *2*(3), 237-242. https://doi.org/10.1002/brb3.50

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wang, X., Duan, Q., & Liang, M. (2021). Understanding the process of data reuse: An extensive review. *Journal of the Association for Information Science and Technology, 72*(9), 1161-1182. https://doi.org/10.1002/asi.24483

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology, 44*(2), 618-623. https://doi.org/10.1037/0012-1649.44.2.618

Wellman, H. M., Philips, A. T., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science, 7*(3), 283-288. https://doi.org/10.1111/j.1467-7687.2004.00347.x

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*(3), 363-375. https://doi.org/10.1016/S0926-6410(01)00129-X

Whipple, N., Bernier, A., & Mageau, G. A. (2011). Broadening the study of infant security of attachment: Maternal autonomy-support in the context of infant exploration. *Social Development, 20*(1), 17-32. https://doi.org/10.1111/j.1467-9507.2010.00574.x

Willatts, P. (1999). Development of means-end behavior in young infants: Pulling a support to retrieve a distant object. *Developmental Psychology, 35*(3), 651-667. https://doi.org/10.1037/0012-1649.35.3.651

Williamson, R. A., & Brand, R. J. (2014). Child-directed action promotes 2-year-olds' imitation. *Journal of Experimental Child Psychology, 118,* 119-126. https://doi.org/10.1016/j.jecp.2013.08.005

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1-34. https://doi.org/10.1016/S0010-0277(98)00058-4

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

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science, 6*(3), 297-311. https://doi.org/10.1111/1467-7687.00286

Zeger, S. L., Liang, K. Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics,* 1049-1060. https://doi.org/10.2307/2531734

Appendix

Fine-grained motionese coding: Quantification of interactiveness and simplification

Coding instructions - steps for ELAN Coding:

All videos should be coded with the sound muted.

Initial setup:
Add the table view and mother view for the participant and synchronise the timings of the videos using Options -> Media Synchronisation Mode. If synchronisation is done after coding, timings will be wrong.

How to synchronise:

- Tick the 'player 1' box. Find the point that the LED light turns off at the beginning of the task. If the mother touches the toy before the light turns off, offsets should be set to this moment.
- Find the exact frame for the moment that the task begins. Press and hold the ctrl (cmd on mac) key and scroll through the video frames using the arrow keys.
- Repeat the same process with 'player 2' ticked.
- Tick 'use absolute offsets' then press the 'apply current offsets' button to save your synchronisation.
- Play the videos to make sure that they are both in sync. If you find that they are out of sync, adjust your offsets.

Go to Tier -> Change Tier Attributes and create a tier for:

- task onset/offset
- gaze
- mother possession
- infant possession
- joint contact
- exchanges

1. **First pass - code task onset/offset**. The task begins from the moment the mother touches the toy or the light turns off to signal the beginning of the task (usually 00:00:00 when you have synchronised the videos) until the LED light signals the end of the task.

2. **Second pass - code gaze bouts.**

A gaze segment begins when the mother looks directly at the face of the infant and ends as soon as she looks away.

Tips:

- This should be done using frame-by-frame coding. Press and hold the ctrl key (cmd on mac) and scroll through the video frames using the arrow keys.
- Try to look very closely at the mother's pupils. You can use either the table or mother view to do this.
- It is easier to enlarge and watch only one of the camera angles when coding gaze. To hide one of the camera views, go to View then uncheck the angle that you do

not want to view. If you need it back at any point (e.g., you lose view of the mother's eyes using the camera angle you have chosen), go back up to View and check it again.
- The table view is usually easy to look at, except for when the mum wears glasses or has loose hair around her face. This can restrict your view of the mother's eyes. Sometimes the mum will get up out of her seat, so you should use the mother view camera angle at these moments.
- The mother view camera is most useful when the back of the infant's head is also in the shot. This makes it easier to see when the mum is looking directly at the infant.

3. **Third pass - code mother possession.**
   A segment of mother possession begins whenever the mother is touching any piece of the toy and ends as soon as she takes her hand off it.

   Note: A single possession can still be coded if mother taps the toy (i.e., you do not have to code a segment each time mother's hand/finger meets the toy when tapping). The time that the mother spends tapping the toy is segmented from first to last tap.

4. **Fourth pass - code infant possession.**
   Just like the mother possession, a segment of infant possession begins whenever the infant is touching a piece of the toy and ends as soon as they take their hand off it.

   Note: A single possession can still be coded when the infant repeatedly hits the toy (a similar principle to when the mother taps the toy). The time that the infant spends hitting the toy is segmented from the first to the last hit.

5. **Fifth pass - code joint contact and exchanges.**
   Joint contact:
   - Find segments of mother and infant possession that overlap with one another, then check to see if they are both touching the same toy at the same time.
   - A segment begins as soon as both the mother and infant are touching the same toy and ends whenever one of them takes their hand off it.
   - Joint contact also counts when the mother and infant are touching either end of the assembled toy (e.g., shape cube, caterpillar, tower of cups, or rings).

   Note: A single segment is coded when one of them is holding the toy whilst the other is tapping or hitting it. This is a similar principle to how you coded tapping/hitting in the mother and infant possession coding.

   Exchanges:
   - An exchange segment is coded whenever the mother relinquishes a toy to the infant. The object must be directly transferred from the mother's hand to the infant's hand to count as an exchange. An exchange is **not** counted if the mum places the toy on the table, then the infant touches it.
   - An exchange segment begins when the infant places their hand on the toy and ends when it leaves the mother's hand.
   - Try not to get exchanges confused with joint contact. There should not be any overlap between exchange and joint contact segments.

6. **Sixth (and final) pass - identify action types.**
   - Go to Options -> annotation mode.
   - In the mother possession tier, annotate every segment for each action that the mother performs. If the mother performs an action that does not fit the action types listed in the table for the given toy, the action can be listed as 'other.'

- The mother can perform multiple actions within one segment of possession. E.g., stacks, unstacks, taps with finger.
- Repeated action types must **not** be counted within one segment. For example, 'stack, stack' is not possible. Only list the action once. The idea is to capture the number of unique action types that mothers demonstrate each time she is in possession of the toy, independent of the number of repetitions.

Potential action types that mothers may perform when demonstrating the stacking cups:

Action type	**Definition**
Stacks	Stacking a cup on top of another to make a tower of at least two cups.
Unstacks	Takes stacked cups apart. Knocking cups over does not count.
Nests cup(s) inside	Mother places smaller cup(s) inside a bigger cup. Rim of cups are facing upwards.
Nests cup(s) over	Mother places larger cup(s) over smaller cups, hiding the smaller one(s) underneath. Rim of cups are facing downwards.
Unnests	Takes nested cups apart (whether they are inside or over one another)
Pushes tower to infant	Mother pushes the tower of two or more stacked cups towards the infant.
Animates	This can involve moving the cups around to gain the infant's attention. E.g., moving cups around on the table, in the air, shaking, etc.
Peek-a-boo	E.g., hiding one cup inside another then revealing it.
Taps cups together	Tapping two cups together.
Taps cup(s) with finger(s)	Mother taps the cup(s) using her finger(s) or whole hand.
Taps cup(s) on table	Tapping the cups against the surface of the table.
Taps infant with cup	Mother touches the infant using a cup. May be the infants' hand, arm, face, etc.
Holds cups in place	This can involve steadying the tower or nest of cups, often when the infant is manipulating it.
Holds cup(s) up to show infant	Mother may hold cup(s) up high or in front of the infants' face.
Puts cup(s) in front of infant	This can involve placing/pushing cup(s) on the table in front of the infant. This does not count when mother pushes the stacked tower of cups to infant.
Arranging	This can involve turning cups over, lining cups up, changing their position on the table etc.

Potential action types that mothers may perform when demonstrating the caterpillar:

Action type	**Definition**
Clip	Clipping one piece of the caterpillar into another.
Unclip	Unclipping one piece of the caterpillar out of another.
Examine	Moving the toy around, looking at it, etc. E.g., figuring out how to clip the toy together.
Peek-a-boo	E.g., hiding the toy under the table then bringing it above.
Pull	E.g., when infant is holding the toy and the mother tries to pull a piece off or is using the caterpillar to play tug of war.
Shake	Mother shakes the toy. Particularly a single piece of it.
Wiggle	Mother 'wiggles' connected pieces of the caterpillar.
Spin	Some parts of the caterpillar have spinning pieces.
Taps pieces together	Tapping two pieces of the toy together.
Taps toy with finger	Mother taps the toy with using her finger(s) or whole hand.

Taps toy on table	Tapping the toy against the surface of the table.
Taps/pokes infant with toy	Mother touches the infant using a piece of the toy. May be the infants' hand, arm, face, etc.
Twist	The orange and blue tail of the caterpillar can be twisted.
Holds toy up to show infant	Holds single or connected piece(s) of the caterpillar up to show the infant. Might be held up high or in front of the infants' face.
Places toy in front of infant	Mother places/pushes single or connected piece(s) of the toy on the table in front of the infant.

Potential action types that mothers may perform when demonstrating the shape sorter:

Action type	**Definition**
Pushes cube to infant	Mother pushes the cube towards the infant, often so that the infant can manipulate the cube/attempt to put a shape through a hole.
Animates	This can involve moving the shapes around to gain the infants' attention. E.g., moving shapes around on the table, in the air, shaking, etc.
Taps shapes together	Tapping two shapes together.
Taps shape(s)/cube with finger(s)	Mother taps the shape(s) or cube using her finger(s) or whole hand.
Taps shape(s) on table	Tapping the shapes against the surface of the table.
Taps infant with shape	Mother touches the infant using a shape. May be the infants' hand, arm, face, etc.
Taps shape against cube	Mother taps the shape against the cube. Often next to the correct hole for the shape.
Holds shape(s) up to show infant	Mother may hold shape(s) up high or in front of the infants' face.
Arranging	This can involve turning shapes over, lining shapes up, changing their position on the table, and so on.
Puts lid on cube	Mother places the red lid on top of the box.
Takes lid off cube	Mother takes the red lid off the box.
Posts shape through hole	Mother posts a shape through a hole in the box or the lid. This still counts when the lid is flat against the table and the mother places a shape in one of the holes.
Tips shapes out of cube	Mother tips the shapes out of the box.
Holds/balances shape in hole	Mother holds or balances a shape in a hole of the box or lid. This is often so that the infant can push the shape through the hole.
Places lid over shape	Mother places the lid over a shape that's on the table. This is like a 'reverse posting' action.
Puts shape(s) inside cube	Mother puts the shape(s) straight into the box without posting it through a hole.

Potential action types that mothers may perform when demonstrating the stacking rings:

Action type	**Definition**
Stacks	Stacking a ring on top of another to make a tower of at least two rings.
Unstacks	Takes stacked rings apart. Knocking rings over does not count.

Pushes tower to infant	Mother pushes the tower of two or more stacked rings towards the infant.
Animates	This can involve moving the rings around to gain the infants’ attention. E.g., moving rings around on the table, in the air, shaking, etc.
Taps rings together	Tapping two rings together.
Taps ring(s) with finger(s)	Mother taps the ring(s) using her finger(s) or whole hand.
Taps ring(s) on table	Tapping the rings against the surface of the table.
Taps infant with ring	Mother touches the infant using a ring. May be the infant’s hand, arm, face, etc.
Holds rings in place	This can involve steadying the tower of rings, often when the infant is manipulating it.
Holds ring(s) up to show infant	Mother may hold ring(s) up high or in front of the infant’s face.
Puts ring(s) in front of infant	This can involve placing/pushing ring(s) on the table in front of the infant. This does not count when mother pushes the stacked tower of rings to infant.
Arranging	This can involve turning rings over, lining rings up, changing their position on the table etc.
Rolls	This involves rolling the ball-shaped piece. E.g., rolling the ball towards the infant.