Goal Tracking and Prosocial Behaviour in Early Childhood

By

Alexander Patrick Green

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psychology

Department of Psychology

University of Warwick

February 2021

Contents

List of Tables

List of Figures

Acknowledgements

Although my name is on the front cover, this thesis has been built on the foundations provided by a multitude of individuals. My deepest gratitude goes to all of these people.

Firstly, I would like to express my gratitude to the Department of Psychology at the University of Warwick and the European Research Council for funding and supporting my PhD research.

I would also like to thank the nurseries, parents, and children who took part in my developmental research studies reported in this thesis, as well as the adult participants who took part in the research aimed at an adult population. This work would not have been possible without them.

My deepest thanks go to my primary supervisor, Sotaro Kita, and my secondary supervisors, Michaela Gummerum and John Michael, for their unending patience and guidance over the course of the last four years. The quality of both my empirical and theoretical research has improved drastically as a result of working with them. I would also like to thank them for their advice in my professional development, giving me suggestions on which journals, conferences, and funding bodies to target, as well as where my profile could be improved.

My gratitude also extends to other collaborators for their advice in improving the research in this thesis: Barbora Siposova, Keith Jensen, Stephen Butterfill, and Sam Clarke. Their advice in the writing and experimental design phase of research has shaped my way of approaching problems. I would especially like to emphasise the contribution of Stephen Butterfill, with whom I have been working since I was an undergraduate at Warwick, and who has guided my professional development since then.

Beyond direct collaborators, I would also like to extend my thank you to the research assistants involved in the different parts of my projects over the years: Nicole Zhang, Priyanka Modi, Lewis Coyne, Georgiana Puscas, and Jinnie Ooi. Their hard work in collecting data, recruiting participants, and coding, was essential for the completion of this thesis.

My gratitude also goes out to members from both the *sense of commitment* project and Kita's lab meetings, for establishing a welcoming environment within which research ideas could be discussed and improved upon. These include Francesca Bonalumi, Matt Chennells, Wayne Christensen, Michael Hattersley, Simon Myers, Melissa Reddy, Marcell Székely, Suzanne Aussems, Jacob Barker, Jiahao Yang, and many others. Henrik Singmann deserves a special mention for his advice on approaching statistics.

In this vein, I would also like to thank journal editors and reviewers I have received feedback from for their insightful comments that improved my research further still, as well as participants from conferences I attended in Hungary, the USA, the UK, Italy, Germany, and Croatia. I am deeply thankful to all these individuals for their thought-provoking questions and useful suggestions.

Outside of research, I would like to thank Multiverse (formerly WhiteHat) for incentivising my prompt submission by offering me a job starting 1st March 2021.

On a more personal level, I would like to thank my family (Adrian, Clare, Finlay, Nell, and Peggy Green) for supporting me during my research, even in the most difficult of times. For being able to take my mind off work, my thanks go to the members of S2.35 (Tristan Kreetz, Henry Powell, and Jack Shardlow). Finally, I would like to thank Rebecca Carnevali, with whom I have shared the last 5 years of my life, and with whom I look forward to the rest.

Declaration

This thesis is submitted to the University of Warwick in support of the application for the degree of Doctor of Philosophy in Psychology. It has been composed by the author and has not been submitted in any previous application for any degree. The work presented (including data collection and data analysis) was carried out by the author, or principally by the author in collaboration with others as acknowledged. Parts of this dissertation have been prepared for publication and submitted to journals by the author:

Chapter 2: Green, A., Siposova, B., Kita, S., Michael, J. (under review). Stopping at Nothing: Two-Year-Olds Differentiate Between Abandoned and Interrupted Goals

Chapter 3: Green, A., Siposova, B., Jensen, K., Kita, S., Michael, J. (in prep.). Finish What you Started: Instrumental Helping in Two-Year-Olds Motivated by a Preference for Completing Unfinished Actions.

Chapter 5: Green., A. (under review). Honouring an Interpersonal Commitment: A Reductive Theoretical Framework.

Summary

Prosocial behaviour (behaviour aimed at benefiting other agents) and the ability to track the goals of other agents are foundational in human social life. This thesis investigates the development of these in toddlers. Specifically, this thesis investigates the relation between an underexplored aspect of goal tracking, goal status, and one of the earliest emerging prosocial behaviours, instrumental helping.

Chapter 1 introduces the topic and raises questions to be answered in this thesis.

Chapter 2 empirically investigates children's sensitivity to goal status in an instrumental helping context. This is the first evidence that helping in 2-year-olds is sensitive to the distinction between abandoned and interrupted goals.

Chapter 3 investigates the cognitive mechanisms underpinning instrumental helping in early childhood. This experiment's findings are the first to support the goal slippage hypothesis - that children contribute to others' goals because they want to see goals be completed.

Chapter 4 addresses the role of goal status in determining goal salience. Across three experiments, two hypotheses about the relation between goal status and goal salience were tested with an adult population. The results failed to support either hypothesis or replicate past findings.

Chapter 5 focuses on a more applied context for goal tracking and prosocial behaviours: children's ability to honour a commitment. Most research on commitment in early childhood does not focus in detail on the developmental trajectory of interpersonal commitment. To address this gap, I articulate and evaluate evidence for the hypothesis that social interaction scaffolds the development of commitment over childhood.

Chapter 6 discusses the theoretical implications of these findings. The research presented here extends our understanding of the relation between goal tracking and some of the earliest forms of prosocial behaviour, highlighting that the cognitive and motivational foundations of social cognition and prosocial behaviour in early childhood are multi-faceted.

Chapter 1

Introduction

1.1 General Introduction

1.1.1 Goal Tracking

Tracking the goals of other individuals is a prevalent feature of everyday life. As you sit in a cafe watching other patrons peruse the coffee menu and meet with friends, or as you watch shoppers go about their business, it is almost impossible to observe another agent without forming expectations and predictions about their goals. Indeed, humans seem naturally inclined to make sense of the behaviours and actions of other individuals by reference to the goals that they may have (Baillargeon et al., 2016; Csibra, 2008; Csibra & Gergely, 2007; Lombrozo & Carey, 2006).

Goal tracking is a key component of social cognition because it underpins much of our social activities. My understanding of goals is as outcomes to which an agent's actions are directed (Butterfill, 2019); that is, I am understanding goals in terms of states of the world rather than as mental states. Accurately tracking the goals of others can help us to identify others' mental states, such as their preferences (Choi et al., 2018; Luo & Baillargeon, 2005; Luo et al., 2009, 2017) and beliefs (Király et al., 2018; Southgate et al., 2007; Southgate & Vernetti, 2014), as well as enabling us to make predictions about others' future actions and behaviours (Eisenberg et al., 2018; Flanagan & Johansson, 2003; Liepelt et al., 2008). This helps us to achieve our own goals, such as traversing a busy road without colliding into pedestrians or drivers. This also helps us to achieve joint goals, because identifying others' mental states and predicting the future actions of others facilitates coordination between individuals (Knoblich et al., 2011; Konvalinka et al., 2010; Meyer et al., 2015, 2016; Sacheli et al., 2013; Vesper et al., 2010). Goal tracking thus enables agents to achieve outcomes together that we would not be able to achieve otherwise, such as moving heavy objects (e.g., couches) or organising large-scale social activities (e.g., conferences).

These functions that goal tracking serves in our everyday social lives are suggestive of the evolutionary importance of goal tracking. By facilitating coordination and helping us to identify others' mental states, goal tracking encourages cooperation between agents (Cross et al., 2016; Lang et al., 2017; Rusch & Leutge, 2016; Wiltermuth & Heath, 2009) and prosocial behaviour (Fischer et al., 2013; Kokal et al., 2011; Reddish et al., 2014; Valdesolo & DeSteno, 2011). The depth and degree of cooperation and prosociality found in human societies is distinctive (Bratman, 2014; Henrich, 2004; Melis & Semmann, 2010; Tomasello, 2014), and so the prevalence and flexibility of goal tracking in humans may have proven pivotal in the emergence of the complex social structures that are unique to humans.

The foundational role that goal tracking plays in our everyday social lives and its evolutionary importance have generated much interest in the development of this ability in early childhood. Goal tracking emerges early in childhood and advances rapidly over the first few years of life. From as early as 3 months of age children generate predictions about others' future actions on the basis of their past goal-directed behaviours (Cannon & Woodward, 2012; Kim & Song, 2015; Krogh-Jesperson & Woodward, 2014; Luo, 2011; Sommerville & Crane, 2009; though see Ganglmayer et al., 2019 for a failed replication of such findings), and infants even use these same cues to track the goals of non-human agents (Adam et al., 2017; Csibra, 2008; Luo, 2011; Luo & Baillargeon, 2005). From the middle of the first year of life goal tracking is also sensitive to environmental constraints that the target agent faces (Csibra, 2008; Gergely et al., 2002; Liu et al., 2017; Liu & Spelke, 2017; Meltzoff, 1995; Sommerville & Woodward, 2005), as children expect goal-directed actions to be selected on the basis of their efficiency in achieving the relevant goals. Goal tracking ability in the first year of life is also dynamic; infants are better able to track others' goals if they themselves have prior experience in bringing about those same goals (Elsner & Adam, 2020; Gerson & Woodward, 2014; Krogh-Jespersen & Woodward, 2018; Sommerville et al., 2005; Sommerville & Woodward,

2005). By the end of the first year of life children are even able to track goals at a broader conceptual level. For instance, 12-month-olds can identify that an agent's goal may be to acquire any object of a particular kind, such as a truck, rather than to acquire a specific individual object of that kind (Spaepen & Spelke, 2007).

From around the second year of life goal tracking becomes integrated with more complex information. From around the second year of life, goal tracking is sensitive to others' mental states, such as their preferences (Luo & Baillargeon, 2005; Martin et al., 2017; Spaepen & Spelke, 2007) and beliefs (Southgate et al., 2007; Southgate & Vernetti, 2014). Children also start to take into account communicative utterances as cues to others' goals at this age (Jin & Song, 2017; Song et al., 2014; Tauzin & Gergely, 2018). This is also the age at which children's understanding of shared goals emerges, as children start to form predictions about others' behaviours on the basis of previous collaborative interactions (Henderson et al., 2013; Henderson & Woodward, 2011; Krogh-Jesperson et al., 2020). By the end of the second year of life, then, goal tracking has emerged as a flexible ability that can integrate a variety of information.

Goal tracking underpins much of our socio-cognitive abilities, which means that the development of goal tracking can shed light on the development of further social and prosocial behaviours and abilities. For instance, the development of goal tracking facilitates children's engagement in joint actions (Carpenter, 2009). Developments in goal tracking also facilitate social learning, because identifying others' goals can deepen one's own understanding of which actions and behaviours are worth imitating (Bekkering et al., 2000; Carpenter et al., 2005), as well as helping one to learn the means of achieving goals that others have pursued (Esseily et al., 2013). A further kind of prosocial behaviour that goal tracking underpins is helping behaviour, and specifically, instrumental helping.

1.1.2 Instrumental Helping

To instrumentally help another agent is to deliberately contribute to their goal, typically in scenarios in which the agent is unlikely to achieve

that goal by themselves or if they are struggling to do so (Geller & Bamberger, 2009; Hammond, 2014; Warneken & Tomasello, 2006). An everyday example of this kind of helping may be helping a friend to move an object that is too heavy for them to move by themselves, or assisting an elderly person in crossing the road.

Instrumental helping can be contrasted with other types of helping, such as sharing or comforting another agent. All three of these types of helping behaviour emerge early in childhood (Dunfield, 2014; Paulus, 2014; Svetlova et al., 2010; Thompson & Newton, 2013). However, sharing and comforting do not necessarily require goal tracking, as these kinds of helping instead revolve around others' mental states (namely, their preferences and emotional states, respectively). Goal tracking therefore plays a relatively more important role in the emergence of *instrumental* helping.

There has been a wealth of research on instrumental helping in early childhood, demonstrating that children contribute to others' goals in a variety of contexts. Instrumental helping first emerges early in the second year of life (Warneken et al., 2007; Warneken & Tomasello, 2006), as toddlers spontaneously help others to achieve their goal when it has been interrupted. For example, if an experimenter accidentally dropped an object such that it is out of reach, toddlers return the object to the experimenter without being prompted (Warneken et al., 2007; Warneken & Tomasello, 2006). From 18 months of age infants also start to help others anonymously and before others know that help is required (Hepach et al., 2017; Warneken, 2013), and children even extend this behaviour to non-human agents, such as robots and non-geometric shapes (Kenward & Gredebäck, 2013; Martin et al., 2020).

During the second and third years of life instrumental helping also becomes sensitive to different features of the infants' physical and social environment. For instance, toddlers are more likely to help friends (Engelmann et al., 2019), prosocially-behaving agents (Dunfield & Kuhlmeier, 2010), and agents with whom they have previously coordinated

(Barragan & Dweck, 2014; Cirelli et al,. 2014, 2016; Wan et al., 2018). The mere priming of affiliation can boost instrumental helping in 18-month-olds, even if the affiliation does not involve the infant or the helpee (Over & Carpenter, 2009). Instrumental helping is also sensitive to the behaviours of caregivers, as scaffolding (Brownell, 2013, 2016; Dahl et al., 2017; Hammond & Carpendale, 2015; Pettygrove et al., 2013; Schuhmacher et al., 2018) and praise and encouragement (Dahl, 2015; Dahl et al., 2017; Kärtner et al., 2020) boost helping in the second year of life (although for a failure to find effects of caregiver praise on helping, see Warneken & Tomasello, 2013). Furthermore, although helping behaviour may not be increased by material extrinsic rewards (Warneken & Tomasello, 2008), it may be inversely related to the physical costs of helping others (Sommerville et al., 2018; though Warneken et al., 2007, failed to find this effect). Infants are therefore willing and able to instrumentally help others from early in childhood, and this behaviour is sensitive to a number of social and environmental factors.

Several non mutually exclusive proximate psychological mechanisms have been proposed to underpin instrumental helping in early childhood (Brownell, 2013; Dahl & Paulus, 2019; Eisenberg et al., 2016; Hammond & Brownell, 2018; Martin & Olson, 2015; Michael & Székely, 2019; Paulus, 2018; Warneken, 2015). The mechanism that is most prominent in the literature is psychological altruism (Herrmann et al., 2019; Martin et al., 2020; Paulus et al., 2020; Warneken, 2015; Warneken & Tomasello, 2006, 2009). Under the psychological altruism hypothesis, children instrumentally help others because they care about others' well-being, and contributing to the goals that others are pursuing is seen as a means of improving their well-being. This hypothesis is supported by the physiological arousal that infants experience when they observe that another agent requires help (Hepach et al., 2012, 2016, Hepach, Hedley, & Nuske, 2019), and the finding that instrumental helping is not increased by the presence of extrinsic material rewards (Warneken & Tomasello, 2008). A further mechanism that has been suggested to underpin instrumental helping

is the desire to socially interact with others (Carpendale et al., 2015; Dahl, 2015; Dahl & Brownell, 2019; Hammond & Brownell, 2018; Paulus, 2014, 2018; Paulus & Moore, 2012). Specifically, it is hypothesised that infants help because this will enable them to interact with others, which is intrinsically rewarding. This hypothesis is motivated by children seeming to enjoy taking part in activities simply for the sake of interacting with others (Dahl, 2015; Over, 2016), as well as the positive effects of coordination (Barragan & Dweck, 2014) and affiliation (Over & Carpenter, 2009) on instrumental helping. While there are further mechanisms hypothesised to underpin instrumental helping, such as that reputation management (Leimgruber et al., 2012; Engelmann & Rapp, 2018; Engelmann et al., 2013) and direct reciprocity (Hepach, Vaish, et al., 2019; Leimgruber, 2018), psychological altruism and social interaction are the mechanisms that feature most prominently in the literature (for a more in-depth evaluation of these hypothese, see Dahl & Paulus, 2019; Michael & Székely, 2019; Paulus, 2018).

The importance of understanding the cognitive mechanisms that underpin instrumental helping in early childhood is twofold. Firstly, instrumental helping is one of the earliest emerging forms of prosocial behaviour in early childhood (Callaghan & Corbit, 2018; Dunfield, 2014; Dunfield & Kuhlmeier, 2010; Hammond, 2014; Jensen, 2016; Paulus, 2018; Svetlova et al., 2010). Understanding the cognitive mechanisms that underpin this behaviour will therefore provide greater insight into the developmental trajectory of prosociality (Dahl & Brownell, 2019; Grossman, 2018; Hammond & Brownell, 2018; Jensen et al., 2014; Leimgruber, 2018; Martin & Olson, 2015; Vaish & Hepach, 2019; Warneken & Tomasello, 2009; Waugh & Brownell, 2017), which may in turn provide insight into prosocial behaviours in adolescence and adulthood. Secondly, while instrumental helping may not be unique to humans (Barnes et al., 2008; Bshary & Raihani, 2017; Hepach et al., 2020; Melis, 2018; Warneken et al., 2007; Warneken & Tomasello, 2006; Yamamoto et al., 2009, 2012), the depth and prevalence of instrumental helping found in

human society does seem to be unique (Barnes et al., 2008; Decety et al., 2016; Hepach et al., 2020; Jensen et al., 2014; Melis, 2018; Warneken & Tomasello, 2006), and some studies fail to find evidence for instrumental helping in some non-human animals (Pérez-Manrique & Gomila, 2019; Skerry et al., 2011; Tennie et al., 2016). Uncovering the cognitive mechanisms that underpin this earliest emerging prosocial behaviour may therefore provide insight into the evolutionary distinctiveness of human social and prosocial cognition relative to other animals.

1.1.3 Goal Tracking and Goal Status in Instrumental Helping

Instrumental helping is one of the prosocial abilities for which goal tracking is foundational. This is simply because being able to deliberately contribute to another agent's goal first requires that one can identify what that goal is (Jensen, 2016). And indeed, there has been some investigation into the ontogenetic link between these two core aspects of humans' social lives. For instance, instrumental helping behaviour is influenced by children's understanding of the relation between helpees' goals and the means by which they are aiming to achieve those goals. Specifically, from 12 months of age children prioritise instrumentally helping others achieve their goal instead of helping them with the means by which they are attempting to achieve their goals (Hepach et al., 2020; Knudsen & Liszkowski, 2012, 2013; Paulus, 2019). For example, if an agent is attempting to retrieve a tool that they mistakenly believe will enable them to achieve their goal, children will help by providing the agent with a tool that is more appropriate in enabling them to achieve their goal instead of retrieving the tool that the agent is reaching for (Martin & Olson, 2013). This is one example of how children's ability to track others' goals can shape their instrumental helping behaviour.

Research on the development of goal tracking can also uncover the underlying cognitive mechanisms for instrumental helping in early childhood. For instance, individual differences in the rates of instrumental helping between children can partially be explained by children's ability to identify others' goals and the means by which children can themselves

contribute to those goals (Bridgers & Gweon, 2018; Köster et al., 2019). In particular, children's failure to instrumentally help others may in part be due to their developing ability to track another agent's goal object in more complex scenarios, such as when there are multiple potential goal objects (Grosse et al., 2010; Hobbs & Spelke, 2015; Krogh-Jesperson et al., 2015). Goal tracking in the context of instrumental helping is important because it gives us key insight into how these two cornerstones of social cognition and prosocial behaviors emerge and complement each other.

One relatively underexplored aspect of goal tracking is goal status. Goal status refers to the relation between a goal and the actual state of the world, and broadly speaking, a goal can be in one of two states: complete or incomplete. Children's understanding of goal status emerges from around 9 months of age. Specifically, children at this age understand the distinction between a goal that an agent is willing but unable to achieve, and a goal that an agent is simply unwilling to achieve. For instance, consider the goal of handing a toy to a child. Children are more patient if an experimenter accidentally drops a toy while handing it to them (i.e., when the goal is interrupted) as compared to when an experimenter deliberately drops the toy before the child can reach it (Behne et al., 2005; Kachel et al., 2017; Warneken et al., 2012). In the former case the experimenter's goal (handing the toy to the child) is incomplete, whereas in the latter case the experimenter's goal (to tease the child) was achieved. More broadly, children understand the distinction between goals that have been achieved and goals that have been interrupted (Brandone et al., 2014; Meltzoff, 1995; Skerry & Spelke, 2014).

It is important to investigate children's ability to track goals of different statuses because children's understanding of goal status can be used to probe the cognitive mechanisms that underpin instrumental helping. For instance, in an instrumental helping context, children differentiate between scenarios in which an experimenter's goal is interrupted (when they accidentally drop a toy on the ground) and scenarios in which an experimenter's goal is complete (when they deliberately throw a toy on the

ground). Specifically, children in the second year of life are more likely to help the experimenter whose goal is interrupted rather than completed (Warneken et al., 2007; Warneken & Tomasello, 2006). This has motivated the psychological altruism hypothesis: that instrumental helping in early childhood is underpinned by psychological altruism (Warneken, 2013).

Past research on children's understanding of goal status, especially in the context of instrumental helping, has only compared complete goals with one particular kind of incomplete goal - interrupted goals. Interrupted goals are goals that an agent continues to maintain despite facing an obstacle, e.g., when an agent continues to pursue the goal of placing items into a cabinet even though the cabinet doors are closed (Warneken & Tomasello, 2006). There are further ways in which a goal might be incomplete, however.

1.1.4 Goal Abandonment and the Goal Slippage Hypothesis

One further type of goal status which is common in everyday life, whose understanding in children is yet to be investigated, is goal abandonment. Abandoned goals are goals which an agent has previously adopted before later changing their mind and deciding to no longer pursue that goal. One example of this would be the goal of making a cup of coffee late in the afternoon. One might adopt such a goal before realising that it is, in fact, too late in the day for a cup of coffee, and therefore the goal should be abandoned. Representing another agent's goal as abandoned requires understanding that the agent was pursuing the goal at some point in the past, but that they have changed their mind before completing it and are no longer pursuing it.

Identifying whether children understand goal abandonment will enable the testing of a further cognitive mechanism that may underpin instrumental helping behaviour: goal slippage (Barresi & Moore 1996; Eisenberg et al., 2016; Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2019; Paulus, 2014). Under the goal slippage hypothesis, a cognitive tension is generated when an agent identifies that a goal which another agent was recently pursuing will not be completed. This tension

arises from a general tendency to want to see goals that have been started be achieved, and from the violation of our expectation that goals that have been started will be achieved. Until the goal is completed, there is a lack of closure that promotes cognitive tension, thereby encouraging the observer to contribute to the goal of the observed agent. In this way, the goal 'slips' from actor to observer, and from perception to action (Hornstein, 1972; Lewin, 1951; Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2018, 2019).

Beyond the specific context of instrumental helping in early childhood, there are several sources of motivation for the goal slippage hypothesis. From everyday life, consider scenarios in which we observe another agent struggling with a task that we ourselves find relatively straightforward. This might include, for instance, observing a relative who is struggling with technology that we are familiar with, such as opening the front-facing camera on their mobile phone or successfully navigating a particular software programme on their computer. Of course, we may be motivated to help our relatives in such scenarios because we care deeply about them (and are therefore motivated by a concern for their well-being). However, we may also simply find it quite frustrating or irritating to see them struggle with what we consider to be such a straightforward task, and we may thus be motivated to complete the task ourselves (thereby adopting our relative's goal) in order to relieve this tension. While observing someone struggle with a task that we find straightforward may not *always* generate feelings of frustration or irritation, there are certainly scenarios in everyday life where we can be motivated to contribute to and adopt another agent's goal in order to alleviate this cognitive tension.

Further motivation for the goal slippage hypothesis comes from past research with adults about goal status, indicating that the goal status of other agents can influence our own motivation to complete tasks. For instance, simply framing a goal as having been started by another agent can provide extra motivation for an agent to complete the goal, as compared to if the goal is presented as not having been started yet. Consider an experiment by

Kivetz and colleagues with an adult population (Kivetz et al., 2006). Participants were offered loyalty cards for a café (participants would receive a stamp on their loyalty card every time they purchased a coffee at the café, and once the loyalty card was filled with stamps they would receive a free coffee). Participants were given one of two kinds of loyalty card: an initially blank card requiring ten stamps to be filled, or a loyalty card that required twelve stamps which already had the first two stamps filled in (participants were told that these first two stamps had been filled in by a previous customer). Both cards required participants to purchase ten coffees before they received a free coffee, but the goal (receiving a free coffee by filling one's loyalty card with stamps) had already been started in one condition (the condition in which participants were given the twelve-stamp card). Participants with the partially completed twelve-stamp loyalty cards completed their cards faster than those who had been given blank cards, suggesting that simply framing the task as having been started by another agent may have boosted participants' motivation to achieve the goal of filling their loyalty card. There is also research indicating that seeing another agent's goal remaining incomplete can boost one's own motivation in achieving a similar goal. For instance, observing another agent starting but failing to complete an anagram task can boost one's own performance on another anagram task (McCulloch et al., 2011), as compared to the scenario in which the other agent's anagram task is successfully completed (for a review of further similar studies, see Laurin, 2016). While neither of these studies directly test the goal slippage hypothesis, this hypothesis would explain these findings, insofar as one interpretation of these findings is that participants are adopting others' goals for themselves upon identifying that others' goals are incomplete.

A further area of past research which provides some motivation for the goal slippage hypothesis is the co-representation effect in the context of joint action. Specifically, the finding that when engaging in a joint activity with another agent, one's own goal-directed activities can be influenced by those of one's collaborative partners (Sebanz et al., 2003). For example,

when pairs of participants are required to perform actions synchronously (such as placing an object in a particular location at the same time), and one participant must manoeuvre around an obstacle, the second participant will also move their arm *as if* they also had to navigate around an obstacle (Schmitz et al., 2017). In addition to this effect being found in adult populations, this effect has also been found for children as young as 4 years old (Milward et al., 2014; Saby et al., 2014). Goal slippage may play a role in this effect, insofar as adopting another agent's goals would explain why observing others performing goal-directed actions can influence our own goal-directed actions.

At this stage there are several important questions about goal slippage that are yet to be answered, as the goal slippage hypothesis is yet to be directly investigated. One such question is the degree to which goal slippage can explain our behaviours. For instance, it is clear that we do not *always* feel motivated to complete goals that others have started; it is not always irritating or annoying to see someone struggle with a task or to see someone leave a task unfinished. There are also further instances in which we might want to avoid completing a goal that another agent has started, such as if that goal is bad for the agent's well-being (e.g., if the goal in question is damaging to the agent's health) or if the goal is bad for our own well-being.

This is an important topic to address in future research, as it is challenging to answer at present with the lack of research on goal slippage. For present purposes, it is important to bear in mind that, theoretically, goal slippage is not proposed to operate within a vacuum – much like other cognitive mechanisms, it is not expected to necessarily over-ride all other sources of motivation that an agent might have. Indeed, one could make the same point and raise the same question about other proposed mechanisms for instrumental helping, such as psychological altruism or the social interaction hypothesis. We certainly do not *always* behave altruistically towards others. And indeed, just as instrumental helping behaviour begins to emerge in the second year of life, so too does deliberately harmful

behaviour such as biting or kicking (Dahl, 2019). We are also certainly not *always* motivated to seek social interaction with others, as there are times when we prefer some time by ourselves. The degree to which goal slippage underpins our behaviours, and the ways in which this mechanism might interact and mesh with other sources of motivation in our everyday lives, is an open question for empirical research.

A further question about goal slippage that is important to address in future research is the types of goals that we should expect to slip from actor to observer. While we might adopt others' goals when these goals are relatively straightforward to achieve, it is less clear that we would adopt the goals of others that are more long-term and less straightforward to achieve. For instance, I might adopt my partner's goal of finishing a puzzle or doing the washing up (such that I am now internally motivated to achieve these goals myself), but I am not similarly motivated to contribute to their goal of completing their PhD thesis or having a successful career. And indeed, the studies cited above that motivate the goal slippage hypothesis focus on goals with a relatively straightforward means of completion, such as filling out a café's loyalty stamp card or completing a puzzle. Given the lack of research on goal slippage at present it is difficult to identify the degree to which these more complex and long-term goals would slip. However, it is also important to bear in mind that the same concern can also be raised of alternative mechanisms for instrumental helping behaviour, such as the social interaction hypothesis or psychological altruism. Although I do care deeply about my partner's well-being and enjoy socially interacting with them, I am nevertheless not completing their PhD thesis write-up myself or ensuring that they have a successful career. It is therefore an open question for empirical research regarding the degree to which these cognitive mechanisms that have been proposed to underpin instrumental helping behaviour motivate us to help others with more complex and long-term goals.

Turning to instrumental helping in early childhood, specifically, there are several reasons for which it is worth investigating the possibility

that the goal slippage mechanism may underpin at some least some instances of instrumental helping. The goal slippage hypothesis can explain several past findings regarding instrumental helping behaviour in early childhood. Broadly speaking, adopting another agent's goal for oneself (and thereby being internally motivated to contribute to that goal) constitutes helping behaviour. For instance, if an agent accidentally dropped an object, goal slippage also predicts that children would retrieve the object and give it back to the agent (Warneken & Tomasello, 2006). This means that many of the studies that have been cited as providing evidence in support of the psychological altruism hypothesis could also be interpreted as supporting the goal slippage hypothesis, as many past helping studies involve an experimenter struggling to complete a goal themselves (Svetlova et al., 2010). The cognitive tension that is generated when one observes another agent struggling to bring about a goal themselves would also account for the physiological arousal that children experience in such circumstances (Hepach et al., 2012), and goal slippage would also predict that this arousal subsides once the target goal is completed (because once the target goal is complete, any cognitive tension would be relieved). The explanation here afforded by the goal slippage hypothesis is that this physiological arousal is a result of the tension that is created when one observes a goal remaining incomplete.

The goal slippage hypothesis is also consistent with a further set of findings around instrumental helping, namely, that instrumental helping is motivated by intrinsic rewards rather than extrinsic rewards. At 20 months of age, children who have received a reward (a fun toy to play with) for instrumentally helping another agent during a treatment phase will subsequently help at lower rates than children who received no reward or praise during the treatment phase (Warneken & Tomasello, 2008). This suggests that children's motivation to instrumentally help others is intrinsic rather than motivated by external rewards. Under goal slippage the motivation to help others arises from a desire to address the cognitive tension that arises when one identifies that another agent's goal will not be

completed. In this sense, goal slippage is also an intrinsic source of motivation for helping behaviour, and therefore goal slippage is consistent with this data.

The goal slippage hypothesis is also particularly well-placed to account for certain recent findings on instrumental helping in early childhood that alternative mechanisms may struggle to explain. Notably, the goal slippage hypothesis can account for children's willingness to help a non-human agent (including a geometric shape) to achieve its goal (Kenward & Gredebäck, 2013; Martin et al., 2020), because the goal slippage mechanism does not require that the helpee be a human agent. All that the goal slippage mechanism requires is that children are attributing a goal to these non-human agents, and there is evidence that children can attribute goals to non-human agents (including geometric shapes) from the third month of life (Luo, 2011). This finding could only be accounted for under the social interaction hypothesis if there were evidence that children found social interaction with non-human agents (including geometric shapes) as satisfying or enjoyable as with humans; it is not clear that children find interacting with an expressionless and unemotive geometric shape intrinsically rewarding in the way that interacting with other humans might be. In order for this to be explained by psychological altruism, we would need evidence that children are as inherently concerned with the well-being of all agents, including non-human agents, as they are with humans (Warneken, 2015). While this may be the case, there is as yet no evidence to support the hypothesis that children are concerned with the well-being of geometric shapes by 17 months of age. Goal slippage, in contrast, requires no further empirical or theoretical steps in order to account for this finding.

There are two further sets of findings that goal slippage is well placed to account for. First is the finding that instrumental helping in early childhood is subject to the bystander effect. That is, by 5 years of age, children are less likely to help others if other agents are also in a position to provide the required help (Plötner et al., 2015). This is straightforwardly

accounted for under the goal slippage hypothesis because if one sees that others are *also* in a position to provide help, then the chances that the goal will be brought about without one's help is considerably higher than if there are no other agents who can provide the necessary help. This is important because if the goal is likely to be achieved, then we would expect less cognitive tension to arise as a result of observing another agent's goal being incomplete, and therefore goal slippage predicts that children would not be as motivated to contribute to the helpee's goal themselves. Explaining this finding under alternative mechanisms for instrumental helping is more challenging, however. Under psychological altruism one's concern for the helpee should motivate one to help that agent, and it is not clear why the diffusion of responsibility to help that arises from the presence of others who are also in a position to help would diminish one's concern for the well-being of the helpee. In addition, if one were motivated to help others due to a desire to interact with them (as under the social interaction hypothesis), it is not clear why this source of motivation would be diminished by the presence of others who are also in a position to help. Indeed, if anything, the presence of others who are also in a position to help might increase the chances that one would be able to interact with others by helping the helpee, and so we might expect higher levels of helping in this context under the social interaction hypothesis.

Goal slippage is also well placed to explain the fact that instrumental helping in early childhood is not sensitive to others' needs (at least, initially). 1.5-year-olds were presented with the option of helping two agents by retrieving an object for them, but children were only able to help *one* of these agents because they were requesting the same object (Paulus, 2019). Both agents were requesting this object in order to achieve a further goal, but only one of the agents *needed* the object to achieve this goal. Participants did *not* preferentially help the more needy agent. This behaviour cannot be straightforwardly accounted for under the social interaction hypothesis or the psychological altruism hypothesis. One might expect the more needy agent to be more grateful of receiving help than the

non-needy agent. The social interaction hypothesis should therefore predict that children would preferentially help the more needy agent in the hope that this additional gratitude would result in further social interaction with the needy agent in future. In addition, psychological altruism should predict that children would preferentially help the more needy agent because this would have a greater impact on their well-being (as the non-needy agent does not require help to achieve their further goal). Goal slippage can straightforwardly account for this behaviour, however. Both helpees in this study had the same incomplete goal (retrieving the out-of-reach object), and goal slippage does not predict that children's helping would be sensitive to others' needs beyond the target goal. Goal slippage would therefore not predict that children would preferentially help the more needy helpee in this context.

Beyond past empirical research, there are also theoretical reasons for which it may be important to directly investigate the goal slippage hypothesis. Specifically, it is important because of the implications of this hypothesis for our understanding of the emergence of prosociality in early childhood. The goal slippage hypothesis is, at its core, a self-oriented mechanism for instrumental helping (in contrast to, for instance, psychological altruism, under which one offers help to other agents for the sake of the other agents). Evidence supporting the hypothesis that goal slippage may underpin at least some instances of instrumental helping in early childhood would therefore suggest that the origins of prosocial behaviours may be more self-oriented in nature than previously thought, and it may therefore require that theoretical accounts of the development of prosociality be reconsidered (Warneken, 2015).

It is also important to note that there are instrumental helping studies whose findings goal slippage is not straightforwardly able to account for. For instance, consider children's preferential helping of others. 21-month-olds will preferentially help others who have behaved in a prosocial manner in a previous interaction (e.g., by intending to provide a toy to the child) than those who have not (Dunfield & Kuhlmeier, 2010), and 3-year-olds

will preferentially help friends over non-friends (Engelmann et al., 2019). Such findings are not readily explained by goal slippage. Provided that the other agent's goal is identified, goal slippage would predict that children will be as likely to help prosocial others or friends as much as anyone else. The only scenario in which goal slippage may be able to explain these findings is if the goals of friends or prosocially-behaving agents are more salient for children than those of non-friends or non-prosocially-behaving agents. These findings are thus better accounted for by the social interaction hypothesis (as one might expect more social interaction with friends or agents who have previously behaved more positively towards one) or psychological altruism (as one might have more concern for friends or others who have behaved prosocially in prior interactions).

In addition to this, goal slippage is not well placed to account for the differences in helping behaviour that arise as a result of the emotions of both children and the helpee. 18- to 25-month-olds are more motivated to help an agent who has been harmed than an agent who has not been harmed (Vaish et al., 2009), which suggests that helping behaviour is sensitive to children's sympathy for the target agent. In addition, from the second year of life children are more likely to help others who exhibit negative emotions, such as distress, in the face of an incomplete goal, as compared to agents who exhibit neutral emotions (Bandstra et al., 2011; Bischof-Köhler, 1991; Campbell et al., 2015; Hepach et al., 2013; Johnson, 1982; Roth-Hanania et al., 2011; Walle et al., 2017; Zahn-Waxler et al., 1992). Goal slippage may be able to account for such findings *if* such negative emotions, or feelings of sympathy, make these goals more salient, but otherwise it is not clear that goal slippage could account for these findings.

It is clear that there are helping behaviours that goal slippage is not best placed to explain, then. How problematic this is for the goal slippage hypothesis depends on one's expectations about individual proposed mechanisms for instrumental helping. As has been articulated above, there are findings that both the psychological altruism hypothesis and the social interaction hypothesis struggle to account for as well, in the context of

instrumental helping and beyond. For instance, it is not clear under the social interaction hypothesis why 18-month-olds would help an agent who is not present and therefore not aware that they have been helped (Hepach et al., 2017). Psychological altruism cannot straightforwardly account for the finding that helping at 18 months of age is sensitive to the physical costs of helping (Sommerville et al., 2018), or that deliberately harmful behaviour emerges alongside helping behaviour in the second year of life (Dahl, 2019). Goal slippage cannot straightforwardly explain *all* instances of helping behaviour, but this is also true for other proposed mechanisms for instrumental helping. This is therefore only problematic if one were to endorse a view under which only a single cognitive mechanism underpins all instrumental helping, and this is not a view endorsed in this thesis.

Further discussion of the goal slippage hypothesis can be found in Chapter 3. For the moment, it is important to be aware that testing goal slippage requires first identifying whether children understand goal abandonment. This is because it is only under conditions of goal abandonment that goal slippage generates unique predictions about children's behaviour in an instrumental helping scenario (Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2019) (again, this will be spelled out in Chapter 3). This line of reasoning generates the first and second research questions (RQ) of the present thesis:

> RQ 1: Do children understand goal abandonment? More specifically, are young children sensitive to the distinction between abandoned goals and goals of other types of status, e.g., interrupted goals?
>
> RQ 2: Is goal slippage one of the cognitive mechanisms that underpins instrumental helping in early childhood?

1.1.5 Goal Status and Goal Salience

The importance of goal status is not limited to the context of instrumental helping behaviour in early childhood, however. Past research has indicated that goal status also plays a role in determining the relative salience of different goals in an adult population. For instance, adult participants are more likely to mention a target goal that was incomplete (as

compared to complete) when asked to summarise stories verbally and in writing (Asiala et al., 2020; Suh & Trabasso, 1993; Trabasso & Suh, 1993), and participants determine whether a target word is associated with an agent's goal more quickly and accurately if the goal is incomplete as compared to complete (Lutz & Radvansky, 1997; Magliano & Radvansky, 2001; Radvansky & Curiel, 1998). Goal status therefore seems important in determining goal salience, with past findings indicating that incomplete goals are more salient than completed goals (Asiala et al., 2020; Magliano & Radvansky, 2001; Moss et al., 2007; Trabasso & Suh, 1993).

It is important to understand what determines goal saliency for several reasons. Goal saliency may determine which goals we preferentially track, such that we prioritise tracking goals that are relatively more salient than those that are less salient. This is significant because of the prevalence of goal tracking in our everyday social lives and because goal tracking underpins much (pro)social cognition and behaviour. For example, one may be more likely to help another agent if that agent's goals are more salient, and therefore one is tracking that agent's goals more closely than one would otherwise. In this way, goal salience may impact which goals we decide to act upon, and therefore which prosocial behaviours we will exhibit in everyday life. In addition, both children from the age of 12 months and adults are better able to predict others' actions if their goals are more salient (Adam et al., 2016; Eshuis et al., 2009; Henrichs et al., 2012). This means that goal saliency also plays a role in one's ability to coordinate successfully with others. Therefore the relation between goal status and goal salience is worth addressing.

Past research indicating that incomplete goals are more salient than complete goals does not draw a clear distinction between different types of incomplete goals, however. Instead, the focus of this past research primarily compares interrupted and completed goals (Asiala et al., 2020; Magliano & Radvansky, 2001; Trabasso & Suh, 1993) without considering the role of abandoned goals in determining goal salience. As we shall see (in Chapter 4), this distinction between abandoned and interrupted goals carries

implications for the reasons for which goal status determines goal salience. This generates our next research question:

> RQ 3: What is the relation between goal status and goal salience? Specifically, how salient are abandoned goals in relation to interrupted and completed goals?

1.1.6 Interpersonal Commitments

This thesis focuses primarily on the development of goal tracking and instrumental helping. What about the developmental trajectory of other kinds of prosocial behaviour involving instrumental helping and goal tracking? One candidate that has recently been the focus of developmental research is honouring an interpersonal commitment.

Broadly speaking, an interpersonal commitment can be understood as a relation between two agents and an action or a behaviour, under which one of the agents has an obligation to perform the action (or exhibit the behaviour) (Michael & Salice, 2017). Interpersonal commitments can be contrasted with individual commitments, such as committing to learn a language or exercise regularly, under which an agent makes a commitment to themselves to perform an action (Clark, 2006; Michael & Pacherie, 2015). Examples of interpersonal commitments from everyday life would be the commitment to honour the terms of one's job contract, or promising to help a friend move house. Honouring an interpersonal commitment requires similar skills to goal tracking in an instrumental helping context - one must be able to identify what the other agent's goal is, and be willing and able to contribute to that goal. For instance, in helping a friend move house, one must be able to identify the goal that your friend wants you to contribute towards (e.g., moving furniture from their old home to their new home), and then be both willing and able to contribute to that goal (e.g., being able and motivated to carry heavy furniture).

One's ability to honour interpersonal commitments is an important component of our everyday social lives because of the role that commitments play in shaping our motivations and others' expectations about our future actions. Forming a commitment stabilises our motivations

to perform certain actions and to persist with tasks that we might find boring, effortful or unpleasant (Castro & Pacherie, 2020; Michael & Pacherie, 2015; Rachlin, 2016). This may be because they help to solidify and clarify our intentions to ourselves and we might then feel bad if we fail to achieve the goals that we set ourselves (Bratman, 2014; Bryan et al., 2010; Morgan et al., 1990; Schrift & Parker, 2014; Siegel & Rachlin, 1995), or because not fulfilling a commitment may harm one's reputation and violate others' expectations about one's future actions (Bonalumi et al., 2019; Heintz et al., 2015; Michael & Pacherie, 2015; Michael, Sebanz, & Knoblich, 2016). Commitments also act as signals to other agents regarding our intended future actions, thereby enabling others to trust us and rely on us to perform certain actions (Bratman, 1999; Michael & Pacherie, 2015; Michael & Székely, 2018). Many actions and behaviours in everyday life are a result of interpersonal commitments. For example, we may do house chores because we have committed to sharing this labour equally with our partners.

Commitments may have also played an important role in the evolution of human society. By shaping our motivations and others' expectations about our future actions, commitments facilitate cooperation and coordination between agents (Michael & Pacherie, 2015), from small-scale dyadic interactions to large-scale groups (Bratman, 2014; Bryan et al., 2010; Gilbert, 2006, 2018; Tomasello, 2019). By facilitating cooperation on a broad scale, commitment may have enabled agents to achieve outcomes together that they would have struggled to achieve individually, as well as underpinning the uniquely complex social structures that human societies have developed relative to non-human animals (Balliet et al., 2017; Gilbert, 2006; Han, 2013; Roberts, 2005; Tomasello, 2014).

Given the role of interpersonal commitments in everyday life and from an evolutionary perspective, it is important to understand how children's ability to honour interpersonal commitments develops over childhood. This may provide key insight into the cognitive architecture of interpersonal commitment in adulthood, constraining future theorising about

commitment and generating testable hypotheses about the factors that modulate commitment. Such an investigation would contribute to ongoing research into the emergence of prosocial behaviour and motivation in early childhood (Dahl & Brownell, 2019; Köster et al., 2019; Spinrad & Gal, 2018; Paulus, 2018; Melis, 2018; Sommerville et al., 2018; Wan et al., 2018), as well as carrying practical implications for improving persistence in adults.

Empirical research into the development of commitments can be divided into two groups: research focusing on children's expectations about others' willingness to commit to actions, and children's own abilities to honour interpersonal commitments. Most developmental studies about commitment have focused on children's expectations about whether other agents will remain committed to a joint task (Astington, 1988a, 1988b; Chin & Lin, 2018; Gräfenhain et al., 2009; Mant & Perner, 1988; Siposova et al., 2018). For instance, such studies have investigated the conditions under which children believe that commitments can be dissolved (Kachel et al., 2019), or the types of protests and judgements that children make in response to the violation of a commitment (Hussar & Hovarth, 2013; Kachel et al., 2017; Kanngiesser et al., 2017).

Other studies have focused on children's own ability to honour an interpersonal commitment (Gräfenhain et al., 2009; Kachel & Tomasello, 2019; Kanngiesser et al., 2017) and on children's ability to remain committed to a joint task (Gräfenhain et al., 2013; Hamann et al., 2012). The studies that focused on children's ability to honour interpersonal commitments indicate that children dramatically improve at honouring interpersonal commitments between the ages of 3-5 years old (Gräfenhain et al., 2009, 2013; Kachel & Tomasello, 2019; Kanngiesser et al., 2017), and children's understanding of commitments continues to improve into late childhood (Chin & Lin, 2018; Hussar & Hovarth, 2013; Siposova et al, 2018).

The studies that have focused on children's ability to honour interpersonal commitments do not directly focus on what changes during

these early years which underpins children's improvement in honouring commitments, however. Previous studies therefore do not explain the developmental trajectory of interpersonal commitment in early childhood. Instead, these studies employ cross-sectional designs to investigate differences between age groups without focusing on accounting for such differences (Gräfenhain et al., 2009; Kachel & Tomasello 2019; Kanngiesser et al., 2017). Even theoretical articles on the development of commitment largely neglect this issue, choosing instead to either focus on which cognitive mechanism(s) motivates children to honour interpersonal commitments (Castro & Pacherie, 2020) and children's conceptual understanding of commitment (Michael & Székely, 2018). There is therefore a lack of research into the developmental trajectory of children's abilities to honour interpersonal commitments over childhood, generating the final research question for this thesis:

> RQ 4: What is the developmental trajectory of interpersonal commitment? What changes, and why does it change, such that children become better at honouring interpersonal commitments over the course of childhood?

1.2 Research Questions and Hypotheses

I have identified four research questions that concern (pro)social cognition and behaviours over development and in adulthood. By answering these questions, this thesis will contribute to ongoing research into our understanding of (pro)social cognition and behaviours in early childhood. Specifically, this thesis will extend our understanding of the relation between goal tracking and some of the earliest emerging forms of prosocial behaviour (namely, instrumental helping and interpersonal commitment). This will generate testable hypotheses about the ontogenetic and phylogenetic roots of goal tracking and prosocial behaviours, as well as raising new questions about prosociality and social cognition in adulthood. Here, I will spell out the hypotheses are associated with each research question, and which will therefore be tested and articulated in this thesis.

RQ 1: Do children understand goal abandonment? More specifically, are young children sensitive to the distinction between abandoned goals and goals of other types of status, e.g., interrupted goals?

No research has directly investigated this question. However, research illustrates that children understand the distinction between interrupted and completed goals from 9 months of age, and this understanding guides their helping behaviour from around the second year of life. As discussed earlier, children are willing and able to instrumentally help an agent achieve a goal that has been interrupted by the presence of the external obstacle from around the second year of life (Hepach et al., 2017; Kenward & Gredebäck, 2013; Warneken et al., 2007; Warneken & Tomasello, 2006). However, children younger than 2 years of age struggle to help appropriately if there are multiple helping affordances or when cues for an agent's goal are ambiguous between multiple goal objects (Hepach et al., 2016; Hobbs & Spelke, 2015; Krogh-Jespersen et al., 2015; Waugh & Brownell, 2017). That is, children may struggle to help appropriately until the age of 2 when there are multiple possible ways in which they may be able to help an agent (e.g., if the agent is asking for one of multiple available objects). These findings motivate the hypothesis that children's instrumental helping behaviour should be sensitive to the distinction between abandoned and interrupted goals by the age of 2. This hypothesis carries implications for our second research question:

RQ 2: Is goal slippage one of the cognitive mechanisms that underpins instrumental helping in early childhood?

As has been discussed previously (and will be further discussed in Chapter 3), testing the goal slippage hypothesis first requires identifying the age at which children understand goal abandonment (Michael, Knoblich, & Sebanz., 2016; Michael & Székely, 2019). This is because the goal slippage hypothesis generates unique predictions about children's behaviour (relative to other cognitive mechanisms for instrumental helping) once the helpee abandons their goal. Given that we are testing the hypothesis that children understand goal abandonment from the age of 2 years old, we will test the

hypothesis that the goal slippage mechanism underpins instrumental helping behaviour in children of the same age (i.e., 24- to 30-month-olds).

RQ 3: What is the relation between goal status and goal salience? Specifically, how salient are abandoned goals in relation to interrupted and completed goals?

We will test two hypotheses that generate different answers to RQ 3. One hypothesis, which I will refer to as the *predictive value hypothesis*, suggests that goal salience is determined by how useful a goal is in predicting the future behaviour of other agents (Asiala et al., 2020; Lutz & Radvansky, 1997; McNamara & Magliano, 2009; Trabasso & Suh, 1993; Zwaan & Radvansky, 1998). Under this hypothesis, interrupted goals should be more salient than abandoned or completed goals, because a goal that an agent is no longer pursuing (i.e., an abandoned or a completed goal) is less useful in predicting the future behaviour of the agent in comparison to a goal that the agent is currently pursuing (i.e., an interrupted goal). Under the predictive value hypothesis, then, goal status determines goal saliency because of the differences in how useful goals of different statuses are in predicting others' actions.

An alternative hypothesis about the relation between goal status and goal salience is the goal slippage hypothesis. Under this hypothesis, a cognitive tension is generated when an agent identifies that a goal which another agent was recently pursuing will not be completed. This cognitive tension arises from a general tendency to want to see goals that have been started be achieved, and the violation of our expectation that goals that have been started will be achieved (Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2018, 2019). Under the goal slippage hypothesis, interrupted and abandoned goals should both be expected to generate cognitive tension in an observer, because in both cases the target goal is incomplete. We should therefore expect these types of goals to be equally salient, and more salient than completed goals.

RQ 4: What is the developmental trajectory of interpersonal commitment? What changes, and why does it change, such that

children become better at honouring interpersonal commitments over the course of childhood?

This research question is fairly broad, insofar as there are a wide number of factors that might influence the developmental trajectory of interpersonal commitment over the course of childhood. I will therefore answer RQ 4 with a two-step theoretical approach. First, I will introduce a theoretical framework for approaching questions about the development of commitment, reducing the ability to honour an interpersonal commitment into its constitutive components - prosocial motivation, cognitive ability, and practical ability. The second step is to adopt a second-person approach to this research question (Carpendale & Lewis, 2004; Moore & Barresi, 2017; Siposova & Carpenter, 2019; Tomasello, 2019) which emphasises the role of social interaction in children's socio-cognitive development. I will spell out and evaluate evidence for the hypothesis that social interaction scaffolds the development of interpersonal commitment in early childhood.

1.3 Structure of Dissertation

This thesis is divided into four main chapters, each of which addresses one of the research questions identified above. Chapter 2 addresses RQ 1, investigating whether 24- to 30-month-old children understand goal abandonment by testing the hypothesis that children's helping behaviour is sensitive to the distinction between abandoned and interrupted goals by the age of two. We designed an instrumental helping task in which an experimenter begins, but does not complete, a goal-directed action. We manipulated the reason for which the experimenter does not achieve his goal: he has either abandoned his goal in favour of an exclusive alternative, or else he has encountered an obstacle. In both cases, the experimenter then asks the child for help, and we measured children's helping behaviour.

Chapter 3 answers RQ 2, by directly testing whether the goal slippage mechanism underpins instrumental helping behaviour in early childhood. This study is also based on previous instrumental helping studies (Hepach et al., 2017; Warneken & Tomasello, 2006), and follows the same

broad structure: an experimenter initiates a goal-directed action but does not finish it. The crucial manipulation involves the experimenter not completing a goal-directed action in the experimental condition because it was abandoned (rather than because it was interrupted, as in most past helping studies). We measured whether participants completed the experimenter's goal after it was abandoned. Participants in Chapter 3's experiment are also 24-30 months old because we could only conduct this experiment on an age group that understood goal abandonment (i.e., the same age group as in Chapter 2).

Chapter 4 focuses more broadly on the relation between goal status and goal tracking, addressing RQ 3. In order to answer this question, we tested two competing hypotheses about the relative salience of completed, interrupted, and abandoned goals: the predictive value hypothesis, and the goal slippage hypothesis. We conducted three experiments based on past studies that indicated that incomplete goals were more salient than completed goals (Asiala et al., 2020; Magliano & Radvansky, 2001; Moss et al., 2007; Trabasso & Suh, 1993). Adult participants were asked to read short vignettes in which the status of a protagonist's goals was manipulated. Participants were then asked to summarise the vignettes, and we measured whether the target goals were mentioned in these summaries.

Chapter 5 addresses RQ 4 by articulating a particular version of the hypothesis that social interaction scaffolds the development of interpersonal commitment in early childhood. First, I introduce a theoretical framework that reduces one's ability to honour interpersonal commitments into its constitutive parts: prosocial motivation, cognitive ability (to persist with tasks), and practical ability. I then suggest that social interaction scaffolds the development of interpersonal commitment by introducing rewards that motivate and reinforce the emergence of these constitutive components of interpersonal commitments over the course of childhood. More specifically, I suggest that others' emotions may be pivotal in motivating and reinforcing the development of prosocial motivation and cognitive abilities for interpersonal commitment over childhood. I then evaluate evidence that

pertains to this hypothesis and propose ways in which this hypothesis might be tested in future research.

Finally, Chapter 6 discusses the implications of the thesis' research for our understanding of goal tracking and prosocial behaviour in early childhood. This includes discussions of how the research presented here may prompt a re-evaluation of current theoretical approaches to prosociality in development. This chapter then identifies fruitful directions for future research in (pro)social cognition in both ontogeny and phylogeny, in addition to some practical implications of the present research.

Chapter 2

Stopping at Nothing: Two-Year-Olds Differentiate Between Interrupted and Abandoned Goals

Previous research has established that goal tracking emerges early in the first year of life. However, it has not yet been shown whether young children can represent others' goals as changing prior to completion. The current study investigates this by probing young children's (24-30 months old) ability to differentiate between goal directed actions that have been halted because the goal was interrupted, and because the goal was abandoned. To test whether children are sensitive to this distinction, we manipulated the experimenter's reason for not completing a goal-directed action - his initial goal was either interrupted, or it was abandoned in favour of an alternative. We recorded whether children's helping behaviour is sensitive to the experimenter's reason for not completing his goal-directed action by recording whether they complete the experimenter's initial goal or the alternative goal. The results showed that children (n = 24) helped complete the experimenter's initial goal significantly more often after this goal had been interrupted than after it had been abandoned. These results support the hypothesis that children can represent others' goals as changing prior to completion from 2 years of age, and specifically that they differentiate between abandoned and interrupted goals.

2.1 Introduction

From navigating a busy street to organising a business meeting, there are many everyday contexts in which it is important to be able to attribute goals to other agents in order to predict their actions and to coordinate our actions and plans with them. This ability emerges early in the first year of life and rapidly becomes increasingly sophisticated. By 3 months of age infants expect agents to pursue goals that are consistent with their prior goal-directed behaviour (Kim & Song, 2015; Luo, 2011; Sommerville et al., 2005). By 12 months of age infants are sensitive to information about other agents' internal states - such as preferences (Luo & Baillargeon, 2005; Spaepen & Spelke, 2007) and beliefs (Southgate et al., 2007; Southgate &

Vernetti, 2014) - as well as external environmental constraints (Csibra, 2008; Liu & Spelke, 2017) in attributing goals. From early in the second year of life children also take verbal (Jin & Song, 2017) and non-verbal (Tauzin & Gergely, 2018) communication into account when attributing goals.

Accurate goal tracking requires us to continuously update representations of others' goals over time - in particular because sometimes agents change their goals prior to completion. For example, suppose that an agent has the goal of placing a book in a cabinet. She picks up the book, walks over to the cabinet, and begins to open the cabinet door - but then halts her action. In order to discern whether she still has the goal of putting the book in the cabinet, it is crucial to identify why she halted her action. Environmental circumstances may have *interrupted* her goal by presenting her with an obstacle (e.g., she may have noticed that she cannot open the cabinet door without a key, and that she must therefore set the book down while going and getting the key). Alternatively, she may have changed her mind, *abandoning* the goal of placing the book in the cabinet (e.g., she may have been attracted by the cover of the book and decided that she would like to read it). Only in the latter case should the goal that we attributed to this agent (of putting the book in the cabinet) be updated; in the former case she has merely halted the action momentarily because of an obstacle.

Past research on goal attribution in early childhood uses experimental scenarios in which the target agent's goal remains constant over the course of individual test trials. Little is therefore known about whether young children can represent others' goals as changing prior to completion. This is a significant gap: the ability to update our goal representations is crucial for many aspects of social cognition, supporting action prediction, coordination in joint action, and flexible prosocial behaviour, in particular instrumental helping. Reconsider the above example: if the agent's goal of placing the book in the cabinet was interrupted by an obstacle (the lock), she would be grateful if a friend were to help by retrieving the key for her. Alternatively, if she has abandoned the

goal, it would be bad if her friend blindly persisted in helping her. Depending on why the initial goal-directed action was halted, then, it may or may not be appropriate to persist in contributing to that goal. It is therefore important that children track others' goals over time and update their representations of others' goals, and in particular, that they distinguish between interrupted goals and abandoned goals. Research investigating whether young children are fluent in making this distinction would therefore deepen our theoretical understanding of the development of goal attribution and flexible prosocial behaviour.

Although there has been no research directly investigating whether children distinguish between interrupted and abandoned goals, or whether children understand goal abandonment, there is substantial research which establishes that children understand goal interruption. Most notably, by 18 months children instrumentally help an agent who encounters an obstacle while pursuing a goal (Drummond et al., 2014; Hepach et al., 2012, 2016; Meltzoff, 1995; Svetlova et al., 2010; Warneken & Tomasello, 2006). In addition, infants as young as 9 months old differentiate between instances in which an adult is unwilling to pass them a toy and instances in which the adult is unable to do so because she has dropped the toy (and has been, in this sense, interrupted; Behne et al., 2005; Brandone et al., 2014). Moreover, 21- and 27-month-olds respond differently to an experimenter who is unwilling to continue playing than one who is willing but who has merely been interrupted by dropping a toy (Warneken et al., 2012). Finally, 3-year-olds help in a manner that is sensitive not only to an adult's immediate request but to their more distal goal (i.e., they decline to pass an adult a requested object if that object is not helpful in light of the adult's distal goal; Martin & Olson, 2013).

Taken together, these findings indicate that children understand when an agent maintains a goal despite the presence of an obstacle which prevents the immediate completion of the goal, and that they can use this information to guide their helping behaviour. None of them, however, directly addresses the question of whether young children understand when

an agent has taken up a goal and then subsequently abandoned it for a different goal. The current study addresses this question.

Because there is evidence that children understand goal interruption, and are motivated to instrumentally help when an agent's goal is interrupted, an instrumental helping paradigm presents an ideal context in which to probe young children's ability to distinguish between interrupted and abandoned goals. In addition, the use of an instrumental helping paradigm enables us to contribute to the literature on prosocial behaviour in early childhood.

If children can differentiate between goal interruption and goal abandonment, and use this information to guide their actions in instrumental helping tasks, then we should expect them to exhibit this ability by the age of 2 years old. At ages younger than 2 years old we would not expect this: children younger than 2 years of age struggle to appropriately help when there are multiple helping affordances (i.e., when there are multiple possible goals that the experimenter might have, and the child must choose the correct one) (e.g., Hepach et al., 2016; Hobbs & Spelke, 2015; Krogh-Jespersen et al., 2015; Waugh & Brownell, 2017).

To test this hypothesis, we developed an instrumental helping paradigm in which we manipulated why the experimenter (E) did not complete a goal-directed action. On test trials E began to place a toy into one of two boxes (initial location), but did not complete this action. In the abandoned goal condition, E indicated that he had changed his mind, and would prefer to place the toy in the other box (alternative location). In the interrupted goal condition, E encountered a physical obstacle which prevented him from reaching the initial box. Despite this, he continued to maintain the initial goal (i.e. his goal was still to see the toy placed in the initial location).

We sought to maximize the number of trials in which children would try to help because we were interested in the cognitive underpinnings of helping behaviour rather than the motivation. Accordingly, E asked the child to help, without specifying which box was his current goal. We measured

where children helped to place the toy, i.e. whether they helped to place the toy into the box that was E's initial goal (initial location), or whether they helped to place the toy in the other box (alternative location). In order to appropriately help E, children had to continue to update their representations of E's goals and to distinguish between abandoned and interrupted goals (i.e. it was not sufficient to notice that E had halted his initial goal-directed action). If children differentiate between these causes of E not completing his initial action, we should expect their helping behaviour to differ between the two experimental conditions. We therefore predicted that children would help place the toy in the initial location more often in the interrupted goal condition than in the abandoned goal condition.

2.2 Method

The hypotheses, sample size, methods, exclusion criteria, analyses, and pilot data were pre-registered before data collection, and can be accessed at: https://osf.io/4k2h9/?view_only=1a1ba0a24a8c4f1fb3ffc98a553d25d1. All aspects of the study were carried out in accordance with the pre-registered protocol unless otherwise stated. The experiment was conducted in accordance with the Declaration of Helsinki, and was approved by the Humanities & Social Sciences Research Ethics Sub-committee (HSSREC) at the University of Warwick.

2.2.1 Participants

Using the *simr* package in *R* (Green & MacLeod, 2016), it was determined that a sample size of 20 would provide power of at least 90% for detecting a medium-sized effect, as observed in a pilot study (see the pre-registration material, https://osf.io/4k2h9/?view_only=1a1ba0a24a8c4f1fb3ffc98a553d25d1), for our primary analysis. After beginning data collection but prior to commencing data coding, we modified the design to counterbalance the last location referred to by E on test trials; this change necessitated an adjustment of the target sample size to 24, because we were then counterbalancing three factors (see section 2.2.3).

Twenty-four participants were included in the final sample (eight females, average age: 26;28, range: 25;02-30;00). Additionally, 19 participants were tested and excluded from final analysis according to pre-registered drop-out criteria (see section 2.2.5). Seven participants were excluded because they helped to place toys in the same box on each test trial, and 12 participants were excluded because they did not complete at least two trials in each condition due to fussiness (seven), shyness (one), or taking too long to help E on too many test trials (four). All participants were recruited from a database of families in the Department of Psychology at the University of Warwick and from nurseries in the surrounding area. The majority of participants came from middle class backgrounds and were Caucasian.

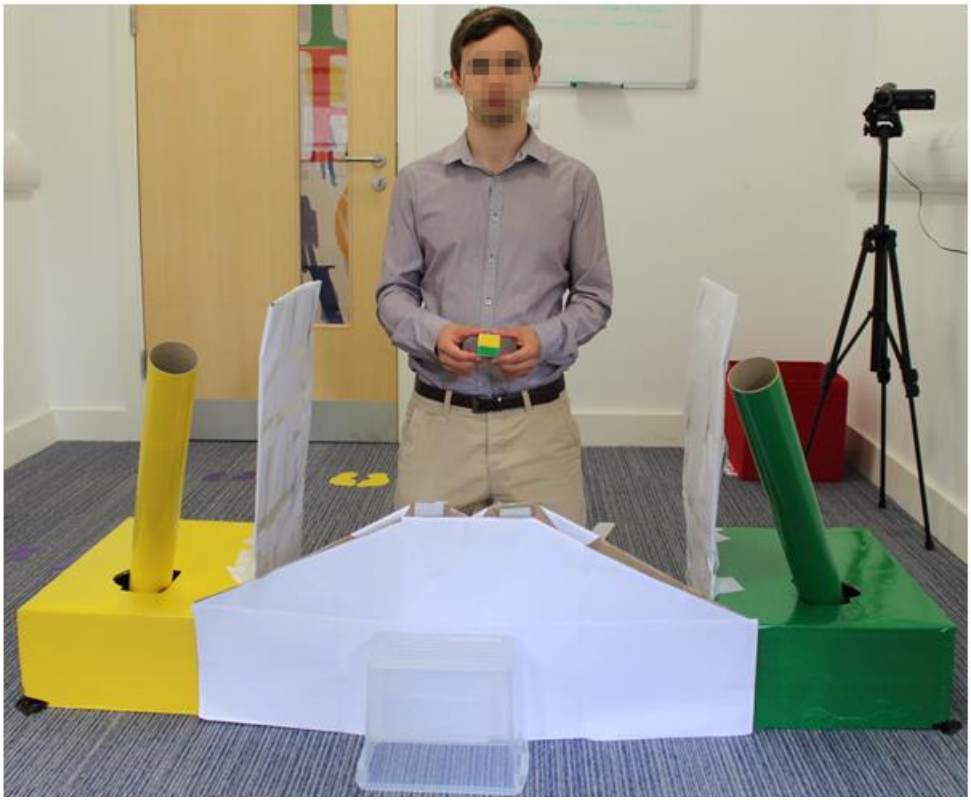
2.2.2 Materials and Apparatus

Participants sat 1.65m away from the apparatus (see *Figure 2.1*) on their caregiver's lap. The apparatus consisted of two coloured boxes, yellow and green (each 35cm x 20cm x 65cm, and each with a white barrier on the inside), with coloured tubes that ran into these boxes (70cm long, 8cm in diameter). Placing the toy in either box required that the toy be dropped into the appropriate tube. The boxes were separated by a white barrier (35cm x 95cm). A small transparent box sat in between this barrier and the participant, which is where E placed the toy after asking for the participant's help. The toys were small cubes (5cm x 5cm x 5cm) that were each equally coloured green and yellow.

A plank ran along the ground between the two boxes, though participants were unable to see this because of the white barrier. The bottom of the tubes rested on this plank, and moving this plank rotated the tubes. This rotation (along with the barriers on the inside of the boxes) prevented E from reaching one of the tubes (see *Figure 2.2*).

Figure 2.1

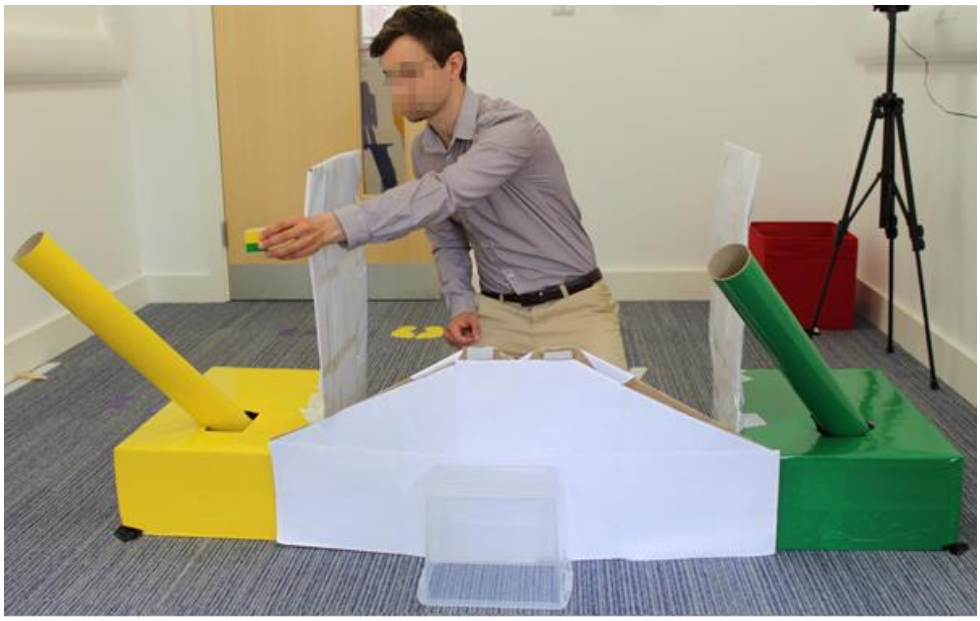
Apparatus From the Participant's Perspective



Note: This figure shows the apparatus and experimenter's position as viewed from the perspective of the participant. The box and tube on the left are yellow, and those on the right are green.

Figure 2.2

Apparatus From the Participant's Perspective After Tubes Have Rotated



Note: This figure shows the apparatus and experimenter's position as viewed from the perspective of the participant, after the tubes have rotated. Tubes could be surreptitiously moved by E.

2.2.3 Design

We used a within-subjects design, with participants performing eight test trials in total. To control for order effects of condition, the eight trials were split into two blocks of four: Block 1 (interrupted goal, abandoned goal, interrupted goal, abandoned goal), and Block 2 (abandoned goal, interrupted goal, abandoned goal, interrupted goal). The order of blocks, E's initial goal location, and whether the final location referred to by E was E's goal by the end of each trial, were counterbalanced.

2.2.4 Procedure

Participants were tested individually in laboratories at the University or else at nurseries. Caregivers gave informed written consent, and participants received a gift for taking part. Sessions lasted approximately 20 minutes.

The apparatus and procedure were validated in a separate study with adults (see the pre-registration material,

https://osf.io/4k2h9/?view_only=1a1ba0a24a8c4f1fb3ffc98a553d25d1). Caregivers were present and played a largely passive role in test trials, with two exceptions: They were instructed to draw the participant's attention to E ("Look at what [E] is doing") if participants were not watching E, and to encourage shy participants to help without giving specific instructions as to which box to help place the toy in ("Can you help [E]? Can you put it where [E] wants it?").

During warm-up participants were acquainted with the boxes and toys, as well as helping to place the toys into the boxes via the tubes. After warm-up, caregivers were asked to sit on a chair with the participant on their lap, whilst E sat in between the two boxes (behind the barrier).

Familiarisation phase

There were six familiarisation trials in which participants were acquainted with helping E, the obstacle that E encounters in the interrupted goal condition (i.e. the rotation of the tubes), and the fact that E sometimes abandons his initial goal. In two of these trials E began to place a toy in one of the boxes, but then decided instead to place the toy in the alternative box. In the next two trials E began to place the toy in one of the boxes, before encountering an obstacle (i.e. the tubes rotated such that he could no longer reach his desired tube). In the two remaining familiarisation trials, the rotation of the tubes moved E's desired tube closer to him (so that children did not always associate the movement of the tubes with E wanting the toy to go into the out-of-reach tube). In all trials, E then asked the participant for help in placing the toy in E's desired box, and placed the toy on the transparent box in front of the white barrier separating the participant from E. The participant was given no specific feedback about where they placed the toy; wherever they placed the toy, E thanked them and clapped, and asked them to sit with their caregiver again.

Test Phase

There were eight test trials in total (four per condition). The number of times that E referred to each box, and the time E spent looking at each box, were kept constant in each test trial, though the type of reference,

gesture, and facial expression made towards each box differed depending on whether this was E's current goal. Each test trial consisted of three phases: (i) establishing E's initial goal, (ii) establishing that E is not going to complete his initial goal-directed action, and (iii) participants helping E.

(i). At the beginning of each trial, in both conditions, E used gesture and verbal reference to indicate his initial goal: "Where will I put it [the toy]? In the green box [or yellow; this was counterbalanced]…? No [E shakes his head and frowns towards the green box]. I want this in the yellow box [E nods and smiles towards the yellow box]". This was done to ensure that E referred to each box an equal number of times. Once E's initial goal (in this example, the yellow box) was established he began the action of placing the toy in the tube connected to the yellow box.

(ii). In the interrupted goal condition, E's goal-directed action was not completed because the tubes rotated (E rotated them surreptitiously), such that E was unable to reach the tube connected to the yellow box. E unsuccessfully reached towards the yellow tube for several seconds, bumping into the barrier on the inside of the yellow box (see *Figure 2.2*). E then explained that: "I want it [the toy] in the yellow box [E nods and smiles at the yellow box], but I cannot reach now! I can only reach the green box, but I do not want it there [E touches the green box with the toy, and shakes his head and frowns]."

In the abandoned goal condition, E did not complete the initial action because he changed his mind: "Actually, I do not want this in the yellow box anymore [E frowns and shakes his head at the yellow box]. Ah, now I want this in the green box [E touches the green box with the toy, and nods and smiles]." In this condition, the tubes did not rotate.

(iii). Both conditions ended with E saying: "You can walk anywhere. Can you help me put it where I want it?". E placed the toy onto the transparent box that is in front of the white barrier separating the participant and E, at which point the caregiver was instructed to set the participant down directly in front of them.

If the participant did not initiate the helping behaviour, E repeated this request. E then signalled to the caregiver to encourage the participant to come forward. Throughout this period E maintained eye contact with the participant and smiled encouragingly. The participant was given no specific feedback about where they placed the toy - E thanked the participant (regardless of which box the participant helped place the toy into), and encouraged them to return to their caregiver's lap. Once they had done so and settled down, the next trial commenced.

2.2.5 Coding and Drop-out Criteria

For each trial, we coded where participants helped to place the toy (initial location vs alternative location). We then coded how participants helped E, distinguishing between two types of helping behaviour: placing the toy in one of the boxes, or moving one of the tubes closer to E. If participants exhibited both of these behaviours in a single test trial, we recorded the first of these behaviours. Since the latter behaviour occurred only five times in total, and since we consider the two helping types to be equivalent, we collapsed these two helping types for all subsequent analyses. We also coded response type (i.e. whether participants' helping behaviour was correct or incorrect): In the interrupted goal condition the correct helping location is the initial location, whilst the correct helping behaviour in the abandoned goal condition is the alternative location.

All sessions were recorded using digital video recorders. Coding was carried out by a research assistant who was naïve to the hypotheses of the study. The coder assessed helping location (initial location vs alternative location), helping type (placing the toy in one of the boxes vs moving one of the tubes closer to E), response type (whether participants' helping behavior was correct or incorrect), and trial exclusion (whether individual trials should be dropped from analysis for any of the reasons listed below). A second research assistant who was naïve to the hypotheses of the study coded a random 6 participants (25%) for reliability. Using the *Kappa.test* function of the *R* package *fmsb* (Nakazawa, 2019), coders were found to be in almost perfect agreement over helping location (judgements matched

95% of the time, κ = 0.90 (95% CI: 0.81, 1.00), $p < .001$) and response type (judgements matched 95% of the time, κ = 0.87 (95% CI: 0.78, 1.00), $p <$ .001), substantial agreement over helping type (judgements matched 93% of the time, κ = 0.66 (95% CI: 0.40, 0.92), p = .001), and moderate agreement over trial exclusion (judgements matched 91% of the time, κ = 0.58 (95% CI: 0.32, 0.86), p = .001). We believe that the findings regarding trial exclusion and helping type are a result of Cohen's kappa being an unreliable statistic for rare observations (Feinstein & Cicchetti, 1990; Viera & Garrett, 2005; Yarnold, 2016), given the distribution of the *helping type* and *trial exclusion* variables (i.e., that participants helped by placing the toy in one of the locations in almost all trials, and that most trials were not dropped) and the high percentage of agreement between coders for these variables.

Participant Drop-out Criteria

Pilot testing revealed that some participants had a strong preference for one of the colours, and always placed the toys in the box of that colour. For this reason we decided to exclude participants who helped place the toy in only one of the two boxes on all test trials (indicating that they have a preference for a particular colour which overrides any motivation to help E). We also excluded participants who did not complete at least two trials out of four in each condition.

Trial Exclusion Criteria

A trial was excluded if any of the following criteria were met: (1) If caregivers instructed the participant as to which box to help place the toy into, (2) if there was a delay longer than 15 seconds between E finishing his script (or the participant being let go by their caregiver, whichever happened last) and the participant helping to place the toy into either one of the boxes, (3) if participants were not watching when E halted his goal-directed action or during the experimental manipulation (i.e. the dialogue specified in (ii)). If this occured, E called the participant by their name and tried again. If the participant still did not pay attention, the trial was excluded, and E moved on to the next trial.

2.3 Results

2.3.1 Data Screening

From the responses of the 24 participants included in our analysis, 18 test trials were excluded (eight in the abandoned goal condition) due to a delay of longer than 15 seconds between E's initial request for help and participants' helping. This left 174 test trials for further analysis (88 in the abandoned goal condition).

Participants helped E by placing the toy into one of the boxes in 169 trials (87 in the abandoned goal condition), and in the remaining five trials participants helped by moving one of the tubes closer to E (one in the abandoned goal condition). We collapsed these helping types for further analysis.

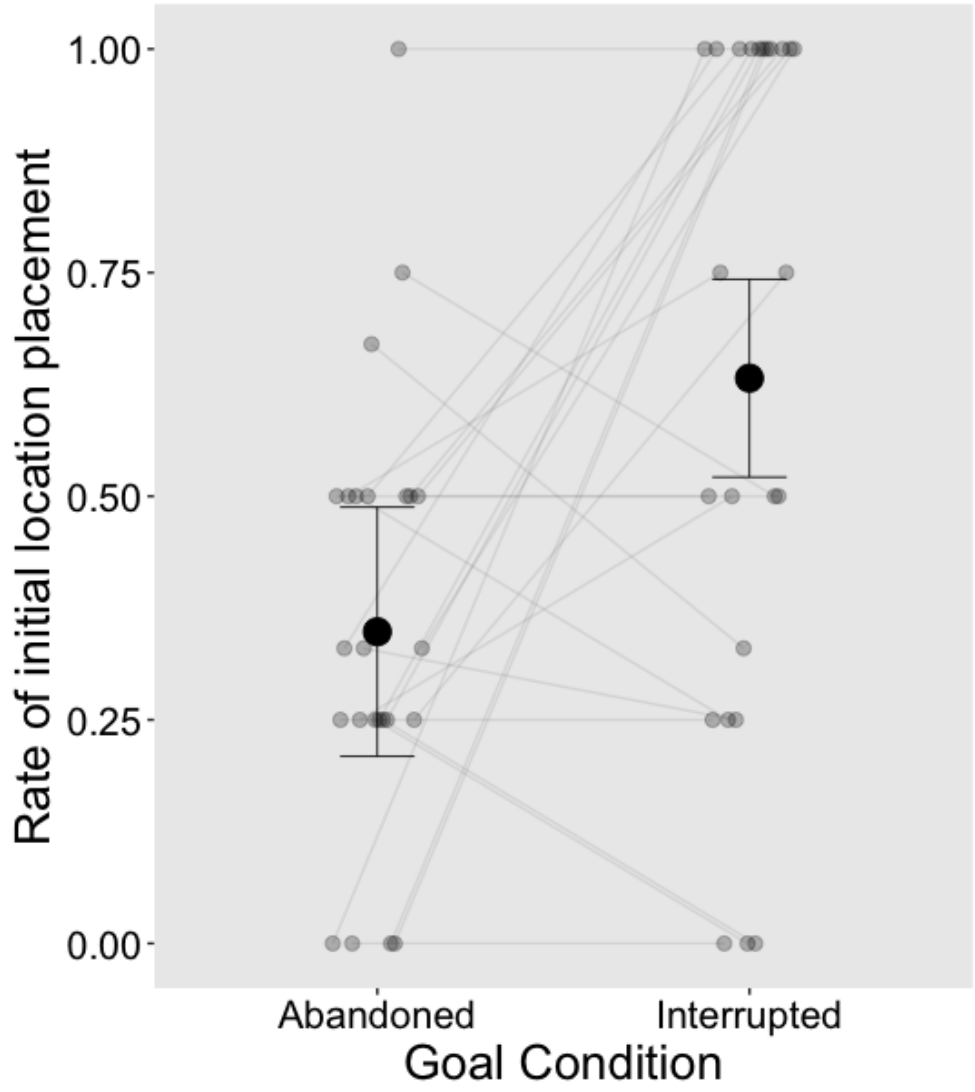
All subsequent analysis was conducted in R (R Development Core Team, 2018; Revelle, 2018). See the supplementary materials (https://osf.io/swtkq/?view_only=6d1ee903aa6b46ca97635c8cbf3660bc) for this experiment's data and *R* code.

2.3.2 Initial vs Alternative Goal Location

To investigate whether participants differentiated between abandoned and interrupted goals we used helping location - whether children helped to place the toy in the location that E was initially trying to place it, or in the alternative location. Children helped place the toy in the initial location on 35% of abandoned goal condition-trials, and in 63% of interrupted goal condition-trials (see *Figure 2.3*).

Figure 2.3

Proportion of Trials on Which Participants Placed the Toy in E's Initial Goal Location



Note: This figure shows the proportion of trials on which participants helped E to place the toy in the initial location, with 95% confidence intervals of the means adjusted for within-subject design (Cousineau, 2005; Loftus & Masson, 1994; Morey, 2008). Jittered dots represent individual participants' performances in a given condition, with light grey lines connecting each participant's performance across conditions.

To test whether the cause of E not completing his initial goal-directed action had an effect on helping behaviour, we used the function

glmer of the *R* package *lme4* (Bates et al., 2015) to create a generalized linear mixed model (GLMM) with a binomial error structure. Participants' helping location (initial location vs alternative location) was the dependent measure. Condition (abandoned vs interrupted) was the only test predictor, and the model controlled for random intercept of participants. We initially included the random slope of participant, but we removed this term because of singularity in the model (Barr et al., 2013; Bates et al., 2015; Singmann & Kellen, 2019). Our pre-registered analysis also included slope and intercept of trial number as random effects, but we realised that this was not necessary given the counterbalanced design of this experiment.

Using the *anova* function, we compared the full model with a null model that was identical, except that it did not include the test predictor (condition). The results indicated that the full model was a significantly better fit for the data than the null model (χ^2 (1) = 15.1, $p < .001$). The odds of participants helping to place the toy in the initial location were over 3 times larger (*OR* = 3.69, *estimate* ± *SE* = 1.30 ± 0.35, $p < .001$) in the interrupted goal condition-trials than in the abandoned goal condition-trials (see *Table 2.1* for a summary of results), indicating a medium effect size (Chen et al., 2010; Chinn et al., 2000; Maher et al., 2013).

Table 2.1

Summary of the Fixed Effects of the Mixed Logistic Model for Helping Location

Predictor	Coefficient	Standard error	Wald Z	*p*
Intercept	-0.70	0.29	-2.40	.016*
Condition	1.30	0.35	3.72	.0002***

Note: log-likelihood = - 104.5

*** < .001

* < .05

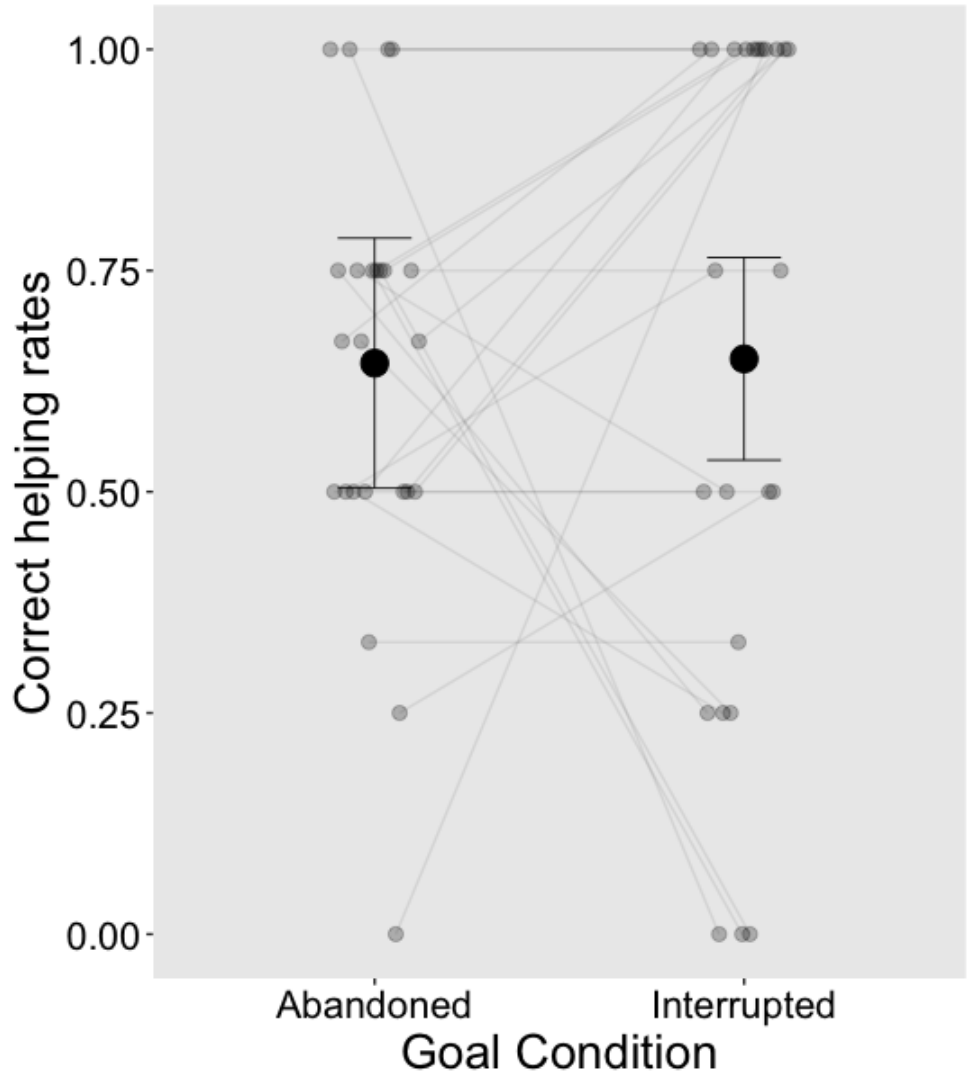
This indicates that participants differentiated between goal abandonment and goal interruption in their helping behaviour. However, the results of the foregoing analysis do not rule out the possibility that the effect found here may be driven by participants being highly accurate in their helping in only one of the conditions. We therefore also tested whether participants are more likely to help E correctly in one of the conditions as compared to the other.

2.3.3 Correct vs Incorrect Helping Behaviour

Children correctly helped E in 65% of abandoned goal condition-trials, and 63% of interrupted goal condition-trials (see *Figure 2.4*). We used our second measure, response type (whether children helped E correctly or incorrectly in each trial) to construct a GLMM identical to that used for our primary analysis, except that response type (correct vs incorrect) was the dependent variable. We compared this to a null model (identical except for the removal of the test predictor, condition) using the *anova* function. This ANOVA showed that the full model was not a significantly better fit for the data than the null model (χ^2 (1) = 0.04, p = .840). This indicates that participants were not significantly more likely to help E correctly in one condition as compared to the other.

Figure 2.4

Proportion of Trials on Which Participants Correctly Helped E



Note: This figure shows the proportion of trials on which participants correctly helped E, with 95% confidence intervals of the means (adjusted for within-subject design). Jittered dots represent individual participants' performances in a given condition, with light grey lines connecting each participant's performance across conditions.

Finally, to determine whether participants helped correctly above chance (50%) in each condition, we compared the proportion of trials with correct helping against 50% with two one-sample Wilcoxon signed-rank tests using the functions *wilcox.test* and *wilcoxonOneSampleR* of the *R*

package *rcompanion* (Mangiafico, 2016, 2018). Wilcoxon tests were used because proportion violates assumptions for t-tests (equal variance across conditions and normal distribution). The median rate of correct helping was 67% in the abandoned goal condition and 75% in the interrupted goal condition. Correct helping differed significantly from chance in both the abandoned goal condition ($V = 125.5$, $p = .010$, $r = 0.81$) and the interrupted goal condition ($V = 150$, $p = .040$, $r = 0.63$).

2.4 Discussion

We tested whether 2-year-old children differentiated between interrupted and abandoned goals in an instrumental helping task. The results reveal that children's helping behaviour differed significantly between the abandoned goal condition and the interrupted goal condition, and that participants correctly helped the experimenter above chance in both conditions. They did not find either condition easier than the other. These results support the hypotheses that 2-year-olds can represent others' goals as changing prior to completion, and that they understand goal abandonment and goal interruption.

Importantly, these results cannot be explained by learning effects from the first condition that children encountered on test trials (Brooks, 2012) because we implemented a counterbalanced block design for the eight test trials and because participants were not provided with any directional feedback. Similarly, because we counterbalanced which box E referred to last, we can rule out the possibility that children were simply placing the toy in the box that E last referred to before asking for help. Moreover, the difference in children's behaviour between the two conditions cannot have been due to children's colour preferences either (Marshall et al., 2006), since those children who helped to place all of the toys in only one of the boxes were excluded prior to analysis.

There are two notable areas for concern about the present study that are important to address. One area relates to participant and trial exclusion criteria, given the relatively high participant drop-out rate. For participant exclusion criteria, we decided to exclude participants who only placed toys

into one of the two boxes and participants who did not complete at least two trials per condition. We excluded participants who did not complete at least two trials out of four in each condition in order to ensure that we had at least two trials to take the average of for our comparisons to chance (section 2.3.3), as a matter of convention. Our motivation for excluding participants who placed toys into only one of the two boxes came from piloting, during which some participants were simply placing toys into the box of their preferred colour rather than helping E. This was made clear by participants' behaviour (e.g., not looking at E when they came forward to place the toy into a box) and their verbalisations (e.g., turning to their caregiver after placing the toy into a box while excitedly referring to the box's colour). This would be problematic because if participants were placing the toy into their preferred box rather than helping E then their behaviour would not be sensitive to the manipulation at the heart of this experiment. Our results would therefore not provide insight into whether children distinguish between abandoned and interrupted goals, because placing the toy into one's preferred box does not require tracking E's goal and responding appropriately. We did not expect the introduction of these criteria to be problematic because we did not expect exclusions to differ between conditions (the colour of our stimuli did not differ between conditions, and each participant had 4 test trials per condition).

For trial exclusion criteria, consider the exclusion of trials in which participants took too long to place the toy into a box. Specifically, our criterion was that trials would be excluded if participants placed the toy into a box more than 15 seconds after E had finished his script or the participant was let go from the caregiver's lap (whichever happened last). There were several sources of motivation for introducing this criterion. Firstly, this criterion would abate concerns that participants would not remember E's goal and the experimental manipulation (part (ii) of the procedure) by the time that they came to place the toy into a box. In this respect, it is important to note that 15 seconds after E had completed his script (or participants were let go by their caregiver) is *not* 15 seconds after the experimental

manipulation. After the experimental manipulation, E must finish his script and then place the toy on the transparent box, and caregivers then remove participants from their lap. This criterion therefore means that there is 20-30 seconds for participants to place the toy into one of the goal containers from the point of experimental manipulation (with the precise timing depending on how quickly E finishes his script and caregivers place the participant on the ground). This exclusion criterion has also been implemented in past helping studies (Hepach et al., 2012; Kenward & Gredebäck, 2013), and *if* children are going to instrumentally help others then they typically do so within the first 30 seconds of being in a position to do so (Hepach et al., 2012; Warneken & Tomasello, 2007). Finally, we expected that trials excluded due to this criterion to be equally distributed across conditions (and indeed, we dropped eight trials in the abandoned goal-condition and ten in the interrupted goal-condition).

While these participant and trial exclusion criteria may seem overly conservative, it should be noted that the results of our analysis did not change when we included participants and trials which were excluded from our original analysis.

The second main area for concern about the present study lies the procedure, and specifically, that E explicitly tells participants where he wants the toy to go. That is, after E either abandons his initial goal or the goal is interrupted by the movement of the tubes, E explicitly states which location he wants the toy to go. The concern would be that because E has explicitly told participants where he wants the toy to go, participants could succeed in the task of correctly helping E by only behaving according to E's request. That is, participants could theoretically succeed at the task without differentiating between interrupted and abandoned goals, as adhering to the request following E's explicit statement does not require understanding whether E's initial goal was abandoned or interrupted.

There are several reasons for being sceptical of this interpretation of our results. Firstly, if children were indeed helping E simply on the basis of his most recent utterance, then the task would be extremely straightforward;

children would only need to place the toy into the box that E had just said he wanted the toy to go. Due to the simplicity of this task we would expect children to perform extremely well, potentially helping correctly in close to 100% of test trials in both conditions. While children were performing above chance in both conditions, they were far from 100% correct helping in both conditions. In addition, this interpretation of our results would require that participants either (a) do not pay attention to parts (i) and (ii) of the procedure, or (b) that they ignore parts (i) and (ii) of the procedure and only behave on the basis of E's final explicit statement. Given that children spontaneously track others' goals (Csibra, 2008; Jin & Song, 2017; Woodward, 2009) it is unlikely that (a) is true, and thus (a) would require more empirical support. The challenge for (b) is that children take into account previous goal-directed behaviours in determining one's current goals (Csibra, 2008; Woodward, 2009), including in helping contexts (Hobbs & Spelke, 2015). That is, children do not typically interpret one's current goal-directed behaviours in isolation. This suggests that children are unlikely to ascribe goals to others, and help others, *only* on the basis of their most recent goal-directed behaviour. In sum, then, this interpretation is inconsistent with our results and with past findings about the spontaneity and sophistication of goal tracking in early childhood.

Despite these reasons for being sceptical that children were simply responding to E's final explicit utterance, this interpretation of our results remains a theoretical possibility, and it is therefore important to spell out how this concern might be addressed in future research. The challenge lies in indicating E's goal status non-verbally. Fortunately, E's goals in the present study were indicated by both verbal and non-verbal cues. This means that the present concern could be addressed in a future study by implementing the present study's procedure but with E's speech removed from parts (i) and (ii) of the procedure. In part (i) E could establish his initial goal by smiling and nodding towards his initial goal location (and frowning and shaking his head towards the alternative goal location). The experimental manipulation in part (ii) could also be implemented without

E's speech. In the interrupted goal-condition E could still reach unsuccessfully towards the initial goal location after the tubes have been rotated, indicating that he maintains this goal. In the abandoned goal-condition E could still indicate that his initial goal was abandoned by halting his goal-directed action, and then shaking his head and frowning towards the initial goal location, before turning to smile and nod at the alternative goal location. Part (iii) of the experiment could then be conducted as in the present study: E asks the participant to help place the toy where E wants it to go. In this way, one could investigate whether children distinguish between abandoned and interrupted goals *without* E explicitly stating his goal before asking the participant to help him.

A potential related concern about the present study is that children might simply be responding to E's most recent goal-cues. That is, even if E's final explicit utterance is removed, children may simply be helping E on the basis of his most recent behaviour (e.g., attempting to reach for the out-of-reach tube in the interrupted condition, or nodding at the alternative location in the abandoned condition). The concern with this possibility is that it would not require that children understand goal abandonment or interruption; children could help E correctly by simply responding to his most recent behaviours.

There are again similar reasons to be sceptical about applying this interpretation to our results. Again, if children were simply responding to E's most recent goal-cues at the point at which E asks for help, then the task would be quite straightforward. We would therefore expect rates of correct helping to be close to 100%, which is not what the data indicates. This interpretation would also require that children ignore E's past goal-directed behaviours or at least that they do not take it into account in determining E's current goal, both of which are inconsistent with past findings about goal tracking in early childhood (Csibra, 2008; Hobbs & Spelke, 2015; Woodward, 2009). This interpretation of events is therefore inconsistent with our present findings and past findings about goal tracking in early childhood.

One might be tempted to address this concern by adjusting the present study's procedure such that E no longer provides cues as to his goal status, but removing such cues would mean that the present study no longer addresses the research question as to whether children distinguish between abandoned and interrupted goals. The most obvious way to address this concern might be to remove all cues of E's goal after E halts his initial goal-directed action. For instance, in the interrupted condition, E's initial initial goal-directed action could still be halted by the tubes moving, but then all cues as to E's goal after this could be removed (e.g., E would no longer reaches for his goal-location). And in the abandoned condition, E could start his initial goal-directed action but then simply halt it. The problem with implementing this change is that whether E maintains his goal or not (i.e., whether the goal is abandoned or merely interrupted) cannot be separated from E's cues; simply halting one's goal-directed behaviour does not indicate whether the goal is abandoned or interrupted. For instance, suppose that I were attempting to drive to a friend's house, and I encounter an obstacle (e.g., my car is broken). I might therefore halt my current goal-directed action (e.g., I might get out of my car), but this by itself is consistent with me abandoning the goal (deciding to no longer go to my friend's house) or maintaining my goal despite the interruption (e.g., I may decide to travel to visit my friend with public transport). My goal status cannot be identified independently of my further cues. Indeed, this is precisely why studies that require children to understand that the target agent's goal is interrupted do not *only* show participants the obstacle, but they also provide further cues to indicate that the target agent is maintaining their goal, such as the target agent bumping into an obstacle (Kenward & Gredebäck, 2013) or reaching for an out-of-reach item (Warneken & Tomasello, 2006). Investigating children's helping behaviour following E simply halting his goal-directed action would therefore not answer the question as to whether children distinguish between interrupted and abandoned goals, because in order for children to identify whether the goal

is abandoned or interrupted there must be *some* indication of E's goal status from E.

The current study contributes to two distinct bodies of previous research. First, this study builds upon research investigating goal tracking in early childhood. This research typically implements scenarios in which the target agent's goals remain constant over individual test trials (Cannon & Woodward, 2012; Csibra, 2008; Liu & Spelke, 2017; Southgate & Vernetti, 2014; Woodward, 1998). Our findings extend that research by changing the target agent's goals not only between test trials, but also sometimes *within* test trials.

Furthermore, our findings contribute to research on early prosocial behaviour. Most previous research probing the emergence of instrumental helping behaviour presents children with only one goal affordance (i.e., only one possible means by which the child could help the experimenter) during test trials, and children are typically faced with a helpee whose goals do not change during test trials (e.g., Barragan & Dweck, 2014; Cirelli et al., 2016; Dahl et al., 2017; Dunfield & Kuhlmeier, 2010; Hepach et al., 2017; Kenward & Gredebäck, 2013; Over & Carpenter, 2009; Warneken & Tomasello, 2006) - though, as noted earlier, 3-year-olds help in a manner that is sensitive to the distinction between an adult's proximal and distal goals (Martin & Olson, 2013). Our findings extend this research by revealing that 2-year-olds are able to help appropriately in dynamic social settings in which an agent's goals sometimes change.

Moreover, the current study provides new avenues for research investigating the cognitive and motivational underpinnings of prosocial behaviour in early childhood. For example, our findings are relevant for research exploring the hypothesis that the identification of an agent's goal leads young children to take up that goal as their own, and accordingly to be motivated to complete unfinished actions (Michael & Székely, 2019; Paulus, 2014). This is because this hypothesis generates the prediction that "an infant would continue helping ... if an agent were to become distracted, lost interest or otherwise abandon the goal" (Michael & Székely, 2019,

p.181). By establishing that 2-year-olds understand goal abandonment, the current study provides a basis for testing this prediction. Such studies would clarify whether helping behaviour in early childhood can be interpreted as a behaviour driven by the child's altruistic motivation (Warneken & Tomasello, 2006) or as a behaviour motivated by a goal that has now become the child's own.

Our findings also raise new questions about *how* children track the goals of others. As in most previous research on goal tracking, this study used multiple cues to indicate E's goals: gestural cues (Sodian & Thoermer, 2004), verbal cues (Jin & Song, 2017), goal-directed actions (Kim & Song, 2015; Luo, 2011; Sommerville et al., 2005), gaze and emotional cues (Phillips et al., 2002), and external constraints (Csibra, 2008; Liu & Spelke, 2017). Future research should investigate how these different cues are integrated, and which cues children prioritise at different ages.

A further avenue for future research is how children's understanding of goal abandonment relates to social learning. Children infer the value of goals on the basis of the costs that agents are willing to invest in the pursuit of those goals (Jara-Ettinger et al., 2015; Liu et al., 2017), but what do children infer about the value of goals which agents abandon? For example, they may interpret goal abandonment as stronger evidence that the goal is not worth pursuing, as compared to the case in which the goal was never adopted in the first place. To what extent do children generalise such inferences about the values of goals towards other goals of similar types (Csibra & Gergely, 2009; Gergely & Csibra, 2005; Martin et al., 2017; Spaepen & Spelke, 2007)? For example, if children observe an agent abandoning the goal of acquiring an apple and infer that that goal is not valuable, will they extend this to the goal of acquiring other apples, or other fruits? And do children infer long-standing psychological states, e.g., preferences, on the basis of goal abandonment (Hamlin et al., 2007; Luo et al., 2017; Michael & Christensen, 2016; Sommerville & Crane, 2009)?

2.4.1 Conclusion

For the development of flexible social cognition and prosocial behaviours (such as instrumental helping), it is important to be able to differentiate between abandoned and interrupted goals. The current findings provide the first preliminary evidence that children as young as 2 years of age have these abilities, although it is important that further studies replicate these results and address concerns raised in section 2.4.

The present chapter contributes to this thesis by uncovering preliminary evidence that 2-year-olds are sensitive to goal abandonment in an instrumental helping context. By doing so, this chapter lays the groundwork for the experimental manipulation that lies at the core of the subsequent chapter.

Chapter 3

Goal Slippage: Two-Year-Olds Complete Others' Unfinished Goals

Children are willing and able to help others achieve their goals from the second year of life (Dahl & Brownell, 2019; Warneken & Tomasello, 2006). Several proximate psychological mechanisms have been proposed to account for this behaviour. The present study investigates whether a previously untested mechanism, goal slippage (Michael & Székely, 2019), may underpin instrumental helping in early childhood. Participants (24- to 30-month-olds) observed an experimenter filling a container with toys. We manipulated whether the experimenter's goal of filling the container was completed before he left the apparatus. We measured whether participants (n = 34) were more likely to place a toy into E's goal container if E abandoned his goal before its completion, as compared to if E had completed his goal before leaving the apparatus. The results showed that participants were more likely to place the toy into E's container when this constituted completing E's abandoned goal, as compared to if E's goal had already been achieved. These results support the hypothesis that the goal slippage mechanism may underpin at least some instances of instrumental helping in early childhood.

3.1 Introduction

Humans lead uniquely rich and complex social lives (Gilbert, 2006; Tomasello, 2014, 2019). One prominent aspect of our social lives is that we are particularly prosocial towards other members of our own species (Jensen, 2016; Tennie et al., 2016). In particular, relative to other animals, humans are better able and more willing to help others achieve goals (i.e., outcomes to which their actions are directed) (Barnes et al., 2008; Bshary & Raihani, 2017; Jensen et al., 2014; Melis, 2018; Sznycer et al., 2019). This willingness and ability to instrumentally help others emerges early in life, and is one of the earliest developing forms of prosocial behaviour (Callaghan & Corbit, 2018; Cirelli et al., 2016; Dunfield, 2014; Dunfield & Kuhlmeier, 2010; Hobbs & Spelke, 2015; Paulus, 2019). For instance, from

the second year of life children spontaneously give out-of-reach objects to agents who are reaching for those objects (Warneken & Tomasello, 2006).

It is important to identify the proximate psychological mechanisms that underpin instrumental helping in early childhood because it is one of the earliest developing forms of prosocial behaviour (Callaghan & Corbit, 2018; Dunfield, 2014; Hammond, 2014; Paulus, 2019). Understanding the cognitive mechanisms that underpin this behaviour may therefore provide insight into the development of further kinds of prosocial behaviour that emerge later in development (Barragan et al., 2020; Dahl & Brownell, 2019; Köster et al., 2019; Paulus, 2018; Schuhmacher et al., 2018). In addition, uncovering the cognitive mechanisms that underpin instrumental helping may shed light on the evolutionary distinctiveness of human prosocial behaviour relative to other animals. While instrumental helping may not be uniquely human (Barnes et al., 2008; Bshary & Raihani, 2017; Warneken & Tomasello, 2006), it does seem to be more common in humans than non-human animals (Hepach et al., 2020; Melis, 2018; Tennie et al., 2016).

Several non mutually exclusive mechanisms have been proposed to account for instrumental helping in early childhood (Hammond & Brownell, 2018; Michael & Székely, 2019; Paulus, 2014; 2018). The two hypotheses that have received the most attention in the empirical and theoretical literature are the psychological altruism hypothesis and the social interaction hypothesis. Under the psychological altruism hypothesis, children are motivated to instrumentally help others achieve goals because children are naturally inclined towards caring about the well-being of others (Warneken, 2015; Warneken & Tomasello, 2006, 2009). Under this hypothesis, children are motivated to help others because contributing to another agent's goal is seen as a means of improving the helpee's well-being. Under the social interaction hypothesis, children are motivated to help others not for the sake of the other agent's well-being, but rather because interacting with others is an intrinsically enjoyable experience for children (Carpendale et al., 2013, 2015; Dahl, 2015; Dahl & Brownell, 2019). This motivates children to contribute to others' goals because doing so will increase the chances that

the helpee will interact with the child. While there are further mechanisms hypothesised to underpin instrumental helping (see section 1.1.2), these are the mechanisms that feature most prominently in the literature.

A further hypothesised mechanism for instrumental helping in early childhood is goal slippage (Barresi & Moore 1996; Eisenberg et al., 2016; Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2019; Paulus, 2014). Under the goal slippage hypothesis children are motivated to contribute to another agent's goal because of a cognitive tension that is generated when one identifies that a goal will remain incomplete; for instance, one may find it frustrating or irritating to see someone struggling to complete a task, or leaving a task incomplete. The hypothesis is that completing the target goal will relieve this cognitive tension, and relieving this tension is what motivates helping behaviour. In this way the goal can be understood as 'slipping' from the helpee to the observer, and from perception to action (Hornstein, 1972; Lewin, 1951; Michael & Székely, 2018, 2019). In this sense, goal slippage involves children taking on others' goals for themselves, and thereby being intrinsically motivated to bring about the relevant outcome.

There are several factors that motivate the study of the goal slippage hypothesis within the context of instrumental helping in early childhood. As discussed in section 1.1.4, there are several findings that goal slippage is particularly well-placed to account for. For instance, goal slippage can straightforwardly account for the finding that children are motivated to help non-human agents (Martin et al., 2020), including mere geometric shapes (Kenward & Gredebäck, 2013). The goal slippage hypothesis is also well placed to account for the emergence of the bystander effect (Plötner et al., 2015), and the fact that children do not initially preferentially help more needy agents over less needy agents (Paulus, 2019).

In addition, the predictions generated by the goal slippage hypothesis are consistent with the findings of many past instrumental helping studies (also discussed in section 1.1.4). Many instrumental helping studies proceed by the experimenter encountering some obstacle and thus requiring the help

of a further agent to complete their goal (Hepach et al., 2020; Svetlova et al., 2010; Wanrken & Tomasello, 2006). While helping in these scenarios is often interpreted as supporting the psychological altruism hypothesis, children's behaviour in these circumstances is also consistent with the goal slippage hypothesis. Broadly speaking, adopting another agent's goal for oneself (and thereby being internally motivated to contribute to that goal because it is now one's own goal) constitutes helping behaviour. For instance, if an agent accidentally dropped an object, goal slippage would also predict that children would help by providing the agent with the object, thereby completing the agent's goal of retrieving the object (Warneken & Tomasello, 2006). In addition to children's behaviour, the goal slippage hypothesis is also consistent with children experiencing physiological arousal when they observe another agent requiring help (Hepach et al., 2012). Specifically, goal slippage predicts that a tension would be generated when one observes another agent failing to achieve a goal, which would be relieved only with completion of the goal. The physiological arousal that children experience in such circumstances would be explained by such tension. If a direct test of the goal slippage hypothesis yielded empirical support, this may therefore prompt a re-evaluation of many past instrumental helping experiments.

Another important set of findings that the goal slippage hypothesis can account for is that children preferentially help agents to achieve their goals instead of helping with the *means* by which agents are attempting to achieve their goals (Buttelmann et al., 2009; Hepach et al., 2020; Knudsen & Liszkowski, 2012, 2013; Paulus, 2019). For example, if an agent is attempting to achieve a goal by using a tool that they mistakenly believe will enable them to achieve the goal, children will instead help by providing the agent with a different tool that does enable them to achieve their goal (Martin & Olson, 2013). As long as children are able to identify the agent's goal, then the goal slippage hypothesis can account for these findings, as goal slippage predicts that children would adopt these goals for themselves. This explanation does not require that children as young as 12 months old

make inferences about whether the goal is beneficial for the agent (as the psychological altruism hypothesis does) or that achieving the goal will encourage the agent to socially interact with the child (as the social interaction hypothesis does).

The finding that children preferentially help others to achieve their goals, rather than to complete their actions, is also important because it provides an answer for one of the questions about goal slippage that could be raised at this point: why think that children are motivated to instrumentally help others by a goal-specific cognitive mechanism rather than an action-specific cognitive mechanism? That is, why think that there is a cognitive mechanism that motivates goal completion-behaviour rather than action completion-behaviour? The answer to this is, simply, that children's helping behaviour seems to focus on helping others to achieve the outcomes to which their actions are directed (i.e., the state of the world to which helpee's actions are directed) rather than others' actions themselves (i.e., the specific means by which others are attempting to achieve their goals). This suggests that helping behaviour is motivated in terms of outcomes rather than specific actions.

The goal slippage mechanism would also carry several evolutionary benefits. For instance, a mechanism that motivates children to contribute to others' goals may support social learning and affiliative behaviour. This is because contributing to others' goals may help children to better understand what goals are useful to bring about and how to bring about such goals (Michael & Székely, 2019; Over, 2016; Yu & Kushnir, 2019). Helping others to achieve their goals may also elicit reciprocal behaviour from others in future, thus fostering cooperation and interdependence between agents (Bshary & Raihani, 2017; Melis, 2018; Roberts, 2005).

As the goal slippage mechanism is yet to be directly tested, there are several questions surrounding it that will require further empirical research to answer. One of these is the degree to which goal slippage would interact and mesh with other sources of motivation. Suppose that an agent had a goal that was ultimately bad for their well-being, for example, a friend wanting to

acquire cigarettes but being unable to do so because they have forgotten their wallet at home. Goal slippage would predict that you would contribute to your friend's goal (by either providing your friend with the money to purchase cigarettes or purchasing them yourself), whereas psychological altruism would predict that you would want to avoid contributing to this goal because smoking is unhealthy and therefore detrimental to your friend's well-being. As goal slippage is yet to be tested in such a scenario, it is not yet clear in what contexts the motivation to contribute to this agent's goal (due to goal slippage) might mesh with motivations to act in the interests of the agent's well-being (and consequently, to *not* contribute to their goal).

The other main question for goal slippage focuses on the degree to which it accounts for past instrumental helping findings. That is, when goal slippage and other mechanisms, such as altruism or the social interaction hypothesis, predict the same behaviour, to what degree is goal slippage underpinning the behaviour? For example, suppose that an experimenter is attempting to stack a book onto a pile of books and accidentally drops the book such that it is out of reach (Warneken & Tomasello, 2006). Children's subsequent behaviour (i.e., contributing to the experimenter's goal by either returning the book to the experimenter or else placing the book onto the pile) is predicted by all three mechanisms; children may help here because they care about the experimenter, or because they want to complete the experimenter's goal, or because they believe that contributing to the experimenter's goal will yield social interaction. In the absence of research directly addressing this question it is difficult to identify the degree to which goal slippage underpins this behaviour. The proposal of this chapter is certainly not that goal slippage is the only explanation for *all* instances of helping behaviour (see section 1.1.4 for more discussion about findings that goal slippage can and cannot account for). It is worth noting that this same question applies to all proposed mechanisms for instrumental helping, i.e., independently of goal slippage, there is a question about the degree to which

each individual proposed mechanism for instrumental helping underpins past findings.

For both of these questions there is simply not yet the empirical research required to provide answers. The first step is simply to test for evidence of the goal slippage hypothesis in an instrumental helping context in early childhood. This is the aim of the present study.

3.1.1 Present Research

The aim of the present study is to test the goal slippage hypothesis: that when an agent observes another agent acting to bring about a particular goal, the goal might slip from observation to action such that the observer takes on the goal for themselves. The observer would then be internally motivated to pursue the goal, even if the original actor has lost interest in the goal.

In order to test this hypothesis with young children we developed an experimental paradigm based on past instrumental helping studies (Warneken & Tomasello, 2006; Dunfield & Kuhlmeier, 2010; Hobbs & Spelke, 2015). Given that instrumental helping paradigms typically focus on children within the second year of life, and there is some preliminary evidence supporting the hypothesis that children understand goal abandonment at 24-30 months of age (Chapter 2), we tested the goal slippage hypothesis with a population of 24- to 30-month-olds.

We implemented two changes from past instrumental helping studies which enable the goal slippage hypothesis to generate unique predictions about children's behaviour in comparison to the psychological altruism hypothesis and the social interaction hypothesis. The first change is to have the experimenter's (E's) goal be incomplete because it was abandoned, rather than because it was interrupted (as in past instrumental helping studies). The psychological altruism mechanism should not activate if the experimenter abandons a goal that he was recently pursuing. This is because, under psychological altruism, children help the experimenter achieve a goal *because* the experimenter has that goal. Psychological altruism therefore does not predict that children will help bring about an

abandoned goal. Goal slippage predicts that children will still help bring about an abandoned goal, however, because under this hypothesis children will have adopted the goal for themselves, and they should be internally motivated to achieve it. If the helpee's goal is abandoned, then, goal slippage generates different predictions than psychological altruism about children's helping behaviour.

The second change is to have E leave when he abandons his goal (i.e., once E abandons his goal, he disengages from the test apparatus and moves elsewhere). If an agent has lost interest in an object and is no longer in close proximity to it, then interacting with that object is unlikely to encourage social interaction with that agent. This means that children will not complete E's goal in order to socially interact with E. Goal slippage, however, predicts that E leaving the apparatus will not affect whether children complete E's goal. This is because once the participant has adopted E's goal, then they will be internally motivated to complete the goal independently of whether E has lost interest in it. Therefore if E leaves the apparatus, then goal slippage generates different predictions than the social interaction hypothesis.

In all test trials E has the goal of filling one of three possible containers to a specified level (indicated with a red line) with toys. E leaves the apparatus when there is one toy that has not yet been placed into a container. The key manipulation lies in the status of the goal by the time E leaves. In the experimental condition, E abandons his goal before filling the container to the specified level (i.e., the goal is incomplete). In the control condition, E fills the container to the specified level before leaving the apparatus (i.e., the goal is complete). In *both* conditions there is space for the final toy to be placed in the container, but only in the experimental condition would placing the remaining toy in E's goal container constitute completing E's goal.

We measured whether children placed the remaining toy in E's goal container, or in one of the two alternative containers. The goal slippage hypothesis predicts that children will be more likely to place the final toy

into E's goal container in the experimental condition than in the control condition. This is because placing the final toy in E's goal container would constitute completing E's abandoned goal in the experimental condition, whereas in the control condition E's goal has been completed before he leaves the apparatus. Therefore the goal slippage mechanism should only activate in the experimental condition.

3.2 Method

The hypotheses, sample size, methods, exclusion criteria and planned analyses were pre-registered before data collection and can be accessed at: http://aspredicted.org/blind.php?x=qz8dy6. All aspects of the study were carried out in accordance with the pre-registered protocol unless otherwise stated. This experiment was conducted in accordance with the Declaration of Helsinki, and was approved by the Humanities & Social Sciences Research Ethics Sub-committee (HSSREC) at the University of Warwick.

3.2.1 Participants

Thirty-four participants were included in the final sample (16 females, average age: 27;17, range: 24;12-29;05). We pre-registered a target sample size of 40 toddlers between the ages of 24 and 30 months in anticipation of a small-to-medium effect size on the basis of a pilot study. However, because our labs were closed indefinitely due to Covid-19, we declared data collection complete with our current sample of participants. This decision was taken prior to data analysis. An additional 10 participants were tested and excluded from final analysis because they did not complete at least two trials in each condition due to shyness (one) or because they took too long to place the toy into a container (nine) (see section 3.2.5). All participants were recruited in the Department of Psychology at the University of Warwick. The majority of participants came from middle class backgrounds and were Caucasian.

3.2.2 Design

We used a within-subjects design, with eight test trials in total. Condition was our independent variable, with two levels (experimental

condition and control condition). Our main dependent variable was whether or not participants placed the toy in the same container as E in each test trial. We also measured whether or not participants placed the toy in any container at all in each test trial.

To control for order effects of condition, the 8 trials were split into two blocks of four: Block 1 (experimental condition test trials), and Block 2 (control condition test trials; for the distinction between experimental and control condition test trials, see section 3.2.4). Block order was counterbalanced between participants. E's goal container alternated sequentially between trials (e.g., if E's goal container on the first test trial was on E's left-hand side, then the goal container on the subsequent trial was in the middle), and the goal container's location for the first test trial alternated sequentially between participants.

3.2.3 Materials and Apparatus

Participants sat approximately 1.5m away from the apparatus on their caregiver's lap. The apparatus consisted of eight games, each of which was used for one test trial (four per condition). Each game consisted of a type of toy (wooden balls, or wooden blocks affixed with pictures of cars, trains, or trees) and a corresponding set of three containers ('homes'). Each container could fit three of the corresponding type of toy inside it. In addition to the toys and containers, there was a small pad on which three toys were placed at the beginning of each test trial. For each trial the containers were set equidistant from the pad (see *Figure 3.1*).

There was also a screen behind which E disappeared when disengaging from the apparatus, such that E could no longer be seen by participants. This served two purposes. Firstly, it emphasised that E was no longer interested in playing further with the apparatus. Secondly, it made it impossible for children to socially interact with E while engaging with the apparatus themselves.

Each container had a red line, indicating the point at which E's goal to fill the container would be reached. In the experimental condition this line would be reached with three toys (see *Figure 3.1*). In the control condition,

however, this line would be reached with only two toys. To visually emphasise this difference, the space for the third toy in the control condition was made of plastic rather than cardboard (see *Figure 3.2*).

Figure 3.1

Sample Game in the Experimental Condition

Note: This figure shows a sample game (in which the toys are wooden balls) in the experimental condition. In test trials, E knelt behind the apparatus and faced the participant. The containers could be filled to the red line by placing three toys into the container.

Figure 3.2

Sample Game in the Control Condition

Note: This figure showed a sample game (in which the toys are wooden balls) in the control condition. In test trials, E knelt behind the apparatus and faces the participant. The containers could be filled to the red line by placing two

toys into the container, and a plastic extension made it possible to place a third toy on top.

3.2.4 Procedure

Participants were tested individually in laboratories at the University. Caregivers gave informed written consent, and participants received a gift for taking part. Sessions lasted approximately 20 minutes.

Caregivers were present and played a largely passive role in test trials, with three exceptions. Caregivers were instructed to encourage the participant to place the toy in a container (without giving any indication as to which container to place it in) if the participant did not approach the apparatus 5 seconds after E retreated behind the barrier. In this instance, caregivers were instructed to say: "You can put it in the home you want." In addition, caregivers were instructed to either praise the participant's behaviour after every test trial ("good job") or after none of the test trials, to ensure equal praise regardless of behaviour. Finally, caregivers were told to smile encouragingly at participants, regardless of the participants' behaviour.

During warm-up participants were acquainted with the toys and the pad, as well as retrieving toys from the pad and filling a box with toys. After warm-up, caregivers were asked to sit on a chair with the participant on their lap while E set up for the first familiarisation trial.

Familiarisation Phase

There were three aims of familiarisation. Firstly, we wanted to familiarise participants with filling each type of container with the corresponding type of toy. Secondly, we wanted to familiarise participants with the containers used in both the experimental and control condition test trials. Thirdly, we wanted to encourage children to empty the central pad, to reduce the number of test trials in which participants did not place the toy anywhere at all.

There were eight familiarisation trials, one for each game. The familiarisation trials proceeded with a fixed order of toys (balls, cars, trees, trains). The games alternated between the experimental and control

condition containers for each type of toy, and the order of conditions was counterbalanced between participants.

Participants sat approximately 1.5m away from the apparatus on their caregiver's lap. The experimenter knelt on the far side of the apparatus, facing the participant (as in *Figure 3.1* and *Figure 3.2*). At the start of each familiarisation trial there were no toys inside any of the three containers, and three toys on the central pad.

Each familiarisation trial proceeded by turn-taking – first E filled a container to the red line with toys, and then E returned the toys to the pad and encouraged the participant to fill a container (of the child's choice) with toys. Participants were encouraged to fill any container so that they learned that any of the three possible containers could be filled. Particular emphasis was drawn to filling containers to the red line, to ensure that participants understood that this was E's goal. E praised the participant when they had filled a container ('Well done!').

The control condition containers could be filled to the red line with only two toys, which meant that there was always one remaining toy after E's goal was accomplished. During their turn, E encouraged the participant to place the final toy into the plastic part of the container, for the sake of tidying up the final toy. This was done to ensure that participants understood that it was possible to place a third toy in each container in the control condition. E's goal container alternated sequentially between each trial, such that participants did not learn to associate E's goal with a particular location (e.g., right, left, or middle).

Test Phase

There were eight test trials in total (four per condition). The physical set-up for test trials was the same as for familiarisation trials (see *Figure 3.1*). Test trials proceeded with the same fixed order of toys as the familiarisation trials (balls, cars, trees, and trains), and condition was blocked and counterbalanced between participants (see section 3.2.2). E's goal container alternated sequentially between each trial.

At the start of each trial there were three toys on the pad, and two toys in two of the containers. The third container, E's goal container, was empty at the start of each trial. The number of times that E referred to each container, and the time that E spent looking at each container, was kept constant across conditions.

Each trial consisted of three phases: (i) establishing E's goal, (ii) the experimental manipulation, and (iii) E leaving the apparatus.

(i). At the beginning of each test trial, in both conditions, E established his goal with verbal reference and gesture: "[Participant's name], I want to fill this home [E points to the container] to the end [E runs his finger along the container to the red line] with these [toys] [E points to the toys]."

(ii). E then began to place the toys into the goal container: "One, two." In the experimental condition, E then abandoned his goal: "I won't do it" [E picks up the third toy and moves it towards the goal container, shakes his head, and then returns it to the pad]. In the control condition, E noted that his goal has been completed: "Now I've done it" [E runs his finger up to the red line on the goal container, picks up the third toy and moves it towards the goal container, before returning it to the pad].

(iii). The control condition ended with E saying "I do not [E shakes his head] want that [toy] [E points to the remaining toy] in that home [E points to the goal container]." In the experimental condition E said the same thing, except that he added "anymore" to emphasise that the goal was abandoned.

In both conditions E then walked behind the barrier (located behind E's kneeling position) out of the participant's view, after which the caregiver set the participant down in front of them. If the participant did nothing after 5 seconds, E prompted the caregiver to say "You can put it in the home you want" by raising his hand over the barrier.

Once the participant had placed the toy into one of the containers, or after 10 seconds passed after the caregiver prompt, E returned from behind the barrier with the next set of containers. E did not provide any feedback to

the participant, or acknowledge where the participant had placed the toy. E replaced the apparatus with the apparatus for the next trial, and encouraged the participant to return to their caregiver's lap. Once they had done so and settled down, the next trial commenced.

3.2.5 Coding and Drop-out Criteria

For each trial, we coded whether participants placed the toy in any one of the three containers, and if so, whether they placed the toy in the same container as E or in one of the other containers. An action qualified as a placement if the participant put the toy in that container and then let go of the toy. If the participant placed the toy in one container and then moved it to another container, we counted the first placement.

All sessions were recorded using digital video recorders. Coding was carried out by a research assistant who was naïve to the hypotheses of the study. The coder assessed: placement behaviour (yes or no), placement in the same container as E (yes or no), and participant dropout and trial exclusion (see below). A second research assistant who was naïve to the hypotheses of the study coded a random 10 participants (29.4%) for reliability. Using the *Kappa.test* function of the R package *fmsb* (Nakazawa, 2019), coders were found to be in substantial agreement in judging whether the participant placed the toy anywhere at all (judgements matched 93% of the time, $\kappa = 0.78$ (95% CI: 0.60, 0.95), $p < .001$), whether participants placed the toy in the same container as E (judgements matched 90% of the time, $\kappa = 0.80$ (95% CI: 0.66, 0.93), $p < .001$), and whether to drop individual trials (judgements matched 90% of the time, $\kappa = 0.73$ (95% CI: 0.56, 0.91), $p < .001$).

Participant Drop-out Criteria

Participants were dropped from analysis if they did not complete at least two test trials per condition.

Trial Exclusion Criteria

A trial was excluded if any of the following criteria were met: (1) If there was a delay of longer than 10 seconds between the caregiver prompt and the participant placing the toy into one of the containers, (2) if the

experimenter made an error (e.g., used the wrong location as his goal, or returned from behind the barrier too early), (3) if a caregiver instructed the participant as to which container to place the toy, (4) if the participant was not watching E when E established his goal or during the experimental manipulation (i.e., the dialogue and actions specified in (ii)). If the participant was not watching E when E established his goal or during the experimental manipulation, E called the participant by their name and tried again. If the participant still did not pay attention, E called the participant's name once more. If the participant still did not pay attention, the trial was excluded, and E moved on to the next trial.

3.3 Results

3.3.1 Data Screening

From the responses of the 34 participants included in our analysis, 44 test trials were excluded (18 in the experimental condition) due to caregiver interference (five), the participant not watching when E established his goal (one), and a delay of longer than 10 seconds between the caregiver's prompt and participants placing the toy into one of the containers (38).

This left 228 test trials for further analysis (118 in the experimental condition). Participants had to be prompted by the caregiver in 22 of these trials (12 in the experimental condition). All subsequent analysis was conducted in *R* (R Development Core Team, 2018; Revelle, 2018).

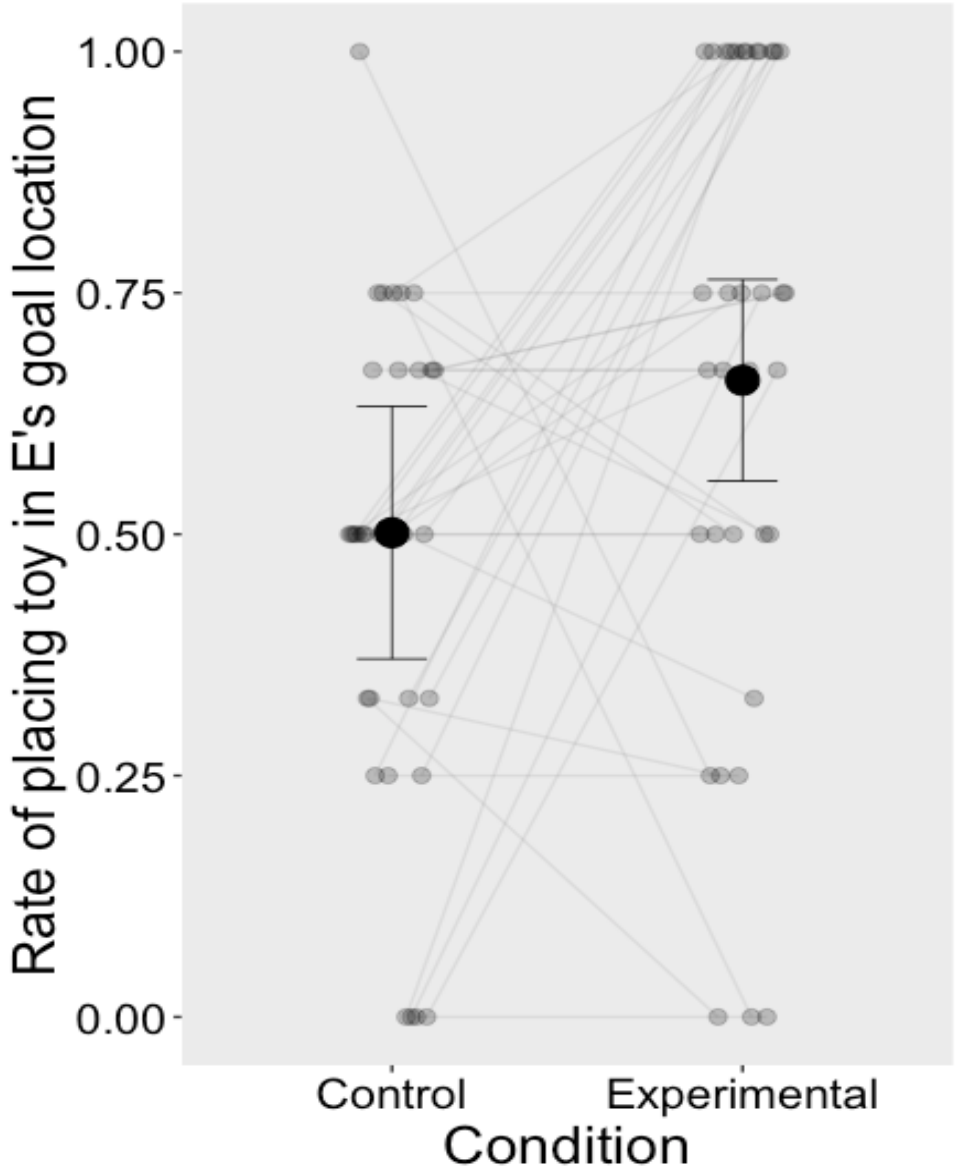
See the supplementary materials (https://osf.io/unrzy/?view_only=8c96e180b574409fb1ad43d446e0898a) for this experiment's data and *R* code.

3.3.2 Placement of Toy in the Same Container as E

To investigate whether condition influenced placement location, we measured whether participants placed the toy in the same container as E or in one of the other containers. Participants placed the toy in the same container as E in 67% of experimental condition test trials, and in 49% of control condition test trials (see *Figure 3.3*).

Figure 3.3

The Proportion of Test Trials in Which Participants Placed the Toy into E's Goal Container



Note: This figure shows the proportion of trials in which participants placed the toy into E's goal container in each condition, with 95% confidence intervals of the means adjusted for within-subject design (Cousineau, 2005; Loftus & Masson, 1994; Morey, 2008). Jittered dots represent individual performances in each block, with light grey lines connecting each participant's performance across conditions.

To test whether condition determined placement location, we used the function *mixed* of the *R* package *afex* (Singmann et al., 2020) to create a generalized linear mixed model (GLMM) with a binomial error structure. Our dependent variable was whether participants placed the toy in the same container as E or not (binary). Condition (experimental vs control) was the only fixed effect. Our random effects included the random intercept of participant and trial number. We initially included the random slope of participant and trial number, but we removed these terms because of singularity in the model (Barr et al., 2013; Bates et al., 2015; Singmann & Kellen, 2019).

The full model was compared to a model that was identical except that the fixed effect was removed. The results indicated that the full model was a better fit for the data as compared to the null model, $\chi 2(1) = 8.95$, $p = .003$. The odds of participants placing the toy in the same container as E were over two times larger (*OR* = 2.37, *estimate* ± *SE* = 0.86 ± 0.29, $p = .003$) in experimental condition test trials than in control condition test trials (see *Table 3.1* for a summary of results), indicating a small-to-medium effect size (Chen et al., 2010; Chinn et al., 2000; Maher et al., 2013).

Table 3.1

Summary of the Fixed Effects of the Mixed Logistic Model for Same Location-Placement

Predictor	Coefficient	Standard Error	Wald Z	*p*
Intercept	-0.757	0.299	-0.253	.800
Condition	0.862	0.294	2.934	.003**

Note: log-likelihood = - 147.4

** < .01

3.3.3 Comparing Placement-Location to Chance

The main analysis indicated that participants differentiated between the two conditions. However, this analysis does not indicate what is driving the effect - participants may be operating at chance (i.e., randomly placing toys in any of the three containers) in one of our conditions. To better understand the effect found in the main analysis, we therefore compared location-placement to chance in both conditions. Participants had a 33% chance to place the toy in the same container as E if they were randomly placing toys into any one of the three containers on each test trial.

We calculated each participant's rate of placing the toy in the same container as E for each condition. We then conducted two one-sample Wilcoxon signed-rank tests (one per condition) using the functions *wilcox.test* and *wilcoxononeSampleR* of the *R* package *rcompanion* (Mangiafico, 2018). The median rate of placing the toy in the same container as E was 75% in the experimental condition and 50% in the control condition. Toy placement differed significantly from chance in both the experimental condition ($V = 553$, $p < .001$, $r = 0.75$) and the control condition ($V = 491$, $p < .001$, $r = 0.57$).

3.3.4 Placement of Toy in Any Container

One potential criticism of the current study's design is that the containers looked different between the two conditions. Specifically, all containers in the experimental condition were entirely made from cardboard, whilst the space for the third toy in each control condition-container was made of plastic. This raises the possibility that our results were driven by differences in the physical affordances of the containers - in particular, that participants may have found it more enticing to place the toy in any container in the experimental condition than in the control condition.

To investigate this possibility, we measured whether participants placed the toy in any container or not in each test trial. The subsequent analysis was not pre-registered, and it comes before trial exclusion criteria were applied to the data (because we excluded trials in which participants did not place the toy anywhere at all for the main analysis). Prior to the

application of exclusion criteria, there were 136 test trials per condition. Children placed the toy in any container in 92% of the experimental condition test trials, and in 88% of the control condition test trials.

To test whether condition had an effect on placement, we ran the same test as for the main analysis, except that our dependent variable was whether participants placed the toy anywhere or not (binary). The results indicated that the full model was not a significantly better fit for the data as compared to the null model, $\chi 2(1) = 1.63$, $p = .201$, i.e., we found no evidence that condition significantly predicted whether participants placed a toy anywhere at all or not.

3.4 Discussion

We tested whether 2-year-olds would be more motivated to place a toy into a container when this constituted completing a goal that was previously abandoned by an experimenter (experimental condition) than when the experimenter had already completed his goal of filling the container to a specified level (control condition). The results revealed that participants' placement-behaviour differed between conditions, and specifically, that participants were significantly more likely to place the toy in E's goal container in the experimental condition (thereby completing E's goal) as compared to the control condition. In addition, comparisons to chance revealed that participants placed the toy in E's goal container above chance in both conditions. These results support the goal slippage hypothesis, that an observer may take on the goal of another agent for themselves when they see that the goal will not be achieved otherwise, such that the observer now has that goal and will pursue it independently of what the other agent is now doing. More specifically, these results provide some support for the goal slippage hypothesis.

The finding that participants were placing the toy in the same container as E above chance in *both* conditions is worth dwelling on. It is an important point in support of the goal slippage hypothesis that children were placing the toy in E's goal container above chance in the experimental condition, because otherwise our results could be explained by children

distributing the toy randomly across the three containers. It may seem surprising that children were placing the toy into E's goal container above chance in the control condition as well, however. Our explanation for this is that E's goal container is likely more salient to participants than the alternative containers simply due to E drawing the participants' attention towards that container by acting upon it in both conditions. Therefore even though E's goal was completed in the control condition, we would still expect children to prefer E's goal container over the alternative containers in this condition because the salience of this container (relative to the alternatives) has been increased. Under our interpretation, then, children are placing the toy into E's goal container above chance in both conditions because of the increased salience of this container relative to the alternatives, and children are placing the toy into the goal container more often in the experimental condition than the control condition due to goal slippage.

Our experimental design further enables us to rule out several alternative explanations for these results. Firstly, these results cannot be driven by psychological altruism or a desire to socially interact with the experimenter. Psychological altruism cannot account for the difference between our conditions because E abandons his goal in the experimental condition (and therefore psychological altruism would not predict that children will complete E's goal). In addition, because E disengaged from the apparatus and walked behind a barrier, participants' behaviour cannot be explained by children seeking to socially interact with E.

Second, these results cannot be explained by differences in the physical affordances of the containers between conditions. If it was more enticing for participants to place toys in any container in the experimental condition than in the control condition due to differences in the stimuli between conditions, then we would expect placement in any container to be higher in the experimental condition than in the control condition. However, we did not find evidence to support this hypothesis.

Finally, due to our experimental design, the results cannot be explained by children imitating E's actions. This is because in both conditions E placed two toys into the goal container, and therefore copying E's actions would not have explained the difference in results between our conditions.

One concern about the design of the present study is the stimuli differences between conditions. These differences are the placement of the red line in a different position between conditions and the difference in the material of the containers between the conditions. First consider the placement of the red line. In the experimental condition the red line was after the space for the third toy, whereas in the control condition the red line was after the space for the second toy. In both conditions E's goal was to fill the container to the level of the red line. This difference in E's goal between the conditions is important because it needed to be possible for participants to perform the same action in both conditions (i.e., placing the remaining toy into E's goal container), where the implication of this action for E's initial goal differed between conditions: in the experimental condition this action constitutes completing E's initial goal and in the control condition it is superfluous to E's goal. This experimental manipulation of E's goal status is at the core of testing for goal slippage, and so it must be possible for participants to perform the same subsequent action in order to identify whether they are more likely to do this when it constitutes completing E's goal.

Our motivation for introducing the red line and altering its height between conditions was to ensure that the difference in E's goal between conditions was sufficiently clear to participants. In particular, we wanted to ensure that participants could visually identify that E's goal was completed in the control condition with only two toys being placed into E's goal container, whereas in the experimental condition placing only two toys into E's goal container would not complete his goal. A further benefit of using the red line is that it is a permanent fixture on the apparatus – even when E retreats behind the barrier, E's goal is still clearly indicated. While this is

not *necessary* to test for goal slippage, we felt more confident in children identifying E's goal with the presence of a visual indicator of the goal. If we had found no difference between conditions in the absence of the red line, we could not be certain as to whether this was because goal slippage does not motivate children at this age or because participants did not identify the difference in E's goal between conditions.

It would be possible for a subsequent study to alter the present study's design to remove this discrepancy, although it would require some other means of establishing the difference in E's goal to participants. For instance, one could remove the red lines from the containers, and simply rely on E's verbalisation of his goal. In the experimental condition E could state "I want to place *all three* of these toys into this home", and in the control condition E could say "I want to place *only two* of these toys into this home". Alternatively, instead of manipulating E's verbalisation, one could rely on E's pointing to establish the different levels to which he wants to fill containers in each condition. That is, E could say in both conditions "I want to fill this home to here", and in the experimental condition E could point towards the level that would require three toys and in the control condition E could point to the level that requires only two toys. Implementing these changes would allow for the removal of the red line altogether, thus removing this difference between conditions.

The other concern relating to our stimuli is that the space for the final toy was made of different material between the two conditions: in the experimental condition the space for the final toy was made of cardboard, whereas in the control condition the space for the final toy was made of plastic. The motivation for this change was to visually emphasise the distinction between conditions. One potential concern with this difference is that one might think that the difference in the stimuli materials alters the salience of the goal container between conditions, and specifically, that this difference makes the goal container more salient in the experimental condition than in the control condition. While we found no evidence that *all* of the containers in the experimental condition were more salient than the

containers in the control condition (see section 3.3.4), one might be concerned that the difference in stimuli material drew participants' attention to the empty space in the goal container, specifically, more in the experimental condition than in the control condition. The problem with this is that it offers an alternative explanation for our findings; this could explain why children are placing the toy into E's goal container more often in the experimental condition than in the control condition.

It is not clear to us why this difference in material would alter the salience of *only* the goal container. In addition, it is not clear to us why this difference would specifically raise the salience of the empty space in the experimental condition rather than the control condition; one might equally expect that the plastic part of the container in the control condition would draw children's attention to the empty space more than the experimental condition because the plastic makes it easier for children to see the previous toy that E had placed into the goal container. In this case, the difference in stimuli would instead increase the frequency of children placing the toy into the goal container in the control condition. Nevertheless, the potential implication of the difference in stimuli material remains a point of concern; if the salience of the goal container were higher in the experimental condition then this would be able to account for our results independently of goal slippage.

Fortunately, it would be straightforward for a subsequent study to address this concern about the difference in the materials used for the containers because this difference is separate for E's goal. That is, while we implemented this change in order to visually emphasise the difference in E's goal between conditions, it is not necessary for establishing E's goal. This means that one could replicate the present experiment without this stimuli difference; if one failed to find a difference between conditions, then it would suggest that the present study's results were driven by this difference in salience between conditions rather than goal slippage.

One might wonder why the goal slippage hypothesis does not predict that the experimenter's abandoning of the goal would also slip to the

participant. There are two main reasons for rejecting this prediction. Firstly, the goal slippage hypothesis is a hypothesis specifically about goals, and motivated by evidence pertaining to goals. Goals, as a particular kind of psychological entity, play a crucial role in children's understanding of the world and others' behaviours (Luo & Baillargeon, 2005; Sommerville et al., 2005; Spaepen & Spelke, 2007; Woodward, 2009). The hypothesis is not that *all* psychological activity slips between agents, and such a hypothesis would require further evidence to motivate it. Secondly, just as with other forms of contagion (e.g., motor contagion, or emotion contagion), a lack of the relevant entity (e.g., emotion or movement) is not predicted to slip to the observer (Bouquet et al., 2011; Ruffman et al., 2019). That is, just because one expects a goal to slip, it does not follow that one should expect a lack of a goal to slip, any more than motor contagion hypotheses predict that a lack of motion might slip to an observer.

The present study extends past research that has investigated the proximate mechanisms that underpin instrumental helping in early childhood (Carpendale et al., 2015; Cirelli et al., 2016; Hammond & Brownell, 2018; Hepach et al., 2012; Leimgruber, 2018; Over & Carpenter, 2009; Paulus, 2014; Paulus & Moore, 2012). There are several mechanisms that have been proposed to underpin instrumental helping in early childhood (e.g., psychological altruism), and goal slippage is not mutually exclusive with these mechanisms. However, as this study is the first to directly test and support the goal slippage hypothesis, it carries significant implications for our understanding of instrumental helping.

Firstly, finding evidence in support of the goal slippage hypothesis prompts a re-interpretation of the findings from some past studies. While the goal slippage hypothesis may not be able to explain the results of *all* past instrumental helping studies, it is consistent with the results of many helping studies (Hepach et al., 2017; Hepach et al., 2020; Warneken, 2013; Warneken et al., 2007; Warneken & Tomasello, 2006, 2013) because goal slippage also predicts that children will be motivated to help agents whose goals have been interrupted. The results of these past studies have often

been interpreted as supporting the psychological altruism hypothesis. However, previous interpretations of these results must now be re-evaluated in light of the current study's findings, as the goal slippage hypothesis is an alternative (and, now, empirically supported) explanation.

In addition to offering an alternative means of reinterpreting past findings, the goal slippage hypothesis may also be better placed to explain the results from certain instrumental helping studies. For instance, the goal slippage hypothesis is well placed to explain why children would be motivated to instrumentally help an animate geometric shape achieve its goal (Kenward & Gredebäck, 2013), as all that is required for goal slippage is that children can identify another agent's goal. Therefore the goal slippage hypothesis not only offers an alternative explanation for some past findings, but it may even offer a preferable explanation of some past findings. This raises the question as to the degree to which past instrumental helping results have been actually underpinned by the goal slippage mechanism.

The present study also raises the possibility that past theoretical positions about the development of prosocial behaviours should be reassessed. It has previously been suggested that prosociality emerges from other-concerning cognitive mechanisms, such as psychological altruism (Hepach et al., 2017; Warneken, 2015). The present findings raise the possibility that the starting point for prosociality may, in part, be more self-centred in nature, insofar as one of the earliest forms of prosocial behaviour may be underpinned by a self-regarding psychological mechanism (goal slippage). Prosociality may then develop from such self-centred origins through the reinforcement of prosocial behaviours by caregivers (Dahl, 2018).

The present research also carries implications for the evolution of prosocial behaviours. A proximate psychological mechanism that motivates agents to get involved with others' goals would be advantageous for social learning (Michael & Székely, 2019), insofar as contributing to others' goals may help children to better understand what goals are useful to bring about

and how to bring about those goals. A cognitive mechanism that motivates agents to get involved in others' goals would also promote cooperation between agents. Therefore the present study extends our understanding of the developmental trajectory and evolutionary emergence of prosociality.

3.4.1 Future Research

There are several avenues for future research following the present study's findings. Firstly, future research could deepen our understanding of goal slippage by investigating the factors that might promote goal slippage. For example, goals may be more likely to slip the closer they come to completion, as the cognitive tension that arises from seeing a goal remain incomplete may increase the closer the goal is to completion. This hypothesis is motivated by the finding that adults are more willing to invest in the completion of their own goals the closer they come to achieving them (Hull, 1932; Kivetz et al., 2006). To test this hypothesis, one could adapt the present study such that the experimenter abandons a task at different stages (i.e., closer or further away from completion). This hypothesis generates the prediction that children would be more likely to complete the experimenter's goal if it was abandoned closer to completion, as compared to if it was abandoned at an earlier stage.

A further hypothesis that could be tested in future research is that goal slippage may be modulated by how valuable the goal is, as indicated by the amount that an agent is willing to invest in the pursuit of the goal. Specifically, a goal may be more likely to slip if the agent originally pursuing the goal is willing to invest more in its completion. This hypothesis is motivated by findings that young children can infer the value of others' goals on the basis of how much they are willing to invest in achieving their goals (Jara-Ettinger et al., 2015; Liu et al., 2017), as well as evidence that commitment to a joint task can be modulated by the degree to which a partner is willing to invest in the task (Chennells & Michael, 2018; Székely & Michael, 2018). To test this hypothesis, the present study could be adapted such that the experimenter invests varying amounts of effort into bringing about a goal before abandoning it. Conducting follow-up studies

that test these hypotheses about how goal slippage works would provide a better understanding of the goal slippage mechanism, thereby improving our understanding of this developmentally primitive mechanism for promoting prosocial behaviour.

A second avenue for future research would be to investigate the goal slippage mechanism beyond the population that was the focus of this study (i.e., 2-year-old humans). Specifically, it would be fruitful to investigate the goal slippage mechanism in human adults, and in non-human animals.

One could test the goal slippage hypothesis in adults by adapting the present study for an adult population. For example, placing toys into different containers could be replaced with filing documents into different trays. This line of inquiry is motivated by evidence that simply framing a goal as having been started by another agent may encourage agents to complete the goal quicker than if the goal is framed as one that is yet to be started (Kivetz et al., 2006). Investigating whether the goal slippage mechanism is present in an adult population would inform our understanding of the cognitive components of prosocial behaviour in adulthood (for a more detailed suggestion of how goal slippage could be tested in adults, see section 6.3).

Testing the goal slippage hypothesis in non-human animals may be more challenging, as it may first require proof that animals understand goal abandonment. However, such an investigation would be fruitful because it may indicate whether goal slippage is a mechanism promoting prosocial behaviour that is unique to humans (see section 6.3). If goal slippage were only found in humans, this may indicate that goal slippage is a key socio-cognitive difference between humans and non-human animals. There is therefore much scope for fruitful research investigating the goal slippage mechanism with different populations.

3.4.2 Conclusion

This is the first study to test the goal slippage hypothesis, under which an observer may adopt another agent's goal for themselves (and thereby be motivated to contribute to that goal's achievement) if the agent

may struggle to achieve the goal by themselves. The results support the hypothesis that the goal slippage mechanism may account for instrumental helping in early childhood.

This chapter contributes to this overall thesis by investigating whether the goal slippage mechanism is active in early childhood. These findings carry implications for research into the developmental origins of prosociality, specifically raising the possibility that the emergence of prosocial behaviours may be the result of a self-concerning cognitive mechanism.

Chapter 4

Goal Status and Goal Salience: The Goal Slippage Hypothesis and the Predictive Value Hypothesis

Previous research indicates that goal status predicts goal saliency. Incomplete goals are more salient than completed goals (Asiala et al., 2020; Suh & Trabasso, 1993; Trabasso & Suh, 1993), because of their greater value in predicting others' future behaviour (predictive value hypothesis). Past studies have not differentiated between types of incomplete goals - interrupted and abandoned – which differ in their predictive value. Additionally, an alternative hypothesis, the goal slippage hypothesis, can also explain past findings, and generates different predictions if abandoned goals are distinguished from interrupted goals. We tested these two competing hypotheses across three experiments with MTurk workers. Participants were asked to read and summarise texts, and the status of a target goal (abandoned, interrupted, or completed) was manipulated between subjects. Goal salience was measured by whether the target goal was mentioned in free recall. In Experiments 1 and 3 goal status was not found to predict goal salience, and Experiment 2 indicated that completed goals were more salient than abandoned or interrupted goals. These results are not predicted by either hypothesis and are inconsistent with past findings. Our findings may be due to alterations that we made to the test materials used in past studies.

4.1 Introduction

Goal tracking is a key component of human social cognition. We use goal tracking in a variety of everyday contexts, from small-scale dyadic interactions (such as helping friends move furniture) to larger-scale social activities like organising meetings and conferences. Goal tracking helps us to identify others' preferences (Choi et al., 2018; Luo et al., 2009, 2017; Luo & Baillargeon, 2005), beliefs (Southgate et al., 2007; Southgate & Vernetti, 2014), and future actions (Eisenberg et al., 2018; Flanagan & Johansson, 2003; Liepelt et al., 2008; Sacheli et al., 2013). In this way, goal tracking fosters cooperation and coordination (Bratman, 2014; Garrod & Pickering, 2009; Henrich, 2004; Knoblich et al., 2011; Melis & Semmann, 2010; Pettit

& Schweikard, 2006; Pezzulo et al., 2018; Tomasello, 2014), as well as commitment (Chennells & Michael, 2018; Gilbert, 2006; Green et al., 2019; Michael, Knoblich, & Sebanz, 2016; Michael, Sebanz, & Knoblich, 2016). Given that goal tracking underpins many aspects of social cognition, it is pivotal in facilitating much of our everyday social activity.

While the benefits of goal tracking are clear, there is a lingering question as to what the main factors are that prompt agents to track the goals that they do track. Why is it that we track certain goals instead of others, or more closely than others? It is important to understand what determines the saliency of different goals given the crucial role that goal tracking plays in everyday life. Identifying the factors that promote goal tracking will enable us to better encourage and facilitate goal tracking in those contexts where it is most important, e.g., in cooperative projects and circumstances in which tightly coordinated behaviour is essential.

More specifically, we might ask what *goal-specific* factors promote goal saliency? We add this qualification because although many factors may promote goal saliency, they are not all equally interesting from a socio-cognitive perspective. Goal salience might be promoted by something as innocuous as how bright an agent's clothes are, or how loud they are talking. However, these factors are not unique to goals - they are also factors that would promote the salience of non-goal events. We are therefore specifically interested in goal-specific factors that promote goal saliency.

One goal-specific factor that has been found to promote goal salience is goal status. More specifically, incomplete goals are more salient than completed goals. For example, suppose that I am going to make myself a cup of coffee. This goal will be more salient to an observer if my goal remains incomplete (e.g., if I am interrupted by a phone call before I can finish making the coffee) than if the goal is completed (e.g., if I receive a phone call after making the coffee). This finding has been replicated in a variety of contexts (Asiala et al., 2020; Burke, 2011; Lewis, 1944; Lewis & Franklin, 1944; Lutz & Radvansky, 1997; Magliano & Radvansky, 2001; Moss et al., 2007; Ovsiankina, 1928; Radvansky & Curiel, 1998; Suh &

Trabasso, 1993; Trabasso & Suh, 1993; Zeigarnik, 1927), and using different research methods. For example, participants are more likely to mention a target goal that was incomplete (as compared to complete) when they are asked to summarise stories verbally and in writing (Asiala et al., 2020; Suh & Trabasso, 1993; Trabasso & Suh, 1993), and participants are faster and more accurate in determining whether a target word was associated with an agent's goal when that goal was incomplete (Lutz & Radvansky, 1997; Magliano & Radvansky, 2001; Radvansky & Curiel, 1998). The cumulative finding from this research is that goal status is an important factor in determining goal salience.

One proposed explanation for this relation between goal status and goal salience is what we will refer to as the *predictive value hypothesis* (Asiala et al., 2020; Bower & Rinck, 1999; Lutz & Radvansky, 1997; McNamara & Magliano, 2009; Trabasso & Suh, 1993; Zwaan & Radvansky, 1998). Under this hypothesis, an incomplete goal is more salient than a completed goal because incomplete goals are more useful in predicting the future actions and behaviours of agents. Reconsider the example of me making myself a cup of coffee. If I fail to achieve my goal because I was interrupted by a phone call then it is likely that I will have retained this goal, and this goal will therefore guide my future behaviour (i.e., I will return to the kitchen to finish making myself my cup of coffee). On the other hand, if I managed to make coffee before the phone rang, then this goal will no longer guide my future actions - I will not return to the sink and boil the kettle again if I already have my coffee. Under this hypothesis, goal status determines goal salience because goal status determines a goal's predictive value.

The predictive value hypothesis has been used to explain a wide range of cognitive phenomena. If agents preferentially track incomplete goals over completed goals (and therefore dedicate more cognitive resources to processing information about incomplete goals), this would explain why an incomplete goal can distract agents from an unrelated task, and thereby lower their performance on the unrelated task (Rothermund, 2003; Shah &

Kruglanski, 2002). The predictive value hypothesis can also explain why agents pursuing incomplete goals receive more attention from observers compared to agents who have completed their goals (Linderholm et al., 2004; Loucks et al., 2017; Zwaan, 2016), and why agents typically model their understanding of an environment and others' behaviours in reference to goals that are incomplete (Kurby & Zacks, 2012; Magliano et al., 2014; Zacks, 2020; Zacks & Tversky, 2001; Zwaan, 2016). In the context of memory, the predictive value hypothesis can also explain why information pertaining to incomplete goals is more readily recalled compared to information about completed goals (Asiala et al., 2020). The finding that incomplete goals are more salient than completed goals, and the predictive value hypothesis that has been put forward to explain this, therefore carry wide-reaching implications for research in social cognition and beyond.

There is a notable flaw in the design of the past studies which have shown that incomplete goals are more salient than completed goals, however. These studies failed to distinguish between different kinds of incomplete goal. Goals may be incomplete because the agent's pursuit of the goal has been interrupted (i.e., the agent retains the goal, but their originally planned means of achieving the goal has been blocked), or because the agent has abandoned the goal (i.e., the agent has changed their mind and no longer retains the goal). Past studies only distinguished between two levels of goal status – complete and incomplete goals.

The distinction between these two types of incomplete goal (abandoned and interrupted) matters for the predictive value hypothesis because the predictive value of an incomplete goal depends on *why* the goal is incomplete. If my goal to make a cup of coffee is interrupted then this goal is still useful in predicting my future behaviour (because I am likely to have retained the goal). However, my goal may alternatively be abandoned (e.g., if I realise that it is too late in the day for me to have a coffee), in which case my goal will no longer guide my future behaviour. Abandoned goals therefore carry the same predictive value as completed goals. In both cases the goal is incomplete, but the goal only guides my future behaviour if

it is interrupted. Therefore under the predictive value hypothesis, the relation between goal status and goal salience is not simply that incomplete goals are more salient than complete goals. Instead, this hypothesis predicts that interrupted goals are more salient than completed or abandoned goals, with no predicted difference in salience between completed and abandoned goals.

An alternative hypothesis about the relation between goal status and goal salience, which generates different predictions about the salience of incomplete and abandoned goals than the predictive value hypothesis, is the goal slippage hypothesis. Under this hypothesis, a cognitive tension is generated when an agent identifies that a goal which another agent was pursuing will not be completed. This cognitive tension is alleviated when the goal is completed, thereby encouraging the observer to contribute to the incomplete goal. In this way, the goal *slips* from actor to observer, and from perception to action (Hornstein, 1972; Lewin, 1951; Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2018, 2019). An example of this would be observing someone struggling with a task that you find straightforward, e.g., observing a relative who is struggling with technology that you are familiar with.

The goal slippage hypothesis generates different predictions than the predictive value hypothesis about the effects of goal status on goal salience, if the distinction is made between abandoned and interrupted goals. Under the goal slippage hypothesis, interrupted and abandoned goals should both be expected to generate cognitive tension in an observer, because in both cases the target goal is incomplete. We should therefore expect these types of goals to be equally salient. In addition, interrupted and abandoned goals should both be more salient than completed goals because completed goals are not expected to generate cognitive tension (Hornstein, 1972). Although no studies have directly addressed the link between goal slippage and the relation between goal status and goal salience, observing another agent's incomplete goal (compared to a completed goal) can boost one's own performance in achieving a similar goal (McCulloch et al., 2011). The goal

slippage hypothesis thus offers a new interpretation of a phenomenon that has been the focus of past research (that incomplete goals are more salient than completed goals), and it generates different predictions than the explanation that has been provided for past findings (the predictive value hypothesis).

4.1.1 Present Research

The aim of the present study was to investigate the effects of goal status on goal saliency. We were specifically testing two competing hypotheses about this relation – the predictive value hypothesis and the goal slippage hypothesis. We compared the saliency of completed, interrupted, and abandoned goals.

The methods and materials that we used were adapted from earlier studies (Asiala et al., 2020; Kurby & Zacks, 2012; Lutz & Radvansky, 1997; Radvansky & Curiel, 1998; Suh & Trabasso, 1993; Trabasso & Suh, 1993). We made two main improvements to these materials.

Firstly, unlike past studies, we clearly distinguished between interrupted and abandoned goals. In past studies, it was often not clear whether an incomplete goal was abandoned or interrupted. This ambiguity in the status of the target goal is problematic because it may affect the predictive value of the target goal. For example, consider the story of Betty, who went to the department store to purchase a birthday present for her mother (Trabasso & Suh, 1993). She discovers that the items on sale are too expensive, so she goes home and knits her mother a sweater. This can be viewed as the interruption of Betty's goal to find a present for her mother, but it could alternatively be viewed as Betty abandoning the goal to *buy* her mother a present. Although some previous studies involved the target agent's goals changing, these changes did not constitute goal interruption or abandonment. They instead involved the target agent taking up a new goal alongside a pre-existing goal or else once a previous goal has already been achieved (Asiala et al., 2020; Kurby & Zacks, 2012; Linderholm et al., 2004; Magliano et al., 2014; Speer et al., 2007).

Secondly, the stories used in previous studies were not consistent with regard to overall length of story (i.e., how many sentences participants were required to read), and multiple sentences were changed between conditions. This is potentially problematic because this raises the possibility of introducing confounding factors that could affect salience independently of goal status (e.g., the number of times that a target goal is mentioned in the story). We standardised these factors, manipulating only one sentence between conditions and keeping story length as consistent as possible.

The present experiments were conducted online using Amazon's Mechanical Turk (MTurk). This enabled us to replicate earlier findings with a broader sample than used in most earlier studies (Asiala et al., 2020; Lutz & Radvansky, 1997; Radvansky & Curiel, 1998; Suh & Trabasso, 1993; Trabasso & Suh, 1993), which typically relied on samples of undergraduate students.

In all experiments, we asked participants to read stories in which a target goal was either complete, incomplete, or abandoned by the protagonist. We then asked participants to summarise each story. The procedure and materials were identical for the first two experiments except for minor changes to the wording of the story. For Experiment 3, we asked participants to read three stories (instead of just one) to increase the number of observations per participant. Our measure of goal saliency was whether or not participants mentioned the target goal in free recall (a binary measure).

The predictive value hypothesis predicts that interrupted goals will be more salient than completed or abandoned goals, and that completed and abandoned goals should be equally salient. Under this hypothesis, we should therefore expect the target goal to be present in free recall more often in the interrupted goal-condition than in the other conditions. Under the goal slippage hypothesis completed goals should be less salient than either abandoned or interrupted goals, and interrupted and abandoned goals should be equally salient. Under this hypothesis, we should expect the target goal to be present more often in abandoned and interrupted goal-conditions (in

comparison to completed goal-condition). *Table 4.1* contains the predictions generated by each hypothesis about the relative saliency of each type of goal.

Table 4.1

Predictions Generated by Competing Hypotheses About the Relative Salience of Goals Depending on Their Status

Predictive value hypothesis	Goal slippage hypothesis
Abandoned < interrupted	Abandoned = interrupted
Abandoned = completed	Abandoned > completed
Interrupted > completed	Interrupted > completed

4.2 Experiment 1

Our goal in this experiment was to investigate whether the status of a target goal (i.e., whether it was interrupted, complete, or abandoned) altered the salience of that goal. Our measure of goal saliency was whether participants mention the target goal in free recall or not.

4.2.1 Method

The hypotheses, sample size, methods, exclusion criteria, and planned analyses were pre-registered on *AsPredicted.org* before data collection began (http://aspredicted.org/blind.php?x=si7wv2). This experiment was conducted in accordance with the Declaration of Helsinki, and was approved by the Humanities & Social Sciences Research Ethics Sub-committee (HSSREC) at the University of Warwick.

Participants

We opted for a large sample size of 100 participants per condition. This was because we expected high variability in our dependent measure because data from online experiments often have high variability and each

participant only contributed one data point. We also made this decision due to calculations with G*Power (Faul et al., 2009) and piloting that showed a small to medium effect size ($\alpha = 0.05$, power = 0.8).

We planned to collect data from at least 150 participants per condition, as we anticipated that approximately one third of participants would be excluded on the basis of pilot data. We requested data from at least 450 participants in total. Due to the randomisation procedure of SurveyMonkey, we could not ensure that there was a precise split of 150 per condition.

We collected data from 545 MTurk workers. Following pre-registered exclusion criteria, participants were excluded from analysis if they failed at least one of the two control questions pertaining to the contents of the story ($n = 68$), if they wrote more than 30 words ($n = 84$), or if they wrote nothing at all ($n = 65$). We decided to limit participants' summaries of the story to 30 words to ensure that they were indeed *summarising* the story, as requested, rather than simply writing out the entire story or copying and pasting the entire story into the answer box. Our concern was that if participants could write as much as they wanted, then we might expect participants to mention the target goal in *all* conditions, and our experimental manipulation would therefore be ineffective. Our final sample ($n = 328$) was made up of 154 females (172 males, one "prefer not to say", one "other"), with a mean age of 35.07 years (range = 19.00 : 69.00 years, $SD = 10.85$ years). All participants received monetary compensation of $0.75 for taking part.

Design

We implemented a between-subjects design with goal status as the independent variable with three levels: the target goal was either completed by the story's protagonist (completed condition), abandoned (abandoned condition), or interrupted (interrupted condition). We manipulated one sentence between conditions in the story. Our dependent variable was whether participants mentioned the target goal in free recall or not (binary).

We used SurveyMonkey's random assignment feature to randomly assign participants to each condition.

Materials

The following story was used as stimulus: "Betty had just woken up. It was a workday so she began her normal routine. She decided to make her morning cup of coffee [target goal]. She put on the kettle. She got milk out of the fridge. But then she saw that it was time to go to work. She had an important meeting that morning. She drank her coffee quickly [completed condition]; She decided that she would have her morning cup of coffee at work [interrupted condition]; She decided that she didn't have time for her morning cup of coffee [abandoned condition]. She got straight in her car. She drove straight to work. She arrived and went straight to her morning meeting. At the meeting, Betty's boss asked her to present her latest findings. She was nervous, but presented them confidently. Her presentation was well received. After the meeting, the boss offered Betty a promotion."

Procedure

The study was carried out using SurveyMonkey. Before starting the experiment, participants provided informed consent and indicated their age and gender. Participants then read a short story (~120 words) about a character named Betty.

After reading this story, the story was removed from the screen and participants were asked the following (in this order):

Free recall question: "Please summarise the story you just read using about 20 words."

Control question 1: "In the story, what was the main character's name?" (correct answer: Betty).

Control question 2: "Where was Betty at the end of the story?" (correct answer: at work).

The control questions were multiple choice and designed to filter out participants who had not read the story with sufficient attention. Participants were thanked at the end of the experiment for their participation and asked to provide their MTurk ID to evidence their completion of the study.

Coding

Coding was carried out by a research assistant who was naïve to the hypotheses of the study, and who assessed whether or not participants mentioned the target goal in their summaries and whether participants satisfied any exclusion criteria. For free recall, the research assistant coded whether participants used the word "coffee" in their summary, if they used a stand-in for coffee, such as "beverage", "drink", or "breakfast", or if they did not mention coffee or a stand-in at all. Only 21 participants used a stand-in for coffee (11 in the abandoned condition, seven in the completed condition, and three in the interrupted condition), so we collapsed these participants with those who mentioned coffee for subsequent analysis.

A second research assistant who was naïve to the hypotheses of the study coded a random 109 trials (20%) for reliability. Using the *Kappa.test* function of the *R* package *fmsb* (Nakazawa, 2019), coders were found to be in almost perfect agreement: coders agreed on whether participants mentioned the target goal in free recall on 97% of trials (κ = 0.94 (95% CI: 0.87, 1.00), $p < .001$) and on whether to drop trials in 94% of trials (κ = 0.90 (95% CI: 0.84, 0.97), $p < .001$).

See the supplementary materials (https://osf.io/2r3tz/?view_only=b651721714104008977960 2b5e1bb0d8) for this experiment's data, *R* code, and the full set of questions (and multiple-choice options) presented to participants.

4.2.2 Results

All analysis was conducted in *R* (R Development Core Team, 2018; Revelle, 2018). In total, 169 participants mentioned the target goal, while 159 participants did not (see *Table 4.2* for the breakdown by condition).

Table 4.2

Percentage of Participants who Mentioned the Target Goal in Free Recall, by Condition

Condition	Target goal in free recall (%)
Abandoned goal (*n* = 124)	59
Interrupted goal (*n* = 92)	47
Completed goal (*n* = 112)	47

In order to investigate whether goal status predicted goal saliency, we conducted a chi-squared test of independence because our dependent variable (whether participants mentioned the target goal in free recall or not) and independent variable (condition) were categorical (we conducted non-parametric equivalents of our pre-registered analysis because our data did not satisfy parametric assumptions). This analysis was conducted using the *chisq.test* function of the *R* package *stats* (R Development Core Team, 2018). The results indicated that there was no relation between condition and whether participants mentioned the target goal in free recall (χ^2 (2) = 4.32, p = .12).

4.2.3 Discussion

The results showed no significant effect of goal status on whether participants mentioned the target goal in free recall. This study therefore failed to find evidence of a relation between goal salience and goal status. This means that we failed to find evidence in support of either the goal slippage hypothesis or the predictive value hypothesis.

These results are puzzling given that previous studies found that incomplete goals are more salient than completed goals (Asiala et al., 2020; Lutz & Radvansky, 1997; Radvansky & Curiel, 1998; Suh & Trabasso,

1993; Trabasso & Suh, 1993). These results are also surprising given the results of our own pilot study, which showed the predicted tendency, especially because the only methodological difference between the present study and our pilot study was that the target goal was making a cup of tea instead of a cup of coffee.

We hypothesised that our failure to replicate the findings of past studies may be due to the target goal in our experiment - getting a cup of coffee. Specifically, participants may have believed that Betty's goal was completed in all conditions due to how easy it is to get a cup of coffee. Getting a cup of coffee can be done relatively easily while one is performing other activities. For instance, it is often easy to buy a cup of coffee from a cafe or a vending machine while on the way to work, or else to make one while at work. Participants may have believed that Betty did get the coffee by the time she went to her meeting, or else that she already had a plan for acquiring coffee with little hassle. As far as participants were concerned, the matter as to whether Betty is going to get her coffee may be settled (Bratman, 1999, 2009, 2014; Ludwig, 2007; Velleman, 1997). This would render our manipulation ineffective, thereby accounting for our failure to find evidence of a relation between goal status and goal salience. We therefore decided to adjust the story for Experiment 2. Specifically, we changed the target goal so that it was no longer something that could easily be achieved at work or on the way to work.

4.3 Experiment 2

Our aim in Experiment 2 was the same as for Experiment 1: we wanted to investigate whether goal status altered goal saliency.

We altered the test materials from Experiment 1 to address the concern that participants may have believed that the target goal was completed in all conditions. Instead of making coffee, Betty's target goal was now baking a cake.

4.3.1 Methods

Prior to data collection we pre-registered our planned analyses, hypotheses, sample size, exclusion criteria, and methods on *AsPredicted.org*

(http://aspredicted.org/blind.php?x=si7wv2). The University of Warwick's Humanities & Social Sciences Research Ethics Sub-committee (HSSREC) approved this experiment, which was conducted in accordance with the Declaration of Helsinki.

Participants

This experiment was conducted online using MTurk. We aimed for the same sample size as in Experiment 1 (at least 150 participants per condition, 450 participants overall) given the minimal changes between Experiments 1 and 2. We collected data from 626 MTurk workers.

In accordance with pre-registered exclusion criteria, participants were excluded from analysis if they failed at least one of the two control questions (n = 13), if they wrote more than 30 words (n = 101), or if they wrote nothing at all (n = 21). Our final sample (n = 491) consisted of 243 females (244 males, three "prefer not to say", one "other"), with a mean age of 39.75 years (range = 18.00 - 72.00 years, SD = 11.06 years)). Participants received monetary compensation ($0.75) for taking part.

Design

This experiment's design was the same as Experiment 1.

Materials

The following story was used as stimulus: "Betty had just woken up. She decided to bake a cake [target goal]. She put on an apron. She got milk out of the fridge. But then she saw that it was time to go to work. She had an important meeting that morning. She finished making the cake quickly [completed condition]; She decided that she would have to find time later [interrupted condition]; She decided that she would not make the cake after all [abandoned condition]. She got straight in her car. She drove straight to work. She arrived and went straight to her morning meeting. At the meeting, Betty's boss asked her to present her latest findings. She was nervous, but presented them confidently. Her presentation was well received. After the meeting, the boss offered Betty a promotion."

Procedure

The procedure for this experiment was identical to Experiment 1, except that the wording of the free recall question was changed to: "Please summarise the story you just read using 10 to 20 words." We made this change to reduce the number of participants excluded due to writing more than 30 words in free recall.

Coding

The coding procedure in this experiment was identical to that of Experiment 1. Thirty-seven participants mentioned a stand-in for cake (e.g., "cookie", "bake"): 13 in the abandoned condition, two in the completed condition, and 22 in the interrupted condition. We collapsed these participants with those who explicitly mentioned cake for subsequent analysis.

A second research assistant who was naïve to the hypotheses of the study coded a random 125 trials (20%) for reliability. Coders were in almost perfect agreement about whether the target goal was present in free recall (agreement on 95% of test trials (κ = 0.87 (95% CI: 0.77, 0.97), p < .001)) and whether to exclude trials (96% of trials (κ = 0.88 (95% CI: 0.78, 0.98), p < .001)). This experiment's data, *R* code, and experimental materials can be found in the supplementary materials (https://osf.io/2r3tz/?view_only=b651721714104008977960 2b5e1bb0d8).

4.3.2 Results

All analysis was conducted in *R* (R Development Core Team, 2018; Revelle, 2018). In total, 416 participants mentioned the target goal while 75 participants did not mention the target goal (see *Table 4.3* for a breakdown by condition).

Table 4.3

Percentage of Participants who Mentioned the Target Goal in Free Recall, by Condition

Condition	Target goal in free recall (%)
Abandoned goal (*n* = 184)	81
Interrupted goal (*n* = 157)	80
Completed goal (*n* = 150)	95

We ran the same analysis in this experiment as in Experiment 1 because the data were of the same type (we conducted non-parametric equivalents to our pre-registered analysis because our data did not satisfy parametric assumptions). The results of our chi-squared test of independence indicated that there was an association between condition and presence of the target goal in free recall (χ^2 (2) = 17.79, $p < .001$).

We computed *Cramer's V* as an effect size using the *CramerV* function in the *R* package *rcompanion* (Mangiafico, 2018). *Cramer's V* is a measure of association between two variables for a chi-squared test in which one of the variables has more than two levels (Kim, 2017; Mangiafico, 2016). *Cramer's V* is bounded between 0 and 1, with 1 indicating perfect association between the variables. The strength of association between condition and presence of the target goal in free recall was small to medium, *Cramer's V* = 0.19 (Akoglu, 2018; Cohen, 1988; Kim, 2017; Mangiafico, 2016).

We performed posthoc comparison tests using the function *pairwiseNominalIndependence* in the *R* package *rcompanion*. We adjusted for multiple comparisons using the Benjamini-Hochberg method of adjustment (Chen et al., 2017; Jafari & Ansari-Pour, 2019; Mangiafico,

2016), which controls the false discovery rate. We used this method because our hypotheses were not independent, and because it minimises false negatives as well as false positives. Results indicated that participants mentioned the target goal in free recall more often in the completed condition as compared to the interrupted condition ($p < .001$, *Cramer's V* = 0.23) and the abandoned condition ($p < .001$, *Cramer's V* = 0.22).

4.3.3 Discussion

The results showed an effect of goal status on presence of the target goal in free recall. Specifically, completed goals were mentioned in free recall more than interrupted or abandoned goals. The present experiment therefore found evidence of a relation between goal status and goal salience; namely, that completed goals were more salient than incomplete goals (both interrupted and abandoned).

This result was not predicted by either of our hypotheses, and it differed from past findings that incomplete goals are more salient than completed goals (Asiala et al., 2020; Lutz & Radvansky, 1997; Radvansky & Curiel, 1998; Suh & Trabasso, 1993; Trabasso & Suh, 1993). For these reasons this was an unexpected result.

We hypothesised that our results may have been caused by a factor that can boost goal salience independently of goal status - surprise. The target goal in this story is making a cake, i.e., a task that takes some time and planning. In the interrupted and abandoned conditions, Betty halts this activity when she realises that she is late for work. This is an appropriate response to realising that one is late for work. In the completed condition, however, Betty finishes making the cake before going to work. This may have struck participants as surprising, given that baking a cake is not a task that can be quickly finished before one leaves for work. Betty's unusual response in the completed condition may account for the greater salience of the target goal in that condition as compared to the other conditions. Surprise is therefore a factor that we wish to control for in Experiment 3, as it can boost salience independently of goal status (Bordalo et al., 2013; Feigenson, 2016; Green, 1956; Wasserman & Castro, 2005).

4.4 Experiment 3

In Experiment 3 we were again investigating whether goal status (abandoned, interrupted, or completed) predicted goal salience.

We made two notable changes to the materials used in Experiment 2. Firstly, participants were asked a question that directly probed their surprise at the protagonist's actions (Pekrun et al., 2017; Reisenzein et al., 2019). Secondly, we asked participants to read three stories instead of one, to increase the number of observations per participant.

4.4.1 Methods

This experiment was approved by the Humanities & Social Sciences Research Ethics Sub-committee (HSSREC) at the University of Warwick and was conducted in accordance with the declaration of Helsinki. *AsPredicted.org* was used to register hypotheses, methods, exclusion criteria, sample size, and planned analyses in advance of data collection (http://aspredicted.org/blind.php?x=mr2kt7).

Participants

For this experiment we requested data from 450 participants (150 per condition) on MTurk, collecting data from 435 MTurk workers (1305 trials). We excluded participants who failed at least one control question (n = 376) or who wrote more than 30 words (n = 329). Two hundred and seventy-eight participants made up our final sample (600 trials), 84 of whom were female (190 males, three "prefer not to say", one "other"). The mean age of our sample was 35.05 years (range: 18-69 years, SD = 9.89 years), and all participants received monetary compensation ($1.00) for participating.

Design

This experiment's design was the same as Experiment 1 except that each participant read and responded to three different stories.

Materials

The following three stories were used as stimulus: "Betty had just woken up. She got dressed. She decided to make some brownies [target goal]. She got the baking tray and prepared the brownie mixture. She went to pre-heat the oven. But then she realised that her oven was broken. She

decided to make the brownies in the microwave instead [completed condition]; She decided that she would need to wait until the oven was fixed [interrupted condition]; She decided that she would not make the brownies after all [abandoned condition]. Betty saw that it was time to go to work. She got into her car. She drove straight to work. She arrived and went straight to her morning meeting. At the meeting, Betty's boss asked her to present her latest findings. She was nervous, but presented them confidently. Her presentation was well received. After the meeting, the boss offered Betty a promotion."

"Susan was getting ready to go out. She was thinking about her upcoming date. She was looking for a sock that she had lost [target goal]. She looked in the laundry basket. She looked in the living room. She looked under the bed. But, either way, she needed to leave soon. Then she spotted the sock [completed condition]; She decided that she would need to find the sock later [interrupted condition]; She decided to forget the lost sock [abandoned condition]. She called a taxi and headed out. She told the driver to go to Central Square. Upon arriving she could see James standing in the rain. He had an umbrella and was wearing a nice suit. Susan paid the taxi driver and ran to meet James. He had bought tickets for the theatre. The atmosphere was very nice."

"Will was having a quiet night in. He had put on a film. He decided to drink some nice whisky [target goal]. He got the whisky glass out of the cupboard. He got some ice out of the freezer. Then he heard the phone ring – it was his friend, asking him to come out for the night. Will decided to have his nice whisky quickly before heading out [completed condition]; Will decided that he would have to find time for his nice whisky later [interrupted condition]; Will decided to have a drink with his friend instead of drinking his nice whisky tonight [abandoned condition]. Will put on his jacket. He grabbed his keys and wallet. He went outside. He flagged down the first taxi he saw. When he arrived he greeted his friend. They ordered some drinks. It was a nice bar. Will thought it was going to be a good night."

Procedure

The procedure for this experiment was identical to that of Experiments 1 and 2, except that participants read three stories and answered three corresponding sets of questions immediately after each story. Participants were also asked a third multiple-choice control question for each story after answering the free recall question ("What did [the main character] do when [turning point in the story]?") to ensure that they had paid sufficient attention to the status of the target goal. Finally, participants were asked for each story: "On the following scale, rate how surprising you found [the main character's] reaction to [turning point in the story]? ('0' indicates 'not at all surprising', '11' indicates 'extremely surprising')" (Boone & Boone, 2012; Derrick & White, 2017; Joshi et al., 2015; Sullivan & Artino, 2013; Willits et al., 2016; Wu & Leung, 2017).

Coding

Coding was carried out by the main researcher who was naïve to the condition of each trial, and was identical to the coding in Experiments 1 and 2 except that no distinction was made between the target goal and potential stand-ins for the target goal in free recall (i.e., rather than coding these responses separately and *then* combining them for analysis, they were combined from the start).

For reliability, a random 20% of trials (261) were coded by a research assistant who was naïve to the hypotheses of the study. Coders were in almost perfect agreement: coders agreed on 97% of trials as to whether the target goal was present in free recall (κ = 0.94 (95% CI: 0.89, 0.98), $p < .001$), and they agreed on whether to drop a trial on 99% of trials (κ = 0.99 (95% CI: 0.96, 1.01), $p < .001$). The data, *R* code, and full set of questions presented to participants can be found in the supplementary materials (https://osf.io/2r3tz/?view_only=b651721714104008977960 2b5e1bb0d8).

4.4.2 Results

All analysis was conducted in *R* (R Development Core Team, 2018; Revelle, 2018). Of the 600 test trials included for analysis, the target goal was mentioned in 411 of them (see *Table 4.4* for a breakdown by condition).

Table 4.4

Percentage of Trials in Which the Target Goal was Mentioned in Free Recall, by Condition

Condition	Target goal in free recall (%)
Abandoned goal (*n* = 176)	72
Completed goal (*n* = 201)	70
Interrupted goal (*n* = 223)	65

We first investigated whether surprise differed between conditions. This required calculating the average surprise rating per participant, and then conducting a Kruskal-Wallis one-way analysis of variance test. We ran this analysis instead of the pre-registered ANOVA because our data did not meet the assumptions of parametric testing. The *byf.shapiro* function of the *R* package *RVAideMemoire* (Hervé, 2020) indicated that surprise was not normally distributed in the abandoned condition ($W = 0.94$, $p < .001$) or the interrupted condition ($W = 0.94$, $p < .001$). In addition, the *leveneTest* function of the *R* package *stats* (R Development Core Team, 2018) revealed that variance was not homogeneous, F (2, 275) = 4.22, $p = .02$. For the Kruskal-Wallis ANOVA, our dependent variable was average surprise per participant, and the independent variable was goal status (interrupted, completed, abandoned): interrupted condition, Median = 5.67, interquartile range = 3.00-8.00; completed condition, Median = 5.5, interquartile range = 3.50-7.33; abandoned condition, Median = 5.67, interquartile range = 2.42-

8.00. This analysis was conducted using the *kruskal.test* function of the *R* package *stats* (R Development Core Team, 2018). Results indicated that there was no significant relation between surprise and condition, H (2) = 0.09, p = .96.

We then investigated the relation between goal status (condition) and goal salience (whether the target goal was mentioned in free recall on each trial) using linear mixed logistic regression. This analysis was conducted using the *mixed* function of the *R* package *afex* (Singmann et al., 2020). Our dependent variable was the presence of the target goal in free recall on each trial, and our fixed effect predictor was condition. Our random effects included the random intercept of story, as well as the random intercept of participant (Barr et al,. 2013; Bates et al., 2015; Singmann & Kellen, 2019). We initially included the random slope of story as a random effect, but we removed this because of singularity in the model. The full model was compared to a model that was identical except that the fixed effect (condition) was removed. The results indicated that the full model was not a better fit for the data as compared to the null model, χ^2 (2) = 1.74, p = .42. That is, condition was not found to be a significant predictor of whether the target goal was mentioned in free recall.

4.4.3 Discussion

The results showed no relation between the status of the target goal and whether participants mentioned that goal in free recall. This study therefore failed to find evidence of a relation between goal salience and goal status. These results provided no evidence in support of either the goal slippage hypothesis or the predictive value hypothesis. Although we found no evidence that surprise differed between conditions, our findings were still inconsistent with those of past studies (Asiala et al., 2020; Lutz & Radvansky, 1997; Radvansky & Curiel, 1998; Suh & Trabasso, 1993; Trabasso & Suh, 1993).

4.5 General Discussion

Across three experiments we investigated whether the status of a protagonist's goal (whether it was completed, interrupted, or abandoned)

affected whether that goal was mentioned by participants in free recall (indicating goal saliency). In Experiments 1 and 3 we failed to find evidence of any relation between goal status and whether participants mentioned the target goal in free recall. In Experiment 2 we found that participants mentioned the target goal more often in the completed condition as compared to either the interrupted or abandoned condition. This means that we only found evidence of a relation between goal status and goal saliency in Experiment 2. However, the results of Experiment 2 were not predicted by either of our hypotheses, and they were not consistent with past findings about the relation between goal status and goal salience (Asiala et al., 2020; Lutz & Radvansky, 1997; Magliano & Radvansky, 2001; Radvansky & Curiel, 1998; Suh & Trabasso, 1993; Trabasso & Suh, 1993).

This study makes two novel contributions that build upon past research. Firstly, this was the first study to differentiate between different types of incomplete goal – abandoned and interrupted. Secondly, drawing this distinction enabled the present study to compare the predictive value hypothesis to the goal slippage hypothesis, which is an alternative explanation for past findings.

There are several possibilities as to why we failed to replicate the results of past studies. One possibility relates to our exclusion criteria. Our experiments required participants to answer two or three control questions to ensure that they had paid sufficient attention to the materials. In addition, we excluded trials in which participants wrote no words at all or more than 30 words (to ensure that participants were selective in their summaries). In contrast to this, some past studies either implemented more minimal exclusion criteria (Asiala et al., 2020; Lutz & Radvansky, 1997) or else no exclusion criteria at all (Magliano & Radvansky, 2001; Suh & Trabasso, 1993; Trabasso & Suh, 1993). The potential problem of removing some or all of our exclusion criteria is that if our results indicated no effect of status on salience, we would be unable to identify whether this was due to there being no real underlying effect, or due our manipulation not affecting participants because they were not paying sufficient attention to our test

materials. However, including all participants for the main analysis did not change the results of any experiment (these results can be found in https://osf.io/2r3tz/?view_only=b65172171410400897796 02b5e1bb0d8).

It is worth reflecting, in particular, on the potential impact of limiting participants' summaries to 30 words. We made this decision to ensure that participants were indeed *summarising* the stories rather than simply writing out the entire story. Our concern was that without this restriction, participants would either write out a lengthy summary of the story or else copy and paste the story, in which case the target goal would be mentioned in all trials. This would render our experimental manipulation ineffective. However, implementing this restriction may have meant that particiapnts felt that they did not have enough words to discuss the target goal. Specifically, participants may have felt more confident in their ability to briefly summarise the target goal in only the *completed* condition; summarising the target goal in the completed condition (e.g., 'Betty baked a cake') may take fewer words than in the interrupted condition ('Betty started to bake a cake but had to stop because the oven was broken') or abandoned condition ('Betty started to bake a cake and then changed her mind'). If this were the case, we would predict the target goal to be mentioned more often in free recall in the completed condition than the other conditions. And indeed, in Experiment 2 we found evidence to support this prediction. However, we did not find the same evidence in Experiments 1 and 3, and the rates at which participants were mentioning the target goal in free recall differed substantially between experiments. It is therefore not clear whether the introduction of this exclusion criteria accounts for our results.

There are two possible changes that could be made to our stimuli to address concerns around the 30-word limit. The first is to remove this limit for participants' summaries of the stories entirely, or else raise it (e.g., to 60 words). Our original concern (that participants would not be incentivised to *summarise* the story without a strict word limit) would still be present if the 30-word limit were increased or removed entirely, but one could address

this original concern by introducing a second change: extending the length of the stories. The stories as presented in this study had a two-part structure, in which the target goal comes at the beginning and is followed by some other event (e.g., the events leading up to Betty's promotion). Extending the length of the stories would allow them to become more complex, such that the target goal might come in the middle of the story instead of the beginning. By extending the length of the stories participants may still be incentivised to summarise the stories even if their word limit were either increased or removed. Together, these changes may address concerns surrounding our 30-word limit.

A further possibility as to why we failed to replicate past findings is that our experimental manipulation may have been too subtle. The second half of each of our stories was unrelated to the target goal, and we manipulated only one sentence between conditions. In contrast, past studies typically change multiple sentences (two or three sentences) between conditions, in such a way that the status of the target goal influences how the protagonist's actions are understood in the second half of the story. For example, consider the story of Jimmy (Asiala et al., 2020; Lutz & Radvansky, 1997; Magliano & Radvansky, 2001; Suh & Trabasso, 1993; Trabasso & Suh, 1993), in which the target goal is Jimmy getting a bike. Jimmy asks his mother for a bike, and she either buys Jimmy a bike (completed condition) or else she does not give him a bike (incomplete condition). Regardless of condition, Jimmy then starts working at a local grocery store saving up money. In the incomplete condition, Jimmy uses these savings to buy a bike, whereas in the completed condition Jimmy uses these savings to buy a basketball. If the relevance of the target goal for the second half of the story differs between conditions, then the salience of the target goal may have differed between conditions independently of goal status. The subtlety of our experimental manipulation may therefore account for our results. However, the potential problem with making our manipulations less subtle is that the experiment would have been less well-controlled. That is, the subtlety of our manipulation enabled us to better

control for confounding factors (e.g., the relevance of the target goal for the second part of the story) that may affect salience independently of goal status.

A related factor that may have influenced the results of past studies is the presence of further factors that may have promoted goal saliency independently of goal status. In the texts used in past studies, the protagonist's emotional state often differs between conditions. Reconsider Jimmy and his goal to acquire a bike. In the completed condition participants are told that "Jimmy was very happy", whereas in the interrupted condition participants are told that "Jimmy was very sad". Jimmy's emotional state is related to the status of his goal (insofar as these emotional states are caused by the state of the goal), and emotional states can influence saliency independently of goal status (Carver & Vaccaro, 2007; Öhman et al., 2001). For example, negative emotional responses from other agents may be more salient than positive emotional responses, and therefore the causes of negative emotions (e.g., an incomplete goal) may be more salient than the causes of positive emotions (e.g., a complete goal). Our study differs from past studies insofar as we did not manipulate the protagonist's emotional states between conditions, which may have reduced the differences in goal saliency between conditions. If we had followed past studies in this regard, however, we may have introduced a confounding factor, and our experiments would therefore have been less well-controlled.

4.5.1 Future Directions

The first area for future research should be to take *precisely* the materials used in past studies, and to only adjust them such as to distinguish between interrupted and abandoned goals. This would raise the chances of replicating the results of past studies, whilst also enabling researchers to test the two hypotheses considered in this study. In follow-up studies, these materials could then be iteratively altered to eliminate confounding factors that may be present in the materials of earlier studies.

A related direction for future research would be to use different measures of goal saliency. The measure of goal saliency that we

implemented was whether participants mentioned the target goal in free recall (Asiala et al., 2020; Kurby & Zacks, 2012; Suh & Trabasso, 1993, Trabasso & Suh, 1993). However, other measures of goal salience have been used in other studies, such as response time or and accuracy in answering questions related to a target goal (Lutz & Radvansky, 1997; Radvansky & Curiel, 1998). The generalisability of results would be strengthened if they were replicated across different measures of goal saliency.

A further avenue for future research would be to investigate the replicability of past findings using live-action materials. A similar experiment could be conducted in which participants watch an actor perform goal-directed actions that are either completed, abandoned or interrupted, before summarising the actor's behaviours. This would build upon recent research which has investigated narrative and action comprehension using visual materials (Cohn & Bender, 2017; Eisenberg et al., 2018; Kurby & Zacks, 2019; Magliano et al., 2014), and which have not distinguished between abandoned and interrupted goals. Such an investigation would improve the ecological validity of the present study because in everyday life we mostly encounter goals in live-action scenarios.

Using live-action materials may also provide insight into the underlying representational format and mechanisms that underpin goal tracking. This is because live-action materials would enable the activation of representational formats and cognitive mechanisms that are not available with text-based materials, such as motor representations and motor contagion (Bouquet et al., 2011; Flanagan & Johansson, 2003; Paulus, 2014; Pezzulo et al., 2018; Zwaan, 2016). One hypothesis that could be tested using such methods is that visual representational formats may enhance the effects of goal status on goal salience (relative to cases in which text-based materials are used). This hypothesis is motivated by the observation that participants watching live-action goal-directed activities may be better able to imagine how they themselves would complete goals that are not yet accomplished (Kilner et al., 2003; Rizzolatti & Sinigaglia, 2016). Imagining

oneself achieving a goal might increase the saliency of that goal. Conducting similar investigations with live-action materials may therefore enable the testing of further hypotheses relating to goal status, goal saliency and goal tracking.

Finally, this study raises further questions about the relation between goal status and practical planning. The main methodological change that was introduced between Experiment 1 and Experiment 2 related to the protagonist's target goal. In Experiment 1, this target goal was to make coffee, and the results from Experiment 1 indicated that there was no effect of goal status on goal salience. We hypothesised that this was because participants may have believed that the goal was accomplished in all conditions due to the relative ease with which one can acquire coffee. That is, we believed that participants may have taken the matter to be settled, and for the goal to be as good as completed (Bratman, 2014; Ludwig, 2007; Velleman, 1997). This line of reasoning generates testable hypotheses about the role of practical planning in how agents view the status of goals: an agent being in a position to complete a target goal with relative ease, or having a specific plan in mind for completing the target goal, may be sufficient for observers to treat the goal as if it were already completed. Testing such hypotheses would enhance our understanding of the factors that mediate the relation between goal status and goal salience, as well as providing insight into how agents determine the status of goals.

4.5.2 Conclusion

This chapter built upon earlier research that demonstrated that incomplete goals are more salient than completed goals by distinguishing between two types of incomplete goal (abandoned goals and interrupted goals). This chapter tested two hypotheses about why goal status affects goal salience: the goal slippage hypothesis and the predictive value hypothesis. We failed to find any effect of goal status on goal salience. Nevertheless, this study is valuable insofar as it constitutes an investigation into the cognitive mechanisms that underpin goal tracking, and we have identified fruitful directions for future research.

This chapter contributes to this overall thesis by investigating the implications of goal slippage for the broader research landscape. This chapter demonstrates that even beyond the context of instrumental helping in early childhood, in which this hypothesis is typically considered (Michael & Székely, 2019; Paulus, 2014, 2018), goal slippage can generate valuable new hypotheses and offer new ways of interpreting phenomena that have been the focus of past research.

Chapter 5

A Reductive Second-Person Approach to the Development of Interpersonal Commitment

Previous research on interpersonal commitment in early childhood has focused on differences between age groups (Gräfenhain et al., 2013; Kachel et al., 2017, 2019) rather than accounting for the developmental trajectory of one's ability to honour an interpersonal commitment over the course of childhood. I adopt a two-step approach to investigating the development of interpersonal commitments. First, I establish a reductive theoretical framework which enables questions about interpersonal commitment to be rephrased in terms of the components of interpersonal commitment (e.g., cognitive ability and prosocial motivation). The second step is to adopt a second-person approach by explicitly considering the role of social interaction in the development of interpersonal commitment. Specifically, I articulate the hypothesis that social interaction scaffolds the development of interpersonal commitment by introducing immediate rewards for employing executive function and contributing to others' goals. These rewards are others' emotional expressions (which one may find pleasant or aversive), and they reinforce helping behaviour and executive function use over childhood. Although this hypothesis is yet to be directly tested, it generates predictions that are consistent with empirical findings. I close by considering the implications of this hypothesis for atypically developing humans and non-human animals.

5.1 Introduction

Interpersonal commitments are a foundational aspect of human social life. We form commitments with our neighbours, our friends and family, and even our colleagues and places of work. For instance, we may commit to honouring the terms of our job contract, or promise to help a friend move house. An interpersonal commitment can broadly be understood as a relation between two agents and an action, under which one agent has an obligation to perform the action (Michael & Salice, 2017). Interpersonal commitments can be contrasted with individual commitments,

such as committing to exercise regularly, under which an agent makes a commitment to themselves to perform an action (Clark, 2006; Michael & Pacherie, 2015).

Interpersonal commitments are foundational in human social life because of their role in shaping one's motivations and others' expectations about one's future behaviour. Forming an interpersonal commitment enables others to trust that we will perform actions to which we have committed, and other agents can thus take into account the performance of such actions in their own future plans (Bratman, 1999; Michael & Pacherie, 2015; Michael & Székely, 2018). Forming a commitment can boost one's motivations to perform the actions to which one has committed, facilitating persistence (Michael & Pacherie, 2015) and further motivating one to act in a way that is consistent with others' expectations (Bonalumi et al., 2019; Heintz et al., 2015). Interpersonal commitments thus promote coordination and cooperation between agents (Bratman, 2014; Gilbert, 2018; Michael, Knoblich, & Sebanz, 2016).

The degree to which interpersonal commitments underpin everyday social life has prompted researchers to investigate the developmental origins of interpersonal commitments. Such research is important because it provides insight into the cognitive architecture of interpersonal commitment in adulthood, generating testable hypotheses about the factors that modulate commitment (Michael, Knoblich & Sebanz, 2016). Research into the emergence of commitment may also contribute to ongoing research into the emergence of prosocial behaviour and motivation more broadly (Dahl & Brownell, 2019; Köster et al., 2019; Melis, 2018; Paulus, 2018; Sommerville et al., 2018; Spinrad & Gal, 2018; Wan et al., 2018).

Most of the empirical research on the development of interpersonal commitment focuses on children's expectations about others' commitment to an action (Astington, 1988a, 1988b; Chin & Lin, 2018; Gräfenhain et al., 2009; Mant & Perner, 1988; Siposova et al., 2018). Such studies have focused on children's understanding of how a commitment may be dissolved (Kachel et al., 2019), and children's protests when faced with an

agent who has failed to honour a commitment (Hussar & Hovarth, 2013; Kachel et al., 2017; Kanngiesser et al., 2017).

Other studies have focused on children's own ability to honour an interpersonal commitment (Gräfenhain et al., 2009; Kachel & Tomasello, 2019; Kanngiesser et al., 2017). These studies indicate that children's ability to honour interpersonal commitments improves dramatically between the ages of 3-5 years old (Gräfenhain et al., 2009, 2013; Kachel & Tomasello, 2019; Kanngiesser et al., 2017), and children continue to develop a more nuanced and complex understanding of commitments into late childhood (Chin & Lin, 2018; Hussar & Hovarth, 2013; Siposova et al, 2018).

There has been relatively little focus on explaining the developmental trajectory of children's own ability to honour an interpersonal commitment, however. Previous studies that have focused on children's ability to honour an interpersonal commitment have employed cross-sectional designs which enable them to investigate differences between age groups without focusing on accounting for the development of such differences (Gräfenhain et al., 2009; Kachel & Tomasello 2019; Kanngiesser et al., 2017). The developmental trajectory of children's ability to honour interpersonal commitment has also been relatively neglected in theoretical investigations, with authors primarily focusing on the cognitive mechanisms that motivate children to honour interpersonal commitments (Castro & Pacherie, 2020) and the development of children's conceptual understanding of interpersonal commitments (Michael & Székely, 2018). While these theoretical accounts might help explain the developmental trajectory of commitments, it is not primarily their focus. There is therefore a relative lack of research into the developmental trajectory of children's own abilities to honour interpersonal commitments over childhood.

I will adopt a two-step approach to address this gap in the literature. I will first introduce a theoretical framework for investigating questions about interpersonal commitment, reducing the phenomenon of interpersonal commitment into separate components - cognitive, motivational, and practical. This new conceptual framework for investigating the development

of interpersonal commitment enables questions about the development of interpersonal commitment to be reframed as questions about the development of these separate components.

The second step is to adopt a second-person approach to the development of children's abilities to honour an interpersonal commitment (Carpendale & Lewis, 2004; Moore & Barresi, 2017; Siposova & Carpenter, 2019; Tomasello, 2019). This approach emphasises the role of social interaction (e.g., with caregivers, siblings, and other children) in children's socio-cognitive development. This approach can be contrasted with a third-person perspective (under which children's role in their own socio-cognitive development is that of a purely passive observer) and a first-person perspective (under which children are viewed as individuals acting in isolation). This chapter therefore focuses on the development of interpersonal commitment within the context of social interaction.

5.1.1 A Reductive Theoretical Framework

What is required for an individual agent to honour an interpersonal commitment? If I make a promise to my friend to meet them for lunch, what is required for me to adhere to that commitment? And if I fail to honour the commitment, then what might have gone wrong?

In everyday life, the ability to honour an interpersonal commitment is often determined by one's cognitive ability to persist with boring or effortful tasks. This may involve inhibiting impulses to defect and perform alternative tasks, or it may involve focusing one's attention on the task to which one has committed (Rachlin, 2016; Székely & Michael, 2018). For instance, I may have failed to honour my commitment to meet my friend for lunch because I failed to resist the temptation to stay at home instead. In this case my commitment has failed due to my lacking sufficient cognitive ability to shield my long-term goal (meeting my friend for lunch) from short-term tempting alternatives. There is therefore a cognitive component that is essential in ensuring that agents honour interpersonal commitments.

This cognitive component of one's ability to honour an interpersonal commitment can broadly be understood as executive function (Michael &

Székely, 2018). Executive function refers to a set of cognitive processes that relate to the self-regulation of one's behaviour (Cohen et al., 2007; Diamond, 2013; Doebel, 2019; Miyake & Friedman, 2012; Nigg, 2017). Such executive functions that are relevant for one's ability to honour a commitment include inhibitory control (i.e., being able to resist tempting alternatives to the course of action that one is currently pursuing), attentional control (i.e., being able to maintain one's attention on the task at hand and the aspects of one's environment that are most relevant for that task), and working memory (being able to retain information relevant to one's current task in the forefront of one's mind). When an agent fails to honour a commitment in everyday life, it is often a result of a failure in executive function. For instance, I alternatively might not meet my friend for lunch because I forgot about our agreement to meet for lunch, or else because I failed to pay sufficient attention to the time (and I therefore missed our meeting). Indeed, in everyday life, when we talk about an individual who is particularly committed to a task, we are often referring to the agent's ability to resist temptations, inhibit impulsive behaviours, and persist with the given task (Michael, Knoblich, & Sebanz, 2016). This cognitive component may not be sufficient in ensuring that agents honour interpersonal commitments, however.

In addition to this cognitive component, honouring an interpersonal commitment also requires that an agent be sufficiently motivated to perform the action to which they are committed. That is, one has to care enough about the action or agent to which one is committed in order for the commitment to be fulfilled. For example, I may fail to honour the commitment to have lunch with an acquaintance if I simply do not want to have lunch (e.g., because I am not hungry), or if I do not care about the impact that reneging on this commitment will have on this acquaintance. Even if I am cognitively able to resist temptations in order to persist with the relevant task, I still need to be sufficiently motivated to do so.

Honouring an interpersonal commitment therefore also requires prosocial motivation (Bonalumi et al., 2019; Castro & Pacherie, 2020). In

this context, prosocial motivation is simply motivation to perform actions that other agents want or need to be performed (at least in part) because the other agent wants or needs the action to be performed (Batson & Powell, 2003; Decety et al., 2016; Eisenberg, 1982; Jensen, 2016; Staub, 1978). Note that under this minimal definition, the fact that another agent wants or needs an action to be performed does not need to be the only reason for which I am motivated to perform that action. For instance, I might contribute to your goal (thereby performing an action that you want me to perform) because I know that it will improve your mood, which may elicit a more positive social interaction between us than would otherwise have occurred.

Motivation to behave prosocially can come from a variety of different sources. One source of prosocial motivation that has already been the focus of research on the development of interpersonal commitments is one's sense of obligation to others that emerges as children come to recognise their place within society and the behavioural requirements that come with this (Dahl & Paulus, 2019; Tomasello, 2019). Children's understanding of societal norms emerges from around 2-3 years of age (Dahl et al., 2020), and this understanding both motivates children to adhere to norms themselves and leads children to expect others to do the same (Gräfenhain et al., 2009; Kachel & Tomasello, 2019; Kanngiesser et al., 2017). As it is the norm to follow through with actions to which one is committed, this source of prosocial motivation plays an important role in the development of children's ability to honour an interpersonal commitment. For example, from 3 years of age children are more motivated to persist with a collaborative task and resist the temptation to defect (for additional individual reward) following an explicit commitment to their collaborative partner, as compared to when no explicit commitment was made (Kachel & Tomasello, 2019). In addition, 3- and 5-year-olds who made a promise to finish a cleaning-up task were more motivated to resist distractions than if no promise had been made, and they were more likely to appeal to their obligation to complete the task in justifying their behaviour

(Kanngiesser et al., 2017). Even in the absence of an explicit promise, children may still be motivated to behave prosocially due to a sense of obligation to others. For example, 3- and 4-year-olds were more likely to acknowledge their partner when disengaging from a collaborative game (e.g., by telling the partner that they are going to disengage from the game before doing so) if they had invited the partner to play the game with them, as compared to if the partner had joined in the game unprompted (Gräfenhain et al., 2009). Taken together, these findings indicate that an emerging source of prosocial motivation (especially in the context of interpersonal commitment) comes from children's developing sense of obligation to others. While the research on interpersonal commitments in early childhood has focused on the emergence of this sense of obligation, it is important to note that there are further sources of prosocial motivation that may be relevant for the emergence of one's ability to honour an interpersonal commitment (discussed in section 5.2).

Even if I am sufficiently motivated and cognitively able to honour an interpersonal commitment, I still need to have the practical means of doing so. That is, I need to be physically able to perform the actions that would enable me to honour a commitment. For example, in order to meet my friend for lunch, I still need to be able to travel to that location.

There are several developmentally significant factors that feed into this component of honouring an interpersonal commitment. The early emergence of children's ability to track others' goals is particularly relevant for understanding what actions one must perform to honour a commitment, because goal tracking enables one to understand and identify what outcomes others want you to contribute towards (Csibra, 2008). For instance, if I promise to help a friend move house, I must be able to identify the specific goal that my friend wants me to contribute towards (e.g., moving furniture from their old home to their new home). The development of children's understanding of the relation between goals and the means by which goals can be achieved underpins their understanding of how to perform the actions to which they have committed. This develops from early in the first year of

life (Liu & Spelke, 2017), and accounts for at least some degree of variability in prosocial behaviour in early childhood (Bridgers & Gweon, 2018).

A further practical consideration for honouring an interpersonal commitment is one's physical ability to perform the actions to which one has committed. This is especially important in early childhood as children are experiencing rapid physiological development. Physiological limitations are especially relevant in development because fine motor skills develop over the course of early childhood, and physiological limitations are a barrier for children's ability to behave prosocially (Köster et al., 2019). That is, even if children understand what to do and how to do it, performing the relevant action is often determined by physiological development.

I have thus far identified three components that are necessary for an agent to honour an interpersonal commitment – cognitive, motivational, and practical (see *Table 5.1*). I will set aside the practical component because much of the practical component is a matter of biological development. That is, the first two of these components – cognitive and motivational – are the most relevant within the context of socio-cognitive development.

Adopting this reductive theoretical framework carries several benefits. By bringing together these components, this framework provides new ways in which questions about the development of commitment can be answered. For instance, the separate components identified in this theoretical framework could be independently manipulated, such that their relative influence on one's ability to honour an interpersonal commitment can be identified and examined.

The main advantage of this theoretical framework for present purposes is that it is now possible to investigate questions about the development of commitment by investigating the development of each of these components: how do cognitive and motivational capacities for interpersonal commitment (executive function and prosocial motivation, respectively) develop in early childhood? Breaking down questions about interpersonal commitment into questions about these different components

is an approach that has yet to be adopted in research on interpersonal commitment. As we shall see, adopting this framework generates novel hypotheses regarding interpersonal commitment in atypically developing populations and non-human animals.

Table 5.1

Reductive Theoretical Framework for Honouring an Interpersonal Commitment

Component	Examples of corresponding abilities
Cognitive	Executive function (e.g., inhibitory control, attentional control)
(Prosocial) Motivation	Concern for others' well-being, a desire to engage in positively-valenced social interactions, sense of obligation
Practical	Fine motor skills, ability to travel on one's own, goal tracking

5.1.2 The Second-Person Perspective

I will consider the role of social interaction in the development of interpersonal commitment over the course of childhood (Moore & Barresi, 2017). I will articulate and explore evidence for the hypothesis that social interaction scaffolds the development of interpersonal commitment by scaffolding the development of executive function and prosocial motivation. This is to say that social interaction promotes the development of these components, such that they develop further and faster for children who engage in relatively more and/or higher quality social interaction than for those who engage in relatively less and/or lower quality social interaction, although social interaction may not be necessary for the development of executive function or prosocial motivation. What exactly 'higher quality social interaction' means will depend, of course, on the specific way in

which social interaction facilitates the development of interpersonal commitment. Spelling out this hypothesis will be instructive in several debates and carry implications for both ontogenetic and phylogenetic research.

More specifically, the hypothesis that I will evaluate is that social interaction scaffolds the development of executive function and prosocial motivation by altering the immediate rewards of performing certain actions and exhibiting certain behaviours. In particular, social interaction alters the rewards of behaving prosocially and employing executive function, encouraging these behaviours and raising the costs of *not* acting prosocially and employing executive function. This change in rewards reinforces these behaviours over time, to the point that these behaviours are internalised (Aknin et al., 2018; Godman et al., 2014).

In order for a behaviour to be *reinforced* by social interaction over time, it is key that social interaction changes immediate rewards (Dickinson, 1980). Behaviours are most effectively reinforced if the temporal gap between behaviour and reward is minimised, because this enables agents to form a stronger association between the reward and the behaviour than if there is a greater temporal gap between behaviour and reward (Decker et al., 2016; Delin & Baumeister, 1994; Dickinson, 1980; Miller et al., 2019). For example, teaching my dog to sit will be most effective if I reward my dog with praise immediately after she sits, as compared to rewarding her several minutes after she has sat down.

This temporal relation between behaviour and reward is particularly important in the context of commitments. In everyday life, the difficulty of committing to a task is often due to delays in the rewards associated with the task (Green & Myerson, 2004; Stevenson, 1986). For example, the costs of failing to inhibit one's impulse to eat a slice of cake (i.e., gaining weight) are not felt immediately, which partly accounts for the difficulty that many individuals have with committing to a diet (Critchfield & Collins, 2001; Hayden, 2016; Reynolds, 2006; Stevenson, 1986; Steel & König, 2006; Steel et al., 2018). Altering the rewards of this behaviour such that the costs

of eating cake are immediate (e.g., imposing an immediate financial penalty per slice of cake) would increase one's motivation to resist this impulse, thereby improving the chances that one would adhere to their diet (Odum & Rainaud, 2003; Story et al, 2014).

I will propose that in social interactions, others' positive and negative emotions serve as immediate positive and negative rewards (respectively). This will be further articulated in the following sections, but in brief, the suggestion is that the emotions that others exhibit in response to a child behaving prosocially or employing executive function in a joint task (or failing to do these things) alter the immediate rewards of exhibiting these behaviours. This is motivated by the finding that children's sensitivity to others' emotions is one of the earliest developing and most fundamental aspects of social interaction (Blair, 2017; Hobson, 2002; Reddy, 2008). Affective engagement is therefore key in shaping young children's early understanding of the world. The core message of this chapter, then, is that others' emotions are foundational in how social interaction scaffolds the development of interpersonal commitment.

One might wonder why social interaction is so important under this hypothesis – why think that others' emotions will primarily reinforce prosocial motivation and executive function use in early childhood in the context of social interaction? The reason for this is that it is in the context of social interaction that others' emotions are most salient (i.e., when one is being directly confronted by others' emotions) due to the affective engagement that comes with social interactions (Moore & Barresi, 2017). This means that it is primarily in the context of social interaction that children are directly faced with others' emotions, and therefore the link between others' emotions and one's own actions (or inaction) is most clear in this context (Moore & Barresi, 2017). It is thus in the context of social interaction that children will be most strongly motivated by others' emotions.

The present hypothesis is not that others' emotions will *only* serve to motivate and reinforce behaviour when children are engaged in social

interaction with those individuals; thinking about others' emotions or merely observing others' emotions in response to children's actions (or inactions) may also motivate and reinforce behaviours. As adults, for instance, we often imagine others' future emotional responses to our behaviour, and we can thus be motivated by imagining others' emotions. I might be motivated to complete a work task because I have imagined how proud my line manager or colleague may be of me, or because I wish to avoid my line manager's disappointment if I do not finish the task on time. Instead, this chapter specifically proposes that it is in the context of social interaction that others' emotions will be most effective in motivating and reinforcing behaviour over the course of childhood because of the increased salience of others' emotions, and their link to our actions, in this context.

A further important point to clarify is what the present hypothesis aims to explain. The focus of this chapter is the developmental trajectory of children's ability to honour an interpersonal commitment over the course of early childhood, i.e., the fact that children become substantially better at honouring commitments as they age. This hypothesis aims to account for how prosocial motivation and executive function use is *reinforced* over the course of early childhood, rather than accounting for the first instance(s) of prosociality or executive function use. The suggestion is that whatever motivates prosociality (e.g., see section 1.1.2) and executive function use initially, social interaction can reinforce these behaviours due to others' emotions.

5.2 Prosocial Motivation Scaffolded by Social Interaction

There are a variety of ways in which prosocial behaviours can be scaffolded by social interaction in early childhood. Social interaction between child and caregiver can promote the practical planning and execution of helping behaviours, that is, caregivers can illustrate *how* and *when* children can help other agents through verbal explanation, physical demonstration, and/or modelling. For example, caregivers might point to the target object that the helpee requires, or tell children that the helpee requires their help. Understanding how to help others accounts for at least some

variation in early childhood prosociality (Bridgers & Gweon, 2018; Köster et al., 2019), and this is one of the ways in which prosocial behaviour can be scaffolded by social interaction (Pettygrove et al., 2013; Schuhmacher et al., 2018).

Opinions are divided as to whether the development of prosocial *motivation* is scaffolded by social interaction, however. Some authors hypothesise that prosocial motivation emerges without scaffolding and is not promoted by social interaction (Tomasello & Vaish, 2013; Warneken, 2015; Warneken & Tomasello, 2013). This view is motivated by the emergence of prosocial behaviours, such as spontaneous instrumental helping, by the second year of life (Warneken & Tomasello, 2006). These authors suggest that parents are not likely to encourage and reinforce helping behaviour this early in childhood, and therefore children's motivation to engage in prosocial behaviours must emerge independently of social interaction. The opposing view (Dahl, 2015; Dahl, 2018) instead suggests that social experiences do shape prosocial motivation from early in childhood, and that the motivation for prosocial behaviours is scaffolded through social interaction even from the first year of life (Brownell et al., 2013; Dahl, 2015).

I will articulate and evaluate evidence for a particular version of the latter view: that social interaction scaffolds the development of prosocial motivation in early childhood because others' emotions alter the immediate rewards of contributing to others' goals. That is, the social context brings additional social rewards and social costs (in the form of helpee's emotions) for behaving prosocially. This altered reward structure increases children's motivation for contributing to others' goals, reinforcing helping behaviour over time. Again, note that this hypothesis is not aiming to explain the initial emergence of helping behaviour, but rather how helping behaviour can be reinforced over the course of childhood.

Before evaluating evidence for this hypothesis it is important to note one limitation of the present discussion. There are many different kinds of prosocial behaviour that emerge in early childhood, such as instrumental

helping, sharing, and comforting, and there are distinct developmental trajectories for these behaviours (Dahl & Brownell, 2019; Dunfield, 2014; Paulus, 2018). As such, social interaction may scaffold these prosocial behaviours differently. I will therefore focus exclusively on instrumental helping (i.e., deliberately contributing to another agent's goal). I am focusing on this type of behaviour because it is one of the earliest emerging forms of prosocial behaviour (Callaghan & Corbit, 2018; Dunfield, 2014; Dunfield & Kuhlmeier, 2010; Hammond, 2014; Jensen, 2016; Paulus, 2018; Svetlova et al., 2010), and there is a wealth of studies focusing on instrumental helping in early childhood (Cirelli et al., 2016; Hepach et al., 2017; Hobbs & Spelke, 2015; Kenward & Gredebäck, 2013; Warneken & Tomasello, 2006), including in the context of social interaction (Barragan & Dweck, 2014; Brownell et al., 2013; Carpendale et al., 2015; Dahl, 2015, 2018; Hammond & Carpendale, 2015; Wan et al., 2018). Instrumental helping is also especially relevant for interpersonal commitment, because commitments in everyday life often involve contributing to another agent's goals (e.g., promising to meet a friend for lunch). Therefore while it may be more typical to focus on obligation and normativity when discussing interpersonal commitments, the prosocial behaviour that I will focus on for present purposes is instrumental helping.

5.2.1 Others' Emotions as Motivating Behaviour

In everyday life, others' emotional responses to our behaviour can motivate particular behaviours because others' emotions can be pleasant or aversive. We may be motivated to behave in a certain way in order to proactively avoid another agent exhibiting an unpleasant negative emotion towards us, or in order to alleviate a negative emotion that is currently being directed at us. For instance, we may be motivated to finish our chores because our parents have told us off for not doing so already, and are therefore disappointed with us. We may alternatively be motivated to perform certain actions because we expect others to consequently direct positive emotions towards us, which we would find pleasant. For example, we might want to perform well on an important presentation at work

because our supervisor will then be proud and tell us that we have done a good job. Such expectations may motivate us to perform actions that we would not perform otherwise or that we may be less likely to be motivated to perform, such as doing chores or preparing a presentation. In this way, others' emotions can act as a source of hedonic motivation for our behaviour.

I will first discuss how others' emotions may motivate behaviour in early childhood, before turning to the context of instrumental helping. For others' emotions to motivate behaviour in early childhood, children must (i) identify positive and negative emotions, (ii) form an association between the emotions and the event or objects to which they are directed, and (iii) find others' emotional displays pleasant or aversive.

Children meet criterion (i) by the end of the first year of life. At 7 months of age, children discriminate between basic emotions on the basis of facial expression. Specifically, 7-month-olds discriminate between happy and fearful facial expressions (Nelson et al., 1979; Nelson & Dolgin, 1985), as well as between angry, fearful, and surprised facial expressions (Serrano et al., 1992). At this age children can even identify emotional expressions such as anger on the basis of subtle facial cues (Ichikawa & Yamaguchi, 2014), suggesting that they are sensitive to emotional displays at a fine-grained level from early in life. Children are also able to discriminate between affect in vocalisations in the first year of life. From as early as the first few days of life, neuronal activity is sensitive to the distinction between happy and sad vocalisations (Cheng et al., 2012). The first criterion is therefore met well before the end of the first year of life.

Criterion (ii) is also met by the first year of life. By 12 months of age, children form expectations about the relation between others' emotions and the status of goals that others are pursuing. Specifically, children are surprised when a negative emotion follows an agent successfully completing a goal, and when a positive emotion follows a failed attempt to achieve a goal (Reschke et al., 2017; Skerry & Spelke, 2014; see also Chiarella & Poulin-Dubois, 2013; Hepach & Westermann, 2013; Ruba et al., 2019). A

further example of (ii) in the first year of life is social referencing. When put in an unfamiliar situation (e.g., when facing a novel object), children's behaviour is sensitive to the emotions exhibited by caregivers. If caregivers exhibit negative emotions (such as fear), children will be less likely to engage with novel objects and approach unfamiliar locations than if caregivers exhibit neutral or positive emotions (Carver & Vaccaro, 2007; Vaish et al., 2008; Vaish & Striano, 2004). It is worth noting that in these studies, others' emotional displays are directed at the goal objects or locations rather than children. This suggests that children from the first year of life associate emotions with the events to which they are directed, and even form expectations about others' emotional responses on the basis of goal status.

It may seem evident that children satisfy criterion (iii) by the end of the first year of life just from anecdotal evidence. The suggestion that children enjoy positive emotions from others (e.g. smiling or laughing) and would rather avoid negative emotions (such as anger) might seem fairly obvious. For instance, when a child expects that their past behaviour will lead to them being scolded (e.g., sneakily eating some sugary items before dinner, breaking a plate, or spilling a glass of water), they may try to hide these mistakes or else avoid interacting with others (e.g., by hiding).

There is also empirical support for children satisfying (iii) by the end of the first year of life. For instance, from 9 months of age children exhibit negative affect (as indicated by crying and distressed facial expressions) upon hearing another agent cry (Geangu et al., 2010), suggesting that children find negatively-valenced affective stimuli unpleasant by this age. Similar outcomes have been found with children at 12 (Nichols et al., 2015), 18 (Bandstra et al., 2011), and 24 (Ruffman et al., 2019) months of age. In contrast, from infancy there seems to be a preference for positive emotional expression. By six months of age, children have a preference (indicated by increased looking time) for positive affect as compared to negative or neutral affect in facial expressions (Kim & Johnson, 2013; Serrano et al., 1995) and vocalisation (Fernald, 1993; Singh et al., 2002). It should be

noted that in these studies, the emotions in question were not directed *at* the participants themselves, and they therefore do not specifically relate to the importance of the second person position. Nevertheless, they do provide evidence that children find others' positive emotions pleasant and negative emotions aversive by the end of the first year of life.

5.2.2 Others' Emotions as Motivating *Helping* Behaviour

How might another agent's emotions influence one's decision to help that agent? Imagine that you see an agent struggling to open a door. If you do not help this agent she may exhibit some negative emotions (such as frustration or anger) towards you, but if you help her then she is likely to exhibit some positive emotions (such as happiness or gratitude) towards you. These emotions will alter the rewards of opening the door because they may be aversive or pleasant. This agent's emotions thus increase the rewards of opening the door and raise the costs of *not* opening the door, as compared to a situation in which she is not trying to open the door. This is how the emotional responses of other agents can alter the rewards and costs of performing different actions, thereby altering the rewards of contributing to others' goals. Of course, in such a scenario the helpee's emotions may not be directed at *you* (e.g., the agent may simply be upset at her failed goal, or happy when it is completed), but the pleasantness or unpleasantness of these emotions will likely be more salient, and therefore better at motivating you to open the door, if the emotion is directed at you.

The hypothesis that others' emotions can motivate helping behaviour and thereby scaffold the emergence of prosocial motivation is well supported in theoretical research. Positive emotional responses, such as praise or encouragement from caregivers after children have helped an agent (Aknin et al., 2018; Brownell, 2016; Dahl, 2018; Dahl et al., 2011; Dahl & Brownell, 2019; Dahl & Paulus, 2019; Paulus, 2019; Spinrad & Gal, 2018; Warneken & Tomasello, 2009), as well as negative emotions, such as distress from the failure to achieve a goal (Hammond & Drummond, 2019; Martin & Olson, 2015; Preston & de Waal, 2002), have been argued to promote helping behaviour in early childhood (although it should be noted

that these theoretical accounts typically do not consider negative emotions as directed at children, e.g., scolding). These may promote helping behaviour because seeing others happy may improve one's own mood or increase the joy of helping others (Dahl et al., 2017; Paulus, 2019), or else one may help others in order to alleviate the unpleasantness of seeing others in distress (Batson, 2010).

The empirical research that bears upon this hypothesis falls into two categories: lab-based and ecological. Evidence from lab-based studies is mixed, insofar as some studies point towards a relation between others' emotions and helping behaviour whilst other studies fail to find such an effect.

With regard to positive emotions, emotional responses following children's helping behaviours can increase children's motivation to engage in helping behaviours in future. For example, praise and encouragement from caregivers led 13- to 14-month-olds to help twice as often compared to a control group with no reinforcement (Dahl et al., 2017), and the same effect has been found for 18- to 24-month-olds (Hammond & Carpendale, 2015; Kärtner et al., 2020). In addition to this, praise of helping behaviour at 15 months of age has been shown to increase instrumental helping at 18 months of age (Kärtner et al., 2020). However, some studies failed to find an effect of positive emotional responses on future helping behaviour. For example, the effect was not found in 15- to 18-month-olds (Dahl et al., 2017) or for children late in the second year of life (Vaish et al., 2009; Warneken & Tomasello, 2013).

Evidence is also mixed with regard to the motivational impact of negative emotions on helping behaviour. Negative emotional displays, such as sadness and pain, have been found to promote helping behaviour in the second and third years of life (Bandstra et al., 2011; Bischof-Köhler, 1991; Campbell et al., 2015; Hepach et al., 2013; Johnson, 1982; Roth-Hanania et al., 2011; Walle et al., 2017; Zahn-Waxler et al., 1992). There is also longitudinal evidence of this effect, as concern for others' distress at 3 months of age is predictive of children helping distressed others at 18

months of age (Davidov et al., 2020). Some studies find no such effect for children in this age range, however (Newton et al., 2014; Pettygrove et al., 2013), although 18- to 20-month-olds with a more developed understanding of others' emotional states do seem motivated to help others who exhibit sadness (Newton et al., 2014). It is worth noting, however, that in these studies the negative emotions are not directed *at children*. Therefore although these findings do provide some evidence in support of the hypothesis that *others' emotions* motivate helping behaviour, they do not specifically support the view that others' emotions *in the context of social interaction* motivate and reinforce helping behaviour over childhood. This is a limitation of past research that will be addressed shortly (see below). For both positive and negative emotions, then, evidence from lab-based studies is mixed.

Results from ecological data provide clearer support for the hypothesis that others' emotions motivate helping behaviour in early childhood. One study in which parents were asked to complete a survey about their children's helping behaviour indicated that caregiver praise was the highest source of motivation for children's helping behaviour between the ages of 1 and 4 (Hammond & Brownell, 2018). In addition to this, 18- to 24-month-olds help in order to relieve others' distress in naturalistic settings (Johnson, 1982). In a longitudinal study in which both interviews and recordings of naturalistic caregiver-child interactions were taken (Dahl, 2015), helping behaviour at 19 and 24 months of age was positively related to encouragement and reinforcement of helping behaviour at 13-15 months of age (Dahl, 2015). The ecological data therefore supports the hypothesis that helping behaviour is motivated and reinforced by others' emotions, although the only emotions directed at children in the ecological data were positive. In summary, ecological and theoretical research clearly supports the hypothesis that others' emotions motivate and reinforce prosocial behaviour in early childhood. Interpreting the lab-based research is less clear, however.

How should the empirical evidence be synthesised? Despite the mixed results, I would argue that the available evidence supports the predictions generated by the hypothesis that others' emotions motivate and reinforce instrumental helping over childhood. Firstly, evidence from the longitudinal studies discussed above (Dahl, 2015; Davidov et al., 2020; Kärtner et al., 2020) is crucial here because the hypothesis in question is about developmental trajectory. That is, it is a hypothesis about changes in behaviour over the course of childhood, rather than a hypothesis about individual helping events. The reinforcement of behaviours may take time (Aknin et al., 2018; Dickinson, 1980; Godman et al., 2014), and we may therefore not *always* expect to see the results of reinforcement in an individual laboratory session. These studies therefore address the present hypothesis more directly than any other research discussed here.

Secondly, the ecological research carries much weight here. Ecological data are more insightful than data from highly controlled laboratory studies for understanding how helping behaviours emerge in the everyday life of children and how others respond to such behaviours. This data indicates that positive emotional responses from caregivers following children's helping behaviour are common, and that this positively correlates with future helping behaviour (Dahl, 2015; Hammond & Brownell, 2018). Ecological evidence also points towards children being motivated to help others who exhibit negative emotional displays, such as distress (Johnson, 1982). These data therefore speak strongly in favour of the hypothesis that others' emotions motivate and reinforce instrumental helping.

Finally, although some studies fail to find evidence that others' emotions can motivate helping behaviour, the lack of evidence for an effect is not evidence against the existence of the effect (in the context of null-hypothesis significance testing). That is, failing to find evidence that others' emotions promote helping behaviour is not itself evidence against this hypothesis. Given this, and that much lab-based research *did* find evidence supporting this hypothesis, I would conclude that past research provides

some support for the hypothesis that the emotional responses of others motivate and reinforce helping behaviour in early childhood.

More research that directly addresses the hypothesis that others' emotions motivate and reinforce helping behaviour in early childhood would be highly beneficial. In particular, the collection of more ecological and longitudinal data would serve several important functions. Firstly, it would ensure that the results of the longitudinal studies discussed above (Dahl, 2015; Davidov et al., 2020; Kärtner et al., 2020) replicate.

A second benefit of further longitudinal research is that it would enable an investigation of the role of both negative and positive emotions in the same study (in contrast to most studies discussed above, which typically focus on either positive or negative emotion). For example, a naturalistic longitudinal study could be conducted in which both positive and negative reinforcement of helping behaviour were monitored. In such a study, both positive reinforcement of helping behaviours (e.g., praise) and negative reinforcement of failures to help (e.g., scolding or disappointment directed at the child when they do not help despite being in a position to do so) could be monitored. Repeated recordings over different stages of childhood (say, starting at 12 months of age and occurring every 6 months) would provide greater insight into the relative roles of positive and negative emotions in the development of prosocial motivation and the degree to which different kinds of emotions might promote helping behaviour.

One limitation of the ecological and lab-based research discussed above which future research could address is that there is no investigation of negative emotions that are specifically directed *at* children. For example, these studies have focused on others' sadness as motivating helping behaviour rather than, say, children being scolded for not helping when in a position to do so. This is a limitation that would be relatively straightforward to address in ecological research, as one could capture negative emotional responses directed at children by simply recording child-caregiver interactions. Indeed, there *are* ecological studies that have investigated negative emotional responses directed at children (Berlin et al.,

2009; Senehi & Brophy-Herb, 2020), although they have not focused on the effects of this on *helping* behaviour. Future research could therefore address this limitation of ecological data relatively straightforwardly. Addressing this limitation in lab-based research may be somewhat more challenging, however.

The reason for which no lab-based studies have focused on the effects of negative emotions directed *at* children is likely ethical. It would not be ethical to unnecessarily direct negative emotions at toddlers (e.g., by scolding children who fail to help others when they are in a position to do so), especially given the potential negative effects of this on children's future behaviour (Chang et al., 2003). However, it may be possible to address such ethical concerns by designing a study in such a way that the negative emotions directed at children who are in a position to help but do not do so are fairly mild. For example, after children help an experimenter the experimenter could say "I am happy you helped me" followed by a mild smile, and following a failure to help the experimenter could say "Oh, I wish you had helped me" followed by a mild frown. In this way, one could investigate the role of negative emotional responses directed at children in reinforcing helping behaviour, while overcoming ethical concerns relating to harsh negative emotions (pending informed parental consent and approval from the relevant ethics committees, of course).

One additional direction for future research would be to investigate the degree to which past findings extend to cross-cultural contexts. This is important because most studies discussed here are from a small number of countries (e.g., the United Kingdom, Germany, and the United States of America), and therefore the degree to which the findings discussed here extend to further cultural contexts is not clear. I wish to set aside this particular consideration until section 5.4, as this limitation also applies to research on executive function (i.e., the research to be discussed in section 5.3).

5.2.3 Changes in Expectation

I have thus far set aside a particular characteristic of the ecological data. This is that some studies find that praise and encouragement of helping behaviour begins to decrease from late in the second year of life (Dahl, 2015; Dahl & Brownell, 2018). How can this decrease in praise and encouragement as children age be reconciled with the hypothesis that others' emotions reinforce helping behaviour in early childhood?

This characteristic can be accounted for by changes in the *type* of social interaction that children encounter over the course of childhood. Specifically, as children age, expectations about their behaviour are likely to change (Gralinski & Kopp, 1993). When children are younger (say, up until 24-30 months of age), caregivers may not expect children to contribute to the goals of others without being prompted to. The praise directed at children after they have helped another agent is therefore likely to be sincere and earnest - if you are not expecting help from an 18-month-old and you receive it, it is a genuinely pleasant surprise. Lower expectations mean that helping behaviour will elicit stronger and more sincere positive emotional responses. The sincerity of this feedback is likely to be especially influential because it may surprise children at this age, and reinforcement learning is most effective when outcomes are surprisingly different from what one initially expected (Delin & Baumeister, 1994). Given the relatively low expectations about children's helping behaviour when they are younger, we would also expect a relatively small amount (if any) of negative emotions directed towards them at this stage when they fail to help others.

Caregivers come to expect more of children with regard to helping behaviour as they age, however (Dunn et al., 2014; Hofferth & Sandberg, 2001; Rogoff, 2003; Waugh et al., 2015). Specifically, they may expect children to help without being prompted, and thus caregivers may begin to provide less child-directed positive reinforcement for such behaviours (Dahl, 2015; Hammond & Brownell, 2018). Indeed, praise is most effective at reinforcing a behaviour if it is sincere and surprising, and praise may be detrimental to the reinforcement of a behaviour if that behaviour is expected

(Henderlong & Lepper, 2002). This means that undue praise (e.g., praise for performing actions that one knows that one is expected to perform) may feel patronising, and can therefore undermine one's motivation to perform the praised behaviour. Changes in expectations about children's helping behaviour may explain why praise and encouragement of such behaviours decrease as children age (Dahl, 2015). This also generates the novel prediction that positive emotional reinforcement will be more *effective* in promoting helping behaviour earlier in childhood, while negative emotional reinforcement may become more effective in motivating children to help others as they age.

A further factor that is likely to change over the course of development is *where* the emotions of others are directed during helping scenarios. Emotional responses may not always be directed at the child. Instead, they may be directed at the goal objects or location, especially in the case of negative emotions such as frustration. For example, if I'm struggling to open a door in the presence of a toddler who could help me and I get frustrated, I am more likely to direct my negative emotion at the door than at the toddler. In contrast, positive emotions that result from a goal being completed due to the contribution of an external agent will likely (at least in part) be directed towards the helper. As children grow older and more is expected of them, however, negative emotions may be more likely to be directed towards them. For example, a caregiver may be more likely to reprimand a 3-year-old for not helping when they are in a position to do so than a 2-year-old. This generates the prediction that positive reinforcement will be more *common* with children in the younger age groups (e.g., pre-30 months of age), while negative reinforcement will become more prominent as children grow older. That positive reinforcement decreases after the third year of life is supported by ecological data (Dahl, 2015), but this has yet to be investigated for negative reinforcement.

Much of what has been said so far may depend on individual and cultural differences. That is, the particular developmental trajectory that I have described here may not apply to all cultural or household contexts. For

instance, there may be some households in which helping behaviour is not expected from children until significantly later in life (say, from the age of 6 years). This important point will also apply to how emotions may reinforce and motivate executive function use, and I will therefore address this point towards the end of this chapter (section 5.4).

5.2.4 Conclusion

I have evaluated evidence pertaining to the hypothesis that social interaction scaffolds the development of prosocial motivation in early childhood. Specifically, I have suggested that social interaction alters the immediate rewards of behaving prosocially because of the emotional responses of others to helping behaviour (or a lack thereof). Others' emotions can motivate helping behaviour by raising the rewards of helping and raising the costs of failing to help others (as compared to a situation in which others are not exhibiting emotional responses in relation to their goal status or others' behaviour), reinforcing helping behaviour over childhood.

Having evaluated the empirical evidence that pertains to this hypothesis, there does seem to be some support for it, although there are notable limitations of the data. There is evidence that positive emotions directed at children following instances of helping behaviour reinforce that behaviour over childhood. There is also evidence that others' negative emotions can promote helping behaviour, although whether negative emotions *directed at children* reinforce helping behaviour is yet to be investigated. For positive emotions, there is therefore support for the hypothesis that social interaction scaffolds the development of helping behaviour over early childhood, although there is a lack of evidence that directly bears upon this hypothesis for negative emotions.

What is unique about social interaction here (in comparison to children merely observing another agent's emotional responses to failing or succeeding to achieve a goal) is that affective engagement is heightened during interaction between agents - the link between affect, goal status, and one's own behaviour is more salient from a second-person perspective than from the perspective of an isolated observer (Moore & Baressi, 2017).

Social experiences are key in scaffolding children's understanding of the world, and specifically their understanding of the role of their own behaviour in determining others' emotions (Thompson & Newton, 2013). While children can identify others' emotions from the position of an observer, the increased salience of these emotions in the context of social interaction means that these emotions will better reinforce and motivate helping behaviour in this context. This generates the prediction that positive and negative emotions directed at children will likely be more effective at motivating helping behaviour that emotions directed instead at goal objects or locations.

More research (especially ecological and longitudinal research) would greatly improve our understanding of how emotions influence helping behaviour over the course of early childhood. In addition, I have only considered prosocial motivation in the context of one type of prosocial behaviour - instrumental helping. It would also be beneficial in future to consider prosociality more broadly. Having discussed how social interaction can scaffold the development of one component of interpersonal commitment over the course of childhood (prosocial motivation), I will now turn to the development of the cognitive component of commitment.

5.3 Executive Function Scaffolded by Social Interaction

Executive function plays an important role in enabling agents to honour interpersonal commitments. For instance, inhibitory control ensures that agents are able to resist tempting actions that conflict with those actions that they have committed to. If social interaction scaffolds the development of executive function in early childhood, then social interaction scaffolds interpersonal commitment.

There are a number of ways in which the behaviours of other agents can scaffold the development of executive function (Vygotsky, 1978). Caregivers can scaffold children's use of executive function processes by highlighting when and where children can employ executive function to greatest effect (Broomell et al., 2020; Hughes & Devine, 2019; Mermelshtine, 2017; Neale & Whitebread, 2019), and by modelling patterns

of behaviour (such as persistence) that require executive function (Leonard et al, 2017; Lucca et al., 2019), thereby demonstrating the benefits of employing executive function.

Despite a recent surge in research focusing on the role of the second-person position in facilitating executive function development (Fay-Stammbach et al., 2014; Lewis & Carpendale, 2009; Moriguchi, 2014; Perry et al., 2019; van Lier & Deater-Deckard, 2016), the role of others' *emotions* in facilitating executive function development has yet to be directly investigated. This is worth investigating because children are highly sensitive to others' emotions from early in life (Ichikawa & Yamaguchi, 2014; Nelson et al., 1979; Nelson & Dolgin, 1985; Serrano et al., 1992).

The specific hypothesis that I will spell out and evaluate evidence for is that social interaction alters the immediate rewards of employing executive function due to others' emotions. Specifically, I will suggest that the second-person position increases the rewards of employing executive function while raising the costs of failing to employ executive function, because of the emotions of agents with whom children are interacting in joint activities. As with prosocial motivation, it may be that others' emotions can motivate or reinforce executive function use even if children are in the position of an observer rather than in the context of social interaction. However, it is in the context of social interaction that others' emotions (and the link between these and one's own action or inaction) will be most salient, and therefore it is in this context that others' emotions will be strongest at motivating and reinforcing executive function use. This change in the rewards of employing executive function reinforces executive function use over the course of childhood.

5.3.1 Others' Emotions as Motivating Executive Function Use

For others' emotions to motivate a given set of behaviours, it is necessary that (i) children are able to identify others' emotions, (ii) children can form an association between an agent's emotions and the events or objects to which they are directed, and (iii) children find others' emotions pleasant or aversive. We have already established that children satisfy these

criteria from the end of the first year of life (see section 5.2.1). The focus here will therefore be evidence pertaining to whether others' emotions motivate executive function use in early childhood.

In the context of executive function, my hypothesis is that others' emotions may motivate children to employ inhibitory control (in order to persist for longer at a task) or attentional control (in order to pay greater attention to task-relevant information), among other executive functions. For example, suppose that a child is playing a game. If they are playing the game by themselves and they lose interest in the game (deciding that they would rather play a different game), there is little reason for them to resist this impulse to abandon the game. That is, there are no negative consequences of failing to inhibit the impulse to switch game. However, if the child is playing this game with another agent, then abandoning the game before its completion may elicit a negative emotional response from their partner (e.g., disappointment). This negative emotional response would raise the costs of abandoning the game, thereby encouraging the child to employ inhibitory control. In this way, others' emotions may motivate executive function use.

The role of others' emotions in motivating executive function use is particularly important in development because others' emotional responses will change over the course of development. Specifically, as children age, there will be changes in what others expect from children's behaviours, as well as changes in who children interact with.

Firstly, as children age, more will be expected of them with regards to executive function (Geeraerts et al., 2021; Laible & Thompson, 2007). Social interaction in very early childhood (e.g., until the age of 24 months) will often be child-led, insofar as caregivers may often adapt to whatever game the child wants to play. If the caregiver and child are playing a particular game and then the child loses interest and moves on to playing a different game, caregivers may be unlikely to reprimand the child. This is because at this age, the focus of the caregiver is often simply on interacting with the child. This means that at younger ages, a failure to employ

executive function is less likely to evoke a negative emotional response, while successful executive function use is likely to elicit a positive emotional response. As children age, however, more is expected of them, especially as they transition to school (Geeraerts et al., 2021). For instance, they may be expected to resist acting selfishly. Children are exposed to these changing expectations in their interactions with adults, raising their awareness of what is required of them in terms of self-control (Laible & Thompson, 2007). In this sense, as children age, a failure to employ executive function in joint tasks will be more likely to elicit negative emotional responses from their partners.

In addition to changes in expectations about children's executive function use, there may also be changes in whom children interact with over the course of childhood. At a younger age (say, up until 24 months), children are likely to mostly be interacting with members of their own household. While there may be cases in which children are interacting with agents outside of their household (e.g., there may be playdates with children from other households), it is primarily interactions within the household that children will be enaging in. For the most part, within their household younger children will be interacting with their primary caregiver(s). Caregivers are likely to be tolerant of children's behaviour, and therefore not likely to exhibit negative emotional responses when children fail to successfully employ executive function. That is, at younger ages children's failure to employ executive function is unlikely to elicit negative emotional responses due to who children interact with at younger ages. However, as children start to attend nursery or pre-school, they will start interacting with a wider variety of other agents, especially children, more frequently. Interacting with peers (and being faced with expectations about how one should behave in school and nursery) will contribute to the development of executive function use in early childhood because other children are likely to be less tolerant of impulsive behaviours than caregivers (Geeraerts et al., 2021; Holmes et al., 2016), and therefore peers are more likely to exhibit negative emotions if children fail to employ executive function. Together,

these changes mean that others' emotional responses to children's performance in joint tasks will become more prevalent over the course of childhood.

An important caveat regarding who children interact with, and how this changes over childhood, revolves around individual differences in households. Specifically, the presence of siblings or other young children in one's household (e.g., from one's extended family) are relevant here. The presence of siblings or other young children may influence the way in which others' emotions can promote executive function use, depending on the amount of time that siblings interact with each other within a household and the manner in which siblings interact with one another. For instance, siblings' tolerance for children's behaviour will likely depend on the age of the siblings. We might expect older siblings at, say, 15 years old, to be more understanding of a 2-year-old's executive function limitations than, say, a 3-year-old sibling (Cutting & Dunn, 2006; Recchia & Howe, 2009). We would therefore expect children with siblings of a more similar age to encounter negative emotions from earlier as a result of failures to employ executive function, although this would of course also depend on the quality of interactions with older siblings (e.g., uninterested or uncaring older siblings may also not be as sympathetic or understanding). These effects would of course be mediated by the amount of time that children interact with siblings as compared to their primary caregivers; in some households siblings may be encouraged, or at least able, to interact with each other often, whereas in other households it may be much less common for siblings to engage in extensive social interaction independently of their primary caregivers. Given such differences between households it is difficult to identify the precise effects that sibling presence would have on executive function development, although this is important to be aware of when considering the relation between children's executive function development and who children interact with in early childhood.

5.3.2 Predictions and Evidence

In this section I will evaluate the hypothesis that social interaction scaffolds the development of executive function because others' emotions motivate executive function use in early childhood. This hypothesis is yet to be directly tested, although it has been suggested in previous theoretical research (Pessoa, 2009). Although some past studies have included others' emotions in their analysis of executive function development, it is typically grouped with other factors (e.g., caregiver intrusiveness and caregiver energy; Blair et al., 2011; Fay-Stammbach, 2014; Rhoades et al., 2011). That is, affective interaction is not typically treated separately for the purposes of statistical analysis, and its effect on executive function use has therefore not been directly investigated. For this reason, I will first identify predictions that the present hypothesis generates and evaluate the indirect evidence that bears upon these predictions. I will then return to consider how this hypothesis might best be directly tested.

One prediction that the present hypothesis generates is that children will perform better at tasks requiring executive function if they have previously engaged in social interaction with other agents who are more emotionally expressive, as compared to relatively less expressive agents. This is because engaging with other agents who are relatively more emotionally expressive promotes emotion understanding in early childhood (Ogren et al., 2018; Ogren & Johnson, 2021), and therefore children in such scenarios will develop a better understanding of the effects of their own behaviour on others' emotions. Children engaging with relatively more expressive others will thus develop a better understanding of the consequences of failing to employ executive function successfully.

One way in which this prediction could be directly tested would be to compare the performance on executive function tasks of children whose caregivers are more or less emotionally expressive than typical individuals. For example, individuals who have been diagnosed with depression are often less emotionally expressive than other individuals (Kahn & Garrison, 2009). This generates the testable prediction that performance on executive

function tasks may be lower for children whose caregivers have been diagnosed with depression, in comparison to children whose caregivers have not been diagnosed with depression.

The prediction that children will perform better at tasks requiring executive function if they have previously engaged in social interaction with more emotionally expressive partners is yet to be directly tested. However, there is evidence that others' emotional responses to children's performance in executive function tasks can motivate and reinforce executive function use in early childhood. Specifically, parental praise of effort and persistence (i.e., a positive emotion) is predictive of children's persistence in challenging tasks at 18 months of age (Lucca et al., 2019). This effect even carries over to future tasks in which the parent is no longer present, and is thus no longer in a position to encourage task persistence (Lucca et al., 2019). This study demonstrates that others' positive emotions motivate and reinforce executive function use in early childhood. Although this study provides some indirect support for the prediction that others' emotional expressions motivate executive function use in early childhood, it should be noted that this study does not investigate the motivational effects of negative emotions (e.g., disappointment due to children's failure to successfully employ executive function), and it only addresses inhibitory control.

A further prediction that is generated by the hypothesis that others' emotions motivate and reinforce executive function use over the course of childhood is that children will become better at executive function tasks when they start to interact with agents who are less tolerant than their caregivers. That is, we should expect performance in executive function tasks to improve when children start to interact with peers and non-parental caregivers, such as teachers (e.g., when children start to attend nursery or preschool), and we may expect children with siblings to perform better at executive function tasks than children without siblings. This is because caregivers are perhaps likely to be more tolerant of children's failure to successfully employ executive function in a joint task than peers or siblings

(e.g., caregivers might be less likely to reprimand a child for losing interest in a collaborative game than other children might be). Children will therefore be exposed to more negative emotional responses following failures in executive function use once they start interacting with non-caregivers.

This prediction is consistent with the finding that performance on executive function tasks typically improves over the course of childhood and especially over the years at which children begin to attend school, pre-school, and nursery (e.g., Best & Miller, 2010; Carlson, 2005; Hughes, 2011; van Lier & Deater-Decker, 2016). However, it should be noted that these studies do not specifically support the hypothesis that this improvement in executive function is a result of children encountering others' emotions in the context of social interaction. More broadly, this prediction is difficult to directly test because there are many factors that change around the age at which children start to attend nursery and pre-school (e.g., physical and neural developments) which may also explain why performance in executive function tasks improves at this age (Hughes, 2011).

Evidence is mixed with regard to the effects of sibling presence on executive function development. One longitudinal study found that having siblings (compared to not having siblings) was predictive of improved performance on executive function tasks over a 12 month period for 3- to 6-year-olds, although this effect may be mediated by improvements in theory of mind that arise from growing up with siblings (McAlister & Peterson, 2013). In contrast, a separate cross-sectional study found that the presence of siblings was negatively associated with executive function for 30- to 60-month-olds, although this was found to be mediated by changes in parental behaviour that arise from having multiple children (Rolan et al., 2018). There is therefore mixed evidence regarding the prediction that sibling presence would contribute to executive function development, and it is again worth noting that this evidence does not specifically focus on the role of social interaction.

The hypothesis that social interaction scaffolds the development of executive function because of others' emotions generates a number of predictions. Unfortunately, this hypothesis is yet to be directly tested. This hypothesis could be tested with a longitudinal study conducted over the course of early childhood, in which children engage in a variety of executive function tasks at regular intervals (say, every 3 or 6 months). For an ecological study the emotional responses of agents that children interact with (whether they be caregivers, siblings, or other agents) could be recorded following children's successes and failures at executive function tasks. The relation of these emotional responses to children's future performances on executive function tasks could then be identified.

Translating the above suggested ecological study into a lab-based study may be more challenging. This is due to the ethical concerns around experimental manipulations that involve subjecting young children to negative emotional responses. However, as with prosocial motivation, it may be possible to run such a study provided that the emotional responses were relatively mild (see section 5.2.2). For example, suppose that an experimenter and a child engage in a joint task. If children succeed in persisting with the task to its completion then the experimenter could respond positively (e.g., "I'm happy that you finished it" while smiling), and otherwise the experimenter could respond less positively (e.g., "I wish you had finished it" while frowning mildly). If this were combined with measurements of children's performance in executive function tasks over the course of childhood (say, every 3 or 6 months), then the relation between positive and negative emotional responses to children's successes or failures to employ executive function could be identified.

5.3.3 Conclusion

I have articulated a version of the hypothesis that social interaction scaffolds the development of executive function in early childhood. Specifically, I have evaluated evidence pertaining to the hypothesis that social interaction alters the immediate rewards of employing executive function because of the emotional responses of others. This hypothesis

builds upon recent work that underlines the importance of social interaction in understanding how executive function develops in early childhood (Carlson, 2009; Lewis & Carpendale, 2009). While this hypothesis has yet to be directly tested, there is some evidence which supports the predictions that this hypothesis generates. Nevertheless, this is mostly indirect evidence, and this hypothesis would benefit from being directly tested in longitudinal studies.

5.4 General Discussion

I adopted a two-step theoretical approach to investigate the development of children's ability to honour an interpersonal commitment. First, I established a new theoretical framework for addressing questions about interpersonal commitment. By reducing the ability to honour an interpersonal commitment into its constitutive components, questions about the development of one's ability to honour an interpersonal commitment can be rephrased into questions about the development of these components. The second step was to adopt a second-person approach, investigating the role of social interaction in the development of interpersonal commitment over childhood.

I articulated the hypothesis that social interaction scaffolds the development of interpersonal commitment over childhood by introducing immediate rewards for employing executive function and instrumentally helping others. These rewards are the emotional responses of other agents, and reinforce instrumental helping and executive function use over childhood. This hypothesis is appealing because others' emotions play a central role in how young children come to understand the world around them; children are sensitive to others' emotions from early in the first year of life and others' emotions can motivate children to act in ways that they would not otherwise. In addition, it is in the context of social interaction that others' emotions (and the relation between these and one's own actions) are most salient, and therefore it is in this context that others' emotions will be strongest in reinforcing behaviour.

This hypothesis is yet to be directly tested. However, there does seem to be some support for this hypothesis insofar as empirical evidence is consistent with the predictions that this hypothesis generates. There are notable limitations of past research, however (see section 5.4.1).

This chapter extends past research investigating commitment in early childhood in three ways. Past research on interpersonal commitment in early childhood has primarily focused on differences between age groups (e.g., Gräfenhain et al., 2013; Hamann et al., 2012; Kachel et al., 2017). In contrast, the present chapter directly addresses developmental changes over the course of childhood, focusing on what changes and how children become better at honouring interpersonal commitments.

Secondly, this chapter introduces a novel reductive theoretical framework for investigating the development of children's ability to honour interpersonal commitments. This framework offers a means of answering questions about commitment by rephrasing them in terms of the components of interpersonal commitment, thereby generating novel hypotheses. For instance, while past empirical studies on children's ability to honour an interpersonal commitment have typically explained differences between age groups by appealing to the development of children's understanding of the normative implications of forming an interpersonal commitment (Gräfenhain et al., 2009, 2013; Kachel & Tomasello, 2019; Kanngiesser et al., 2017), the theoretical framework presented in this chapter raises the possibility that such differences may, at least partially, be driven by the development of children's cognitive abilities. The suggestion here is not to say that the normative account is wrong, but rather it is simply to draw attention to the fact that there are multiple factors that change over early childhood that may affect one's ability to honour interpersonal commitments. Executive function development is perhaps one of the more obvious factors that is important for honouring a commitment; indeed, in everyday life, when we describe others as being notably committed to a task, we are not typically referring to their adherence to norms but rather to their perseverance or focus (i.e., inhibitory and attentional control). For

example, if two agents agree to go to the gym together on a weekly basis, describing these agents as being highly committed would be most naturally interpreted as meaning that these agents are highly focused and persistent in going to the gym together, rather than to say that they are highly motivated to adhere to norms (Michael, Knoblich, & Sebanz, 2016). The precise degree to which different factors contribute to the emergence of interpersonal commitment is a topic for future research. In order to establish the degree to which these past findings are indeed driven by normativitiy as opposed to the development of executive function, it would be important to replicate these past studies while controlling for executive function abilities (e.g., by recording children's performance on executive function tasks and controlling for this performance in subsequent analysis; see Doebel, 2019; Nigg, 2017).

One important contribution comes from the present chapter explicitly highlighting and considering the importance of the second-person position in the development of interpersonal commitment (see also Siposova, 2018). Employing this second-person approach has uncovered a possible novel relation between individual and interpersonal commitment - that commitment in the individual case may be parasitic on interpersonal commitment. If social interaction scaffolds the development of children's ability to honour an interpersonal commitment, then children may bring what they have learned from the interpersonal case (e.g., to inhibit their impulsive actions) and use it to improve at commitment to individual tasks. While much socio-cognitive research has progressed by first focusing on the individual case and then moving to the joint case (Carlson, 2005; Cohen et al., 2007; Miyake et al., 2000; Wellman & Liu, 2004; Woodward, 2009), the present hypothesis motivates the opposite approach to understanding the development of one's ability to honour a commitment.

This chapter also contributes to the extensive research on the development of executive function and prosocial motivation in early childhood. While there has been extensive past research on these topics (Dahl & Brownell, 2019; Doebel, 2019; Köster et al., 2019; Mermelshtine,

2017; Nigg, 2017), this chapter articulates and evaluates evidence for the hypothesis that these are motivated and reinforced over childhood because of others' emotions, as well as suggesting ways in which this hypothesis could be tested in future.

5.4.1 Future Research

The most important direction for future research would be to directly test the hypothesis that social interaction scaffolds the development of interpersonal commitment because others' emotions motivate and reinforce instrumental helping and executive function. Much of the evidence that pertains to this hypothesis is either indirect or not exhaustive, insofar as individual studies rarely focus on both positive and negative emotions (Dahl et al., 2017; Warneken & Tomsaello, 2013), and studies often focus on a narrow age range (Campbell et al., 2015; Koomen et al., 2020). The ideal way to test this hypothesis is with longitudinal studies that record (if the study were ecological) or manipulate (if the study were lab-based) the emotions of the child's partner in joint activities (although, as previously stated, such emotions would need to be mild in order to address ethical concerns around directing negative emotions at children). This approach would enable one to investigate the role of others' emotions in reinforcing executive function use (e.g., inhibitory control) as well as prosocial motivation (e.g., instrumental helping).

Directly testing this hypothesis would enable two further directions for research, one of which is phylogenetic and the other is ontogenetic. If others' emotions are key in facilitating the development of interpersonal commitment, then it follows that the development of interpersonal commitment may be hindered in individuals for whom others' emotions are less effective at motivating behaviour, or for individuals that struggle to interpret others' emotions (and therefore struggle to form an association between their own behaviours and others' emotions).

This line of reasoning generates the testable hypothesis that social interaction will be less effective in scaffolding the development of interpersonal commitment for certain atypically developing humans, such as

people with autism spectrum disorder. Specifically, we might expect that such individuals are more likely to struggle to honour interpersonal commitments or that such individuals will take longer to develop the ability to honour interpersonal commitments. While commitment in children with autism spectrum disorder is yet to be empirically investigated, there is research suggesting that the development of prosocial motivation (Campbell et al., 2015; Hepach, Hedley, & Nuske, 2019; although see Liebal et al., 2008) and executive function (Christ et al., 2007; Hughes, 2011) is hindered for such individuals. The hypothesis at the heart of this chapter suggests that this may (partly) be due to these individualsstruggling to interpret others' emotions or being less motivated by them. For such individuals, this means that social interaction would not facilitate the development of these components of interpersonal commitment to the degree that it does in typically developing children.

The hypothesis that social interaction scaffolds the development of interpersonal commitment because of others' emotions also raises the question as to whether this is a uniquely human way in which executive function and instrumental helping can develop. Testing the hypotheses presented in this chapter with non-human animals may be useful in identifying the degree to which interpersonal commitments may be distinctly human, and may point towards one of the ways in which humans evolved to develop uniquely complex social structures (Tomasello, 2014). Although past research has demonstrated that reinforcement learning is present in other species (Dickinson, 1980), there is a lack of research as to whether social interaction (and specifically, others' emotions) facilitates the development of commitment by motivating and reinforcing prosocial motivation and executive function in non-human animals.

Although much phylogenetic research compares humans to non-human primates, the most appropriate species for testing this hypothesis may be dogs. This is because dogs have an understanding of human emotions (Correia-Caeiro et al., 2020; Müller et al., 2015; Szánthó et al., 2017), and human emotional responses seem to be sufficient motivation for

reinforcement learning, especially in the context of prosocial behaviour and executive function (Blackwell et al., 2008; Cook et al., 2016; Rooney & Cowan, 2011). Indeed, in everyday life, emotional responses (e.g., praise or disappointment) are often used to train dogs (e.g., to resist the temptation of eating food that has fallen onto the floor). Dogs may therefore be the most likely non-human animals to perform well in interpersonal commitment tasks.

Future research could also focus on addressing the limitations of past research. These primarily relate to differences in cultural contexts and individual differences of households. The research that has been discussed in this chapter has almost exclusively come from WEIRD (Western, educated, industrial, rich, and democratic) countries (Muthukrishna et al., 2020), meaning that the findings may not extend beyond this context. That is, there may be substantial cultural differences in how emotions are expressed and communicated (Fischer et al., 2004; Matsumoto & Hwang, 2019; Sorenson, 1975), as well as differences in how and with whom children interact with in early childhood (Callaghan & Corbit, 2018; Carlson & Harwood, 2003; Kärtner, 2018; Kärtner et al., 2010; Posada et al., 2002). This is problematic because it substantially limits the generalisability of the accounts put forward on the basis of past research. One essential direction for future research would be to simply aim to replicate the studies discussed in this chapter in broader cultural contexts.

A related concern focuses on individual differences between households. While I have suggested that young children may primarily engage in social interaction with caregivers as compared to siblings or members of their extended family, this may differ between households. For instance, children may spend substantially less time with adult caregivers in households in which all adult caregivers are required to work full-time jobs, and therefore such children are likely to spend relatively more time with some combination of other non-caregiver adults and/or siblings. This will impact the quantity and quality of social interaction that such children engage in, which will carry implications for the scaffolding effects of social

interaction on the development of prosocial motivation and executive function. This raises questions about the degree to which this chapter's hypothesis will generalise. Future theoretical research must also therefore take into consideration individual differences between households (e.g., family size and socioeconomic status), and future empirical research would benefit from controlling for such factors in their design and analysis.

5.4.2 Conclusion

The core message of this chapter is that social interaction is foundational in the development of one's ability to honour an interpersonal commitment because of others' emotions. Specifically, others' emotions may act as immediate rewards that reinforce prosocial motivation and executive function. In this way, social interaction may scaffold the development of interpersonal commitment over childhood.

This chapter contributes to this thesis by addressing the development of interpersonal commitment over the course of childhood. In doing so, this chapter has generated a novel hypothesis about the relation between individual and interpersonal commitments - that one's ability to honour an individual commitment may, developmentally, be built upon one's ability to honour interpersonal commitments.

Chapter 6
Conclusion

Goal tracking and prosocial behaviours are common features of the day-to-day lives of humans. These aspects of (pro)social cognition complement each other insofar as goal tracking underpins much of our social cognition and prosocial behaviour. This is especially true in early childhood, as goal tracking (at least in part) determines which prosocial behaviours children engage in (Hepach et al., 2020; Jensen, 2016; Knudsen & Liszkowski, 2012, 2013; Martin & Olson, 2013; Paulus, 2019) as well as providing insight into the cognitive mechanisms that underpin prosociality (Bridgers & Gweon, 2018; Hobbs & Spelke, 2015; Kachel et al., 2017; Köster et al., 2019). This, in turn, generates hypotheses about prosociality in adulthood and the evolutionary distinctiveness of human (pro)social cognition relative to other non-human animals.

This thesis focused on one particular aspect of goal tracking, goal status, and its impact on prosociality. More specifically, this thesis primarily investigated the link between children's understanding of goal status and one of the earliest forms of prosociality, instrumental helping, in early childhood (Chapters 2 and 3). In addition, this thesis investigated the role of goal status in determining goal salience (Chapter 4). Finally, this thesis explored the developmental trajectory of a further kind of prosocial behaviour that involves both goal tracking and instrumental helping, namely, interpersonal commitments (Chapter 5).

The main findings of this thesis are summarised and discussed in this final chapter. Section 6.1 addresses the four research questions identified in this thesis' introduction. Section 6.2 integrates these findings with past research on goal tracking and prosociality, spelling out some of the theoretical and practical implications of this research. The limitations of the present thesis' methods and directions for future research are discussed in section 6.3. This chapter closes with a summary of the main conclusions in section 6.4.

6.1 Summary and Answers to Research Questions

RQ 1: Do children understand goal abandonment? More specifically, are young children sensitive to the distinction between abandoned goals and goals of other types of status, e.g., interrupted goals?

While past research investigated children's understanding of different types of goal status, this primarily focused on children's understanding of the distinction between completed and interrupted goals (Behne et al, 2005; Brandone et al., 2014; Carpenter et al., 1998; Kachel et al., 2017; Skerry & Spelke, 2014; Warneken et al., 2012). Children's understanding of a further type of goal, abandoned goals, had yet to be investigated.

It is important to identify whether children understand goal abandonment because accurate goal tracking requires differentiating between the different reasons for which an agent may have not completed their directed action, and goal abandonment is a feature of everyday life (e.g., one might adopt a goal before later deciding that one no longer wants to invest the effort required to achieve the goal). In addition, further insight into children's understanding of goal status can shed light on the cognitive mechanisms that underpin early prosocial behaviour (see Chapter 3).

We hypothesised that if children understood goal abandonment, then their helping behaviour should be sensitive to the distinction between abandoned and interrupted goals from the age of 2. This is because children are willing and able to instrumentally help others achieve their goals from the second year of life (Hepach et al., 2017; Kenward & Gredebäck, 2013; Martin et al., 2020; Svetlova et al., 2010; Warneken et al., 2007; Warneken & Tomasello, 2006), though they may struggle to help appropriately if there are multiple helping affordances before this age (Hobbs & Spelke, 2015; Krogh-Jespersen et al., 2015; Waugh & Brownell, 2017). To test this hypothesis, we designed an instrumental helping study in which an experimenter started to place a toy into one of two boxes before halting his directed action. We manipulated the reason for which the experimenter halted his action: he either encountered an obstacle that prevented him from

reaching his preferred box (interrupted condition), or else he abandoned his original goal location in favour of the alternative box (abandoned condition). The experimenter then asked for help, and we measured whether or not children placed the toy in the experimenter's initial goal location in each test trial.

The results indicated that children were more likely to place the toy into E's initial goal location when E's initial goal was interrupted as compared to when E's goal was abandoned. Follow-up analysis indicated that children were helping correctly (i.e., placing the toy into the box that the experimenter wanted the toy to go by the time he asked for help) above chance in both conditions, and there was no evidence that children were more likely to help correctly in one condition compared to the other.

Thus, the experiment presented in Chapter 2 is the first to present some preliminary evidence in support of the hypothesis that instrumental helping behaviour is sensitive to the distinction between abandoned and interrupted goals from the age of 2, and therefore that children understand goal abandonment from this age. The limitations of this study are important to bear in mind, however. Specifically, as discussed in section 2.4, in order to provide stronger evidence in support of this hypothesis, further research is required to address the alternative explanation for these findings (that children are only responding to E's most recent goal-cues at the point that E asks for help).

RQ 2: Is goal slippage one of the cognitive mechanisms that underpins instrumental helping in early childhood?

It is important to investigate the underlying mechanisms for instrumental helping because instrumental helping is one of the earliest emerging forms of prosocial behaviour in childhood (Dunfield, 2014; Paulus, 2014; Thompson & Newton, 2013). The cognitive underpinnings of this behaviour may therefore shed light on the ontogenetic origins of further forms of prosociality.

A core motivation for investigating children's understanding of goal abandonment in an instrumental helping context was to shed light on the

cognitive mechanisms that underpin instrumental helping. More specifically, identifying whether children understand goal abandonment is essential for testing a hypothesised cognitive mechanism for instrumental helping, that is goal slippage. This is because goal slippage generates unique predictions about children's helping behaviour (in comparison to alternative mechanisms, such as psychological altruism) if the helpee has abandoned their goal (Michael, Knoblich, & Sebanz, 2016; Michael & Székely, 2019). Specifically, the goal slippage hypothesis predicts that children will contribute to another agent's goal even after the agent has lost interest and abandoned that goal.

To test the goal slippage hypothesis we designed an instrumental helping study for 2-year-olds (as this was the age at which children had demonstrated an understanding of goal abandonment). On test trials the experimenter placed two toys into one of three possible goal locations, before disengaging from the test apparatus. In the control condition the experimenter had achieved his goal before disengaging from the apparatus, i.e., placing two toys into the goal location satisfied his goal. In the experimental condition, however, disengaging from the apparatus after placing only two toys in the goal location constituted goal abandonment (i.e., the goal was incomplete). In both conditions it was feasible for the participant to place a final toy in the experimenter's goal location. We measured whether children placed the final toy in the experimenter's goal location or in one of the alternative locations.

The results indicated that children were significantly more likely to place the final toy in the experimenter's goal location in the experimental condition (i.e., when it constituted completing the experimenter's abandoned goal) than in the control condition (i.e, when the experimenter's goal was already completed), and that children were placing the toy in E's goal location above chance in both conditions. This result provides some preliminary support for the hypothesis that the goal slippage mechanism underpins at least some instances of instrumental helping in early childhood. As with the study presented in Chapter 2, however, it is important to be

aware of the limitations of this study. Specifically, the differences in stimuli material between the two conditions raise the possibility of alternative explanations for the results (see section 3.4). Future research would be needed to address these concerns and verify the generalisability of these results (as discussed in section 3.4).

RQ 3: What is the relation between goal status and goal salience? Specifically, how salient are abandoned goals in relation to interrupted and completed goals?

Focusing on the role of goal status in guiding one's prosocial behaviours raises the question as to the relation between goal status and goal salience. This is because one may be more likely to contribute to another agent's goal if that goal is more salient, i.e., goal salience may boost prosociality. In addition, goal salience boosts action prediction (Adam et al., 2016; Eshuis et al., 2009; Henrichs et al., 2012). Goal salience may therefore facilitate coordination and cooperation between agents.

Past studies indicate that incomplete goals are more salient than completed goals (Asiala et al., 2020; Suh & Trabasso, 1993; Trabasso & Suh, 1993). These findings have generated the hypothesis that goal salience is determined by the value of goals in predicting the future behaviour of other agents. However, this past research did not distinguish between interrupted and abandoned goals, which differ in their predictive value. Specifically, abandoned goals should be less useful in predicting the future behaviour of other agents than interrupted goals. We wanted to distinguish between these two types of incomplete goals, which would enable us to test two competing hypotheses about the reason for which goal status determines goal saliency: the predictive value hypothesis, and the goal slippage hypothesis.

To address this gap we presented an adult population with short stories in which an agent's goal was either interrupted, abandoned, or completed. Across three experiments we thus manipulated goal status, asking participants to summarise the stories. Our measure of goal saliency was whether the target goal was mentioned in free recall or not.

We found no evidence that goal status was related to goal salience in Experiments 1 and 3, while the results of Experiment 2 indicated that completed goals were more salient than abandoned or interrupted goals. Taken together, these results contradict the findings of past studies, and they are not predicted by either of the hypotheses that we were testing. We speculated that our failure to replicate past findings may have been due to alterations that we made to the materials used in past studies or the confounds inherent in the previous studies. In sum, on the basis of these experiments, we could not draw conclusions about the relation between goal status and goal salience.

> RQ 4: What is the developmental trajectory of interpersonal commitment? What changes, and why does it change, such that children become better at honouring interpersonal commitments over the course of childhood?

A further type of prosocial behaviour that is prevalent in everyday life, and has recently been the focus of developmental research, is honouring an interpersonal commitment. Of the research that has focused on children's abilities to honour interpersonal commitments (Gräfenhain et al., 2009, 2013; Kachel & Tomasello, 2019; Kanngiesser et al., 2017), relatively little has focused on the developmental trajectory of this ability. However, a better understanding of the developmental trajectory of one's ability to honour an interpersonal commitment would contribute to our understanding of the emergence of prosocial behaviour in childhood, as well as providing insight into the cognitive architecture of interpersonal commitment in adulthood.

To answer the above research question, I first introduced a theoretical framework under which one's ability to honour an interpersonal commitment is reduced to its constitutive components: practical ability, prosocial motivation, and cognitive ability. This theoretical framework carried the advantage that the above research question could be rephrased in terms of the developmental trajectory of these constitutive components of commitment. I then adopted a second-person approach, considering the role

of social interaction in the development of the cognitive abilities and prosocial motivation that are required for one to honour an interpersonal commitment across childhood.

I articulated a particular version of the hypothesis that social interaction scaffolds the development of interpersonal commitment: social interaction introduces rewards (in the form of others' emotions) that reinforce and motivate executive function use and motivation for contributing to others' goals over early childhood. I then evaluated the evidence that pertained to this hypothesis. Although the hypothesis is yet to be directly tested, the predictions that it generates are consistent with the findings from empirical data.

6.2 Theoretical and Practical Implications

This section starts with the theoretical implications of this thesis' research for our understanding of prosociality and goal tracking. I will then outline some speculative suggestions about the potential practical implications of this thesis' research on encouraging prosocial behaviour in everyday life.

6.2.1 Theoretical Implications for Goal Tracking Research

The research in this thesis emphasises the flexibility of goal tracking in early childhood. Past research has demonstrated that goal tracking is sensitive to a variety of factors from early in life, such as communicative utterances (Jin & Song, 2017; Song et al., 2014; Tauzin & Gergely, 2018) and environmental constraints (Csibra, 2008; Gergely et al., 2002; Liu et al., 2017; Liu & Spelke, 2017; Meltzoff, 1995; Sommerville & Woodward, 2005). The present thesis extends this past research by providing preliminary evidence that goal tracking in early childhood is also sensitive to goal status. Specifically, while it had previously been shown that children understand the distinction between interrupted and completed goals (Brandone et al., 2014; Carpenter et al., 1998; Skerry & Spelke, 2014), this thesis provides the first evidence supporting the hypothesis that children also understand goal abandonment, though further research is required to rule out alternative explanations for this finding. By providing preliminary

evidence that children understand goal abandonment, this thesis suggests that children can represent others' goals *as changing* prior to completion. Thus, while past research has demonstrated that goal tracking is sophisticated from early in childhood, this thesis suggests that goal tracking in early childhood is more flexible and dynamic than had previously been thought.

6.2.2 Theoretical Implications for Prosociality Research

This thesis' research carries several implications for our understanding of the development of prosociality. The main implication of this research is that it challenges certain theoretical approaches to the development of prosociality in early childhood. It has been suggested that prosociality in early childhood is underpinned primarily, or even exclusively, by other-regarding motivations (Grossman, 2018; Hepach, 2017; Warneken, 2015). For example, some authors have suggested that prosociality in early childhood is primarily driven by a concern for others and that children are naturally disposed to caring about the welfare of others (Hepach et al., 2017; Warneken, 2015), even to the degree that these prosocial tendencies are not affected (or are minimally affected) by the children's environment (Warneken & Tomasello, 2009). The research in Chapter 3 and Chapter 5 suggests that prosocial behaviour may also be due to more self-oriented sources of motivation and that these are subject to environmental factors (such as reinforcement in the context of social interaction). This thesis therefore empirically and theoretically builds upon more recent theoretical accounts that propose that the origins of prosociality may not be entirely other-oriented (Aknin et al., 2018; Dahl, 2015; Dahl & Brownell, 2019; Dahl & Paulus, 2019; Hammond & Brownell, 2018; Paulus, 2018; Spinrad & Gal, 2018). It is important to bear in mind that there are instrumental helping findings that goal slippage is not best placed to account for, as well as findings that other proposed mechanisms for instrumental helping cannot straightforwardly account for (see section 1.1.4). The relative contribution of self- and other-regarding sources of motivation to the development of prosocial behaviour is an open question

for future empirical research (see section 6.3), as is spelling out the contexts in which different cognitive mechanisms account for helping behaviour.

This thesis' research also carries implications for the interpretation of past studies on instrumental helping and interpersonal commitment. As has been discussed in Chapter 3, the goal slippage mechanism can explain why children instrumentally help others in a variety of contexts. This is because if a child adopted another agent's goal for themselves (as the goal slippage hypothesis predicts) then they would be intrinsically motivated to contribute to that goal, and thus help the helpee. While the goal slippage hypothesis may not be able to account for the findings of *all* previous instrumental helping studies, it can account for the findings of many helping studies that have previously been interpreted as arising from other cognitive mechanisms, such as psychological altruism (Hepach et al., 2017; Hepach et al., 2020; Warneken, 2013; Warneken et al., 2007; Warneken & Tomasello, 2006, 2013). This is because the psychological altruism hypothesis and the goal slippage hypothesis generate the same predictions about children's behaviour in many past studies. The present research therefore prompts a re-evaluation of the findings from past instrumental helping studies.

The theoretical framework for interpersonal commitment introduced in Chapter 5 also prompts a re-evaluation of findings from studies investigating children's ability to honour interpersonal commitments. Past empirical studies typically explain differences between age groups by appealing to the development of children's understanding of the normative implications of forming an interpersonal commitment or of engaging in collaborative activities over childhood (Gräfenhain et al., 2009, 2013; Hamann et al., 2012; Kachel & Tomasello, 2019; Kanngiesser et al., 2017). The theoretical framework of Chapter 5 suggests a further (non mutually exclusive) explanation for differences between age groups, which is simply that children's cognitive abilities (specifically, their executive function) develops over childhood. While these authors do not argue *against* the possibility that executive function plays an important role here, this point is notably absent from discussion in these past studies, and therefore the

relative importance of executive function in the development of interpersonal commitment has been downplayed. This crucial component of one's ability to honour an interpersonal commitment has been largely neglected in empirical studies on interpersonal commitment in childhood, and further research would be required to identify the degree to which differences between age groups can be explained in terms of children's understanding of norms and their cognitive ability.

6.2.3 Practical Implications

There are two main practical implications of this thesis' research, both of which concern encouraging prosocial behaviours (though it is important for further research to be conducted before the suggestions put forward here are implemented). First, the suggestion that prosociality can be encouraged and reinforced over the course of childhood points towards practical steps that could be taken to facilitate the emergence of prosocial behaviours. Specifically, Chapter 5 suggests that we may see an increase in prosocial behaviours if children interact with a wider variety of agents, and more emotionally expressive agents, from an earlier age. The precise practical implications of this will depend on cultural and individual differences between households, but in broad terms the recommendation would be that caregivers should be encouraged to expose children to social interaction with a variety of agents from a young age, and that interventions be held in which caregivers are encouraged to be emotionally expressive around children.

There are several ways that social interaction between children and a wider variety of agents might be encouraged. For instance, caregivers with relatively smaller families (e.g., single-child households) could be encouraged to sign their children up to playgroups and nurseries from a young age, and children from such families could be given greater priority than those from larger families (e.g., households with multiple children or with extended families). Perhaps the main challenge here is that this suggestion would require that families are willing to impart a greater amount of their personal information to nurseries and playgroups, which

they may not be willing to do (for the sake of maintaining their own personal privacy). Imparting this personal information to playgroups and nurseries would also carry implications for nurseries and playgroups in terms of data storage; depending on the quantity of personal information that they receive, they may be required to invest in secure database management systems to ensure that information is securely maintained. In addition to encouraging smaller households to sign children up to nurseries and playgroups, it might also be beneficial to encourage caregivers to sign their children up to nurseries or playgroups from a younger age if the family comes from a cultural or individual context in which emotional expressiveness is relatively low. The main challenge here is that this would require that caregivers' emotional expressiveness is measured (see Ogren et al., 2018), which may be challenging to arrange from a practical perspective for large numbers of caregivers.

Alternatively, instead of addressing caregiver behaviour (by encouraging caregivers to sign children up to nurseries and playgroups), steps could be taken to improve the accessibility of nurseries and playgroups. That is, the present thesis suggests that it may be beneficial for local organisations or governments to ensure that there are sufficient numbers of playroups and nurseries, and that these are accessible to caregivers (i.e., that they are either not overly expensive or that they are subsidised). Although, of course, such steps may require changes in how nurseries and playgroups are funded and organised. While further research that directly addresses the hypotheses considered in Chapter 5 would be important before such steps are implemented, this thesis' research suggests that the implementation of these steps may facilitate the emergence of prosocial behaviour in early childhood.

Rather than focusing on the variety of agents that children interact with, one could instead aim interventions at increasing the emotional expressitivity of caregivers. Such interventions could be aimed at both familial caregivers (e.g., parents) and those working in the child-care industry (e.g., individuals working at nurseries or playgroups). Specifically,

it may be beneficial for children's prosocial development if caregivers were encouraged to express their emotions to a greater degree given that this improves children's emotion understanding (Ogren et al., 2018; Ogren & Johnson, 2020, 2021). The kinds of interventions that would be beneficial here could be as straightforward as simply encouraging caregivers to verbalise their emotional states (e.g., "I am disappointed that this happened") and exhibit behavioural responses to their emotional states (e.g., a mild frown or a smile) more often than they typically would around children. Of course, not *all* emotional expressiveness is necessarily beneficial for children, and encouraging certain kinds of emotional expressiveness may be unethical. For instance, although shouting at or spanking children may constitute as increasing one's emotional expressiveness, this should not be encouraged, and indeed, these particular kinds of emotional expression may promote *anti*-social behaviours in children (Ferguson, 2013; Simons & Wurtele, 2010). Instead, caregivers should be encouraged to express negative emotions in a more subdued and mild manner (e.g., "I did not want this to happen", followed by a mild frown). Designing interventions to encourage caregivers to be more expressive in this manner may therefore promote prosocial behaviour. Such interventions may be especially important for the development of executive functions in children whose caregivers are relatively less emotionally expressive, such as caregivers who suffer from depression or anxiety (Kahn & Garrison, 2009).

On a more speculative note, Chapter 3 points towards ways in which prosocial behaviours may be promoted in an adult population. In occupational psychology there has been a growing interest in cognitive mechanisms that encourage employees to contribute to, or adopt, others' goals (Stajković & Sergeant, 2019). While this research typically focuses on the subconscious priming of goals (Aarts et al., 2008; Chartrand & Bargh, 1996; Kesek et al., 2011), the goal slippage mechanism may offer an alternative means of encouraging agents to contribute to others' goals. Specifically, it suggests that if a goal is framed as having been started but

not finished by another agent, and it is clear that the agent is either unlikely to complete the goal themselves or that they may struggle to complete it themselves, then this may increase one's motivation to contribute to the same goal. While this is yet to be directly tested in a workplace environment, there is evidence that simply framing a goal as having been started and not completed by another agent can boost one's own motivation to complete the goal (Kivetz et al., 2006).

There are several ways in which the goal slippage hypothesis could be tested and implemented in a workplace environment. Take a data entry task as an example. Suppose that there are 1500 rows of data to be added into a spreadsheet. The way in which this task is framed could be manipulated; the task could either be framed as it is (e.g., "You need to add 1500 rows of data into this spreadsheet"), or it could be framed as a task that a colleague has started *but not finished* (e.g., "Your colleague was given the task of entering 2000 rows of data into this spreadsheet, but after entering the first 500 rows they have been interrupted by an urgent project"). The goal slippage hypothesis predicts that agents will be more motivated to complete this task in the latter case (i.e., if it is framed as having been started by someone else). There are different ways in which motivation could be measured here. For instance, one might predict that the task would be completed faster if it were presented as being partially completed by one's colleague (Kivetz et al., 2006). Alternatively, if the task were not assigned to an individual and instead simply presented as a task that *someone* within the team will have to complete, then one might predict that more people will volunteer to complete the task if it were framed as being partially completed by someone else. The particular example presented here has been data entry, but one could easily extend this to other kinds of tasks. For instance, marking exam scripts or neatly organising a storage room are tasks that could be framed as either partially completed by one's colleague or not yet started.

There are a few things that one would need to be wary of in implementing this suggestion. Firstly, the explanation for why the task is

incomplete would need to be considered carefully. Specifically, one would need to frame the task such that employees are not jealous or irritated at having to take over their colleague's workload. For instance, if the task were framed as incomplete because the colleague has gone on holiday, this may engender feelings of jealousy in the employee who has to complete the task. This feeling of jealousy may be counter-productive, decreasing motivation to complete the task. It may instead be better to frame the task as incomplete because one's colleague has been given an urgent project at short notice.

Another factor to consider is progress towards the goal. Specifically, it is important that progress towards the goal be measurable, such that it is clear that the task has indeed been partially completed by one's colleague. In the case of entering rows of data into a spreadsheet, this is quite straightforward: one can identify progress by the number of rows of data that have been entered into the spreadsheet thus far. One could also clearly identify progress in the case of marking exam scripts (e.g., the number of scripts that have been marked) or the organisation of a storage room (e.g., the number of boxes that have organised). However, many tasks, especially those that are more creative in nature, may not have such a straightforward path of progression. For example, consider designing a new marketing initiative; in what way would one measure progress towards the design of a new marketing campaign? Even if one's colleague had been working on this task for a number of hours, it is perhaps less clear how progress towards that goal might be measured. Taking over this task from one's colleague, even if they have been working on it for hours, may therefore feel as if one is starting the task from the beginning. This is an important matter because the premise of this line of enquiry is that one can frame the task as having been started by someone else. Due to the challenges in framing more abstract goals as being partially completed, it may be best to initially focus on more straightforward tasks.

Thirdly, one would need to consider the factors beyond goal slippage that might motivate agents to complete tasks that others have started in a workplace context. For example, employees might be motivated to complete

a task that their colleague has started simply because they like their colleague, and would therefore like to lessen their colleague's future workload. Alternatively, one might be motivated to complete a task that one's colleague has started for strategic reasons, such as reinforcing one's reputation as a 'team player' or in the hope of receiving reciprocal behaviour from that colleague in future. In order to address this, it would be best to ensure that the colleague who has 'started' the task is someone that the employee does not know well, or someone that they are unlikely to interact with in future. This could be achieved by presenting the colleague as coming from a different department, or else by simply not naming the colleague. In addition, it would be ideal if the task were one that could be completed anonymously, in order to avoid the possibility that completing the task could improve one's reputation at work (although this may be difficult to practically arrange).

6.3 Limitations and Future Research

Several directions for future research have been spelled out in each chapter of this thesis. In this section, I will briefly reiterate these while adding further suggestions. The main suggestions for future research in Chapter 2 were to investigate how children prioritise and integrate different sources of information (e.g., communicative utterances, emotional cues, goal status) in identifying others' goals, as well as the impact on social learning of goal abandonment. To add to these suggestions, it may be fruitful to investigate children's understanding of goal abandonment using different methods than those used in Chapter 2. In Chapter 2, we studied children's understanding of goal abandonment in an instrumental helping context. This was because we were interested in how children's understanding of goal status guided their prosocial behaviour. However, the instrumental helping context may be too complex for children younger than 2 years old to exhibit their understanding of goal status, and specifically, to exhibit their understanding of the distinction between abandoned and interrupted goals (Hobbs & Spelke, 2015; Krogh-Jesperson et al., 2015). One could alternatively investigate children's understanding of goal

abandonment in different contexts. That is, one could adopt measures used in other goal tracking studies (e.g., predictive gaze) to investigate whether children understand goal abandonment from a younger age. This method would not have been appropriate for Chapter 2 because predictive gaze alone does not provide insight into how goal tracking guides prosocial behaviour. However, if one were solely investigating goal tracking itself, then predictive gaze would be appropriate.

Testing for children's understanding of goal abandonment using predictive gaze could be done by adapting the experiment in Chapter 2. Using only facial expression, the experimenter could establish their initial goal (e.g., by smiling and nodding towards the location that they want to place the toy) before either abandoning the goal (e.g., by halting their action, looking again at the box, and then shaking their head and frowning) or having their goal interrupted (e.g., by encountering an obstacle). The experimenter could then return to a neutral position, at which point predictive gaze could be recorded. In the interrupted condition children should predictively look towards the initial goal location because they should expect the experimenter to continue to pursue the goal of placing the toy into that location. In the abandoned condition, in contrast, children should predictively look towards the alternative location in anticipation of the experimenter's new goal-directed action. One advantage of this design is that it avoids the concerns around the design of the experiment in Chapter 2 (namely, that the experimenter explicitly told children where he wanted the toy to go after the experimental manipulation).

There are several benefits of conducting a study which addresses the same question as Chapter 2, while using predictive gaze instead of instrumental helping as a measure of children's goal tracking abilities. Firstly, it is less challenging for children to indicate their understanding of others' goals with predctive gaze in comparison to, say, helping behaviour (Hobbs & Spelke, 2015; Krogh-Jesperson et al., 2015), insofar as children only need to predict what the experimenter will do next (instead of using this information to guide their own behaviour). Therefore with this method

it should be possible to test for children's understanding of the distinction between abandoned and interrupted goals with a younger population than that in Chapter 2. Secondly, replicating the results of Chapter 2 with different research methods would also increase the generalisability of these results. Thirdly, this would enable the testing of goal slippage in children younger than 24 to 30 months of age (see below).

The main direction for future research suggested in Chapter 3 was to deepen our understanding of the goal slippage mechanism by identifying factors that may promote goal slippage and testing for goal slippage in adults and non-human animals. In addition to these suggestions, one direction for future research would be to investigate whether there is evidence for goal slippage in a younger population than that which we tested. This hinges on first uncovering evidence that children younger than 24 months of age understand goal abandonment. Following this, one could make the necessary stimuli adjustments to the materials used in Chapter 3, such as ensuring that all containers are fully made of cardboard (see section 3.4), before running the same experiment with a younger population (e.g., 12- to 18-month olds). The procedure and measures could otherwise remain the same as in Chapter 3, as past helping studies with a younger population have used similar measures and procedures (Hepach et al., 2017; Svetlova et al., 2010; Warneken & Tomasello, 2006). The predictions about children's behaviour would be the same as in Chapter 3, i.e., we would predict that participants would be more likely to place the remaining toy into the experimenter's goal container in the experimental condition (when this constitutes completing the experimenter's goal), than in the control condition (in which the experimenter's goal has already been completed). This would provide evidence that prosocial behaviours may be motivated by more self-oriented sources from an earlier age.

In Chapter 3 it was suggested that one could test for goal slippage in an adult population. This is an important line of enquiry because the existing empirical support for the goal slippage hypothesis with an adult population is indirect (see section 1.1.4); the only experiment that has directly tested the

goal slippage hypothesis (i.e., the experiment in Chapter 3) focused on 2-year-olds. It is therefore a further question as to whether (and to what degree) goal slippage influences behaviour in adulthood. In a laboratory context, one could test the goal slippage hypothesis in a similar way to which it is tested in Chapter 3. That is, one could design an experiment in which the experimenter's goal-directed actions are the same in two conditions, but the experimenter's goal is either completed or abandoned. For instance, participants could be brought into a waiting room with a confederate and told that they should wait in this room for further instructions. The confederate could initiate some goal-directed action, such as tidying up the coffee table on which there lie several magazines and books. In the control condition the goal might be to place all magazines on a nearby bookshelf, whereas in the experimental condition condition the goal might be to place all magazines *and books* on the nearby bookshelf. In both conditions the confederate's goal could be established verbally (e.g., with an off-hand comment: "While I'm waiting I'm just going to tidy up a bit here, by putting these magazines [and books] on the bookshelf") before placing all magazines on the bookshelf. In the control condition this would constitute completing the goal ("There, all the magazines are on the bookshelf now"), whereas in the experimental condition the confederate could abandon the goal at this point (e.g., "Eh, I can't be bothered putting the books on the bookshelf too, never mind"). In both conditions the confederate could then be called away by the main experimenter (to 'participate in the experiment'). The participant could then be left alone in the waiting room for some further amount of time (e.g., 1-2 minutes), during which time the number of books that the participant places on the bookshelf could be measured. The goal slippage hypothesis predicts that participants will be more likely to place the books on the bookshelf in the experimental condition (when this constitutes completing the confederate's abandoned goal) than in the control condition (when the confederate's goal has been completed). Of course, this is only a single experiment, and it would be best to design multiple experiments with a similar structure (i.e., in which a

confederate's goal is either abandoned or completed) to ensure that any effects are replicated.

A further suggestion from Chapter 3 was to test for goal slippage in non-human animals. This may be somewhat more challenging than in an adult human population because non-human animals' understanding of goal abandonment is yet to be investigated. There is some reason to believe that at least some non-human animals, such as chimps, may be able to understand goal abandonment; chimps understand goal interruption insofar as they are willing and able to help conspecifics when their goal-directed actions are interrupted (Engelmann et al., 2019; Warneken et al., 2007), and chimps understand the distinction between humans being unwilling and unable to perform an action (Call et al., 2004). The same predictive-gaze experiment could be used to test for goal abandonment understanding in non-human animals as with children younger than 24-30 months of age (discussed above). The main challenge here is whether non-human animals will be sensitive to the cues that are used to indicate goal abandonment and goal interruption, namely, facial expression and gesture (e.g., frowning and shaking one's head). Although chimps do take into account facial expression in tracking humans' goals, they do so to a lesser degree than human infants (Myowa-Yamakoshi et al., 2012). Instead of relying on facial expression, then, one could adapt the proposed predictive-gaze experiment in such a way that the experimenter's goal status were established with their actions. That is, one could alter the experiment such as to exaggerate the experimenter's actions. For instance, in the interrupted condition the experimenter could encounter a more substantial barrier and bump into it in a more exaggerated manner. In the abandoned condition the different actions required to achieve the two possible goals could be exaggerated, such that it is more apparent on the basis of the experimenter's actions that he had started to achieve one goal but then switched to an alternative goal. This exaggeration could simply involve the goal locations being further apart, for example.

Once goal abandonment understanding were uncovered in a population of non-human animals, one could test for goal slippage within this population. For instance, one could adapt the measures used in past studies which demonstrate that chimps are willing and able to help humans (Warneken & Tomasello, 2006), such that in one condition the experimenter's goal is completed and in the other condition it is abandoned. It could then be measured whether chimps are more likely to perform some target action when it constitutes completing the experimenter's abandoned goal (experimental condition) as compared to when the action is superfluous to the already completed-goal (control condition). The goal slippage hypothesis predicts that chimps will be more likely to perform the action in the experimental condition than in the control condition. This would indicate whether goal slippage is a uniquely human cognitive mechanism, or if it is also found in non-human animals.

A further direction for future research inspired by Chapter 3 would be to investigate the degree to which different mechanisms account for instrumental helping behaviour, and to investigate whether the relative prominence of different mechanisms in accounting for helping behaviour changes over childhood. Past research has already provided some insight into changes in motivation for instrumental helping over the course of childhood. For instance, the earliest evidence of children helping others strategically (i.e., in order to manage their reputation) indicates that this mechanism emerges from the age of 5 years old (Engelmann & Rapp, 2018). Therefore we would expect this particular mechanism to become more prominent in accounting for helping behaviour as children reach and surpass 5 years of age. However, before identifying changes over the course of childhood, we would first need to identify the earliest age at which each mechanism emerges. For example, goal slippage has only been tested with one age-group (24- to 30-month-olds), and it is therefore not yet possible to identify whether there are changes in the degree to which goal slippage underpins instrumental helping behaviour over the course of childhood.

Independently from investigating changes over childhood, one could investigate the degree to which different mechanisms underpin helping behaviour in an individual age-group. For example, one could investigate the degree to which instrumental helping is underpinned by psychological altruism, goal slippage, and a preference for social interaction in 24- to 30-month-olds. This would require designing a series of experiments to test for each of these mechanisms separately (i.e., in scenarios in which differences between condition could only be explained by *one* of these mechanisms) while keeping experimental design, including apparatus materials and experimenter behaviour (e.g., verbalisations and facial expressions), as similar as possible. The differences in rates of helping behaviour could then be compared across experiments. Testing for goal slippage independently of psychological altruism and social interaction can be done with a version of the experiment presented in Chapter 3 (although changes would need to be made to that experiment's materials; see section 3.4). As discussed in section 3.1.1, we would not expect psychological altruism or a desire to socially interact to explain the results of such an experiment.

In order to investigate the social interaction hypothesis independently of psychological altruism and goal slippage, one would need to design an experiment in which the possibility of interacting with the helpee is manipulated between conditions, but the help required by the helpee to achieve their goal were kept constant. Suppose that an experimenter were performing some task, such as placing toys into a container. The experimenter could 'accidentally' drop a toy such that it is out of reach, while the participant is in a position to help by placing the toy into the container. The experimental manipulation could revolve around whether the experimenter and the child are able to interact with and see each other; in one condition there could be a barrier between them (such that the experimenter cannot see the participant and would not know if the participant completed their goal), and in the other condition there could be no such barrier. The social interaction hypothesis predicts that children will be more likely to help the experimenter when the barrier is absent, allowing

for social interaction with the experimenter, in comparison to when the barrier is present. With this design, any motivation to help the experimenter due to altruism or goal slippage is kept constant between conditions (as the experimenter requires equal help to achieve their goal in both conditions), and therefore any difference in helping behaviour between conditions would be best accounted for by children's motivation to interact with the experimenter.

Finally, in order to investigate the degree to which psychological altruism underpins instrumental helping independently from social interaction and goal slippage, one could design an experiment in which only the helpee's well-being is manipulated between conditions. That is, one could design an experiment in which it is possible to interact with the experimenter and contribute to the experimenter's goal in both conditions, but the completion of the goal itself is not equally beneficial for the experimenter's well-being in both conditions. The main challenge here is in ensuring that children recognise that the experimenter's well-being differs between conditions. Suppose that an experimenter needs help in placing a toy into one of two containers (e.g., if the containers are out of reach). Placing the toy into one container could do nothing, while placing the toy into the other container could be mildly unpleasant for the experimenter (e.g., it will result in a toy being dropped onto the experimenter's foot, causing mild discomfort that is communicated verbally and with facial expression). The experimental manipulation could simply be which container the experimenter wants to put the toy into: the container that will have no effect on the experimenter's well-being (control condition), or the one that has a negative effect (experimental condition). Any differences between conditions here could only be explained by children's concern for the experimenter's well-being – the possibility of completing the experimenter's goal and socially interacting with the experimenter would be kept constant between conditions. The psychological altruism hypothesis predicts that children should be less likely to instrumentally help the experimenter in the experimental condition (where helping would negatively

affect the experimenter's well-being) than in the control condition. This would test the degree to which concern for others' well-being affects children's motivation to help others.

The relative rates of helping in these three experiments could be compared, in order to provide some indication of the degree to which these three mechanisms relatively contribute to instrumental helping behaviour. The main challenge here would be in designing the experiments in order to try to maintain consistency between them as much as possible in terms of, say, the structure of the tasks and the experimenter's behaviour. It is important to be aware that helping rates can differ substantially between experiments (Waugh & Brownell, 2017), and it is therefore important to try to design these experiments as similarly as possible if the comparisons are to be meaningful.

The suggestions for future research put forward in Chapter 4 primarily concern our failure to replicate past findings or uncover evidence supporting either of the hypotheses that we tested for. We suggested attempting to replicate past studies with precisely the same materials that they used. We also suggested using a different measure for goal saliency than that used in our experiments. For instance, some past studies used response time as a measure of goal saliency (Lutz & Radvansky, 1997; Magliano & Radvansky, 2001; Radvansky & Curiel, 1998), and one could also test goal saliency using live-action materials instead of vignettes. Using multiple measures of goal saliency would increase the generalizability of any results, and live-action materials may increase the ecological validity of any findings. Unfortunately, it is difficult to identify further areas for future research given the lack of positive findings in Chapter 4, as well as the failed replication of past studies, beyond those suggested in section 4.5.

The suggestions put forward in Chapter 5 focused on directly testing the hypothesis at the core of that chapter: that social interaction scaffolds the development of interpersonal commitment by introducing rewards (i.e., others' emotions) that motivate and reinforce prosocial motivation and executive function. In addition to this, it may also be fruitful for future

research to apply the theoretical framework outlined in Chapter 5 to research on interpersonal commitment in adults. When adults are faced with the challenge of honouring an interpersonal commitment, they too must be sufficiently prosocially motivated, cognitively able, and practically able to honour the commitment. This framework could therefore be used to extend and contribute to ongoing research on the psychological underpinnings of commitment in adulthood (Michael, Knoblich, & Sebanz, 2016; Powell & Michael, 2019; Székely et al., 2019; Székely & Michael, 2018).

More specifically, this framework can be fruitfully applied to both past and future research focusing on an adult population. It would be insightful to apply this framework to past research because it would draw attention to relevant factors that may have been relatively neglected in the past study of interpersonal commitment. For instance, recent research on interpersonal commitments in adults has typically neglected discussions about executive function (Székely et al., 2019; Székely & Michael, 2018), even though this is a highly relevant factor in identifying whether an agent will persist with a given task. Reconsidering past research with this framework in mind may help to articulate expectations about the relative contributions of each component of interpersonal commitment in explaining the results of these studies, thereby highlighting factors that should be addressed more explicitly in future research.

Looking ahead to future research, the present framework would help to ensure that future investigations focusing on any one of these components control for the other two components. For example, it is important that future experiments that manipulate prosocial motivation ensure that each condition is matched in terms of executive function and practical challenge. For executive function, this may involve ensuring that all conditions are matched in terms of the number of factors that agents must attend to (i.e., attentional control), and that the persistence behaviour required for each condition is kept constant (i.e., inhibitory control). For example, this might require ensuring that agents must attend to an equal number of individual objects across conditions, or that any temptations to defect from a task (e.g.,

for financial reward) are kept constant across conditions. Alternatively, one could control for executive function on the basis of individual participants. That is, one could subject participants to a variety of executive function tasks in order to identify individual differences in performance on executive function tasks, such that these individual differences in performance could be controlled for in subsequent analysis (Diamond, 2013; Hofmann et al., 2012).

Controlling for the level of practical challenge in an experiment may be relatively straightforward given that recent past studies investigating interpersonal commitment in adults have taken place on computers (Chennells & Michael, 2018; Székely et al., 2019; Székely & Michael, 2018). However, the practical element of any future experimental setup would need to be kept in mind if future studies moved away from experiments conducted on computers, especially if task requirements were physically taxing. This is important because identical practical tasks may not be equally challenging for all agents. For example, suppose that an experiment required agents to commit to moving a heavy object from one location to another (Sommerville et al., 2018). Individual differences in participants' strength would influence how practically challenging this task is for them, and therefore it may affect how long they persist with the task. This is important because a common measure of interpersonal commitment is persistence (Chennells & Michael, 2018). In designing future experiments, then, either the practical challenge of a task would need to be kept at a relative minimum for all participants (by using tasks that are not particularly physically taxing, e.g., tasks taking place on a computer) or else the practical ability of participants would need to be controlled for (e.g., participants' strength could be measured and the weight of objects manipulated on this basis). In this way, the framework proposed in Chapter 5 could guide future research design.

A further area for new research comes from the idea put forward in Chapter 5 that commitment in the individual case may build upon commitment in the interpersonal case. One suggestion was that if social

interaction scaffolds the emergence of interpersonal commitment (and specifically, executive function), then one's ability to honour interpersonal commitments may develop prior to one's ability to remain committed to individual tasks, or at least that the individual case builds upon the interpersonal case. This generates some specific predictions. If individual commitment is parasitic on interpersonal commitment, then we should expect improvements in children's ability to honour interpersonal commitments to emerge prior to improvements in individual commitment, or at least that we should not see improvements in individual commitment prior to improvements in interpersonal commitment.

This prediction about the relation between individual and interpersonal commitment development could be tested with a lab-based longitudinal study in which participants are given a task that requires executive function. Using inhibitory control as an example, participants could be given an extremely long task that is boring or effortful, such as a challenging puzzle or moving a large amount of items from one location to another. In one condition this task could be an individual task that participants perform themselves, and in the other condition the task could be joint (i.e., participants perform the task with a confederate). Commitment to the task could be measured by the amount of time that participants persist with the task before giving up. Implementing a within-subjects design, participants would need to take part in this task every, say, 3 or 6 months from 2-3 years of age until 6 years of age (although they would need to come into the laboratory on separate days for each version of the task, in order to avoid cognitive fatigue).

We might expect participants to perform better on the joint task than the individual task (e.g., because participants may find the joint task more engaging than the individual task due to the social element). We might also expect some improvement in performance in both conditions, simply due to general neurological and physiological developments that take place over the course of childhood (Doebel, 2019; Miyake & Friedman, 2012; Nigg, 2017). The main research question, however, would focus on *relative*

improvement in performance over time in each condition, rather than directly comparing performance in each condition. If individual commitment builds upon interpersonal commitment, we would expect participants' performance to either improve first in the interpersonal case and only later in the individual case, or at least we would expect that performance in the individual case will not improve prior to improvements in the performance of the joint task (i.e., they might improve at the same time). We would therefore predict an interaction effect: over time, performance in both conditions would improve, but we would expect performance to improve faster in the joint version of the task. While this prediction is somewhat speculative, it is consistent with the suggestions from Chapter 5.

6.4 Conclusion

Goal tracking and prosocial behaviour are foundational in human social life, and together they have shaped the evolution of human society. This thesis has investigated how these abilities complement each other in development. Specifically, this thesis has focused on the relation between an underexplored aspect of goal tracking, goal status, and one of the earliest emerging forms of prosocial behaviour, instrumental helping, as well as the developmental trajectory of interpersonal commitment. I have addressed four research questions that concern the development of (pro)social cognition. By answering these questions, this thesis contributes to ongoing research into our understanding of (pro)social cognition and behaviours in early childhood. Specifically, this thesis extends our understanding of the relation between goal tracking and some of the earliest emerging forms of prosocial behaviour.

References

Aarts, H., Dijksterhuis, A., & Dik, G. (2008). Goal contagion: Inferring goals from others' actions—And what it leads to. In J. Shah & W. Gardner (Eds.), *Handbook of motivation science* (pp. 265–280). Guilford.

Adam, M., Reitenbach, I., & Elsner, B. (2017). Agency cues and 11-month-olds' and adults' anticipation of action goals. *Cognitive Development, 43*, 37–48. https://doi.org/10.1016/j.cogdev.2017.02.008

Adam, M., Reitenbach, I., Papenmeier, F., Gredebäck, G., Elsner, C., & Elsner, B. (2016). Goal saliency boosts infants' action prediction for human manual actions, but not for mechanical claws. *Infant Behavior and Development, 44*, 29–37. https://doi.org/10.1016/j.infbeh.2016.05.001

Aknin, L. B., Van de Vondervoort, J. W., & Hamlin, J. K. (2018). Positive feelings reward and promote prosocial behavior. *Current Opinion in Psychology, 20*, 55–59. https://doi.org/10.1016/j.copsyc.2017.08.017

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine, 18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

Applin, J. B., & Kibbe, M. M. (2019). Six-month-old infants predict agents' goal-directed actions on occluded objects. *Infancy, 24*(3), 392–410. https://doi.org/10.1111/infa.12282

Asiala, L. K. E., Chan, G. C., Kurby, C. A., & Magliano, J. P. (2020). The role of goals and goal barriers in predicting the outcomes of intentional actions in the contexts of narrative text. *Journal of Cognitive Psychology, 32*(1), 82–92. https://doi.org/10.1080/20445911.2019.1690494

Astington, J. (1988a). Children's understanding of the speech act of promising. *Journal of Child Language, 15*, 157–173. https://doi.org/10.1017/s0305000900012101

Astington, J. (1988b). Promises: Words or deeds? *First Language*, *8*, 259–270. https://doi.org/10.1177/014272378800802404

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

Balliet, D., Tybur, J. M., & Van Lange, P. A. M. (2017). Functional Interdependence Theory: An Evolutionary Account of Social Situations. *Personality and Social Psychology Review*, *21*(4), 361–388. https://doi.org/10.1177/1088868316657965

Bandstra, N. F., Chambers, C. T., McGrath, P. J., & Moore, C. (2011). The behavioural expression of empathy to others' pain versus others' sadness in young children. *PAIN*, *152*(5), 1074–1082. https://doi.org/10.1016/j.pain.2011.01.024

Barnes, J. L., Hill, T., Langer, M., Martinez, M., & Santos, L. R. (2008). Helping behaviour and regard for others in capuchin monkeys (Cebus apella). *Biology Letters*, *4*(6), 638–640. https://doi.org/10.1098/rsbl.2008.0410

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi:10.1016/j.jml.2012.11.001

Barragan, R., Brooks, R., & Meltzoff, A. N. (2020). Altruistic food sharing behavior by human infants after a hunger manipulation. *Scientific Reports, 10(1)*, 1–9. https://doi.org/10.1038/s41598-020-58645-9

Barragan, R., & Dweck, C. (2014). Rethinking natural altruism: Simple reciprocal interactions trigger children's benevolence. *Proceedings of the National Academy of Sciences*, *111*(48), 17071–17074. https://doi.org/10.1073/pnas.1419408111

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, *19*, 107–154. https://doi.org/10.1017/s0140525x00041960

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

Batson, C. (2010). *Altruism in Humans*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195341065.001.0001

Batson, C., & Powell, A. (2003). Altruism and prosocial behavior. In T. Millon, M. Lerner, & I. Weiner (Eds.), *Handbook of Psychology: Volume 5, Personality and Social Psychology*. John Wiley & Sons, Inc.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology*, *41*(2), 328–337. https://doi.org/10.1037/0012-1649.41.2.328

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology Section A*, *53*(1), 153–164. https://doi.org/10.1080/713755872

Berlin, L. J., Ispa, J. M., Fine, M. A., Malone, P. S., Brooks-Gunn, J., Brady-Smith, C., Ayoub, C., & Bai, Y. (2009). Correlates and consequences of spanking and verbal punishment for low-income White, African American, and Mexican American toddlers. *Child Development, 80(5),* 1403–1420. https://doi.org/10.1111/j.1467-8624.2009.01341.x

Bernier, A., Carlson, S. M., & Whipple, N. (2010). From external regulation to self-regulation: Early parenting precursors of young children's executive functioning. *Child Development*, *81*(1), 326–339. https://doi.org/10.1111/j.1467-8624.2009.01397.x

Best, J., & Miller, P. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Biro, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: Development through cue-based bootstrapping.

Developmental Science, *10*(3), 379–398. https://doi.org/10.1111/j.1467-7687.2006.00544.x

Bischof-Köhler, D. (1991). The development of empathy in infants. In M. E. Lamb & H. Keller (Eds.), *Infant development: Perspectives from German-speaking countries* (pp. 245–273). Routledge.

Blackwell, E. J., Twells, C., Seawright, A., & Casey, R. A. (2008). The relationship between training methods and the occurrence of behavior problems, as reported by owners, in a population of domestic dogs. *Journal of Veterinary Behavior*, *3*(5), 207–217. https://doi.org/10.1016/j.jveb.2007.10.008

Blair, C., Granger, D. A., Willoughby, M., Mills-Koonce, R., Cox, M., Greenberg, M. T., Kivlighan, K. T., & Fortunato, C. K. (2011). Salivary cortisol mediates effects of poverty and parenting on executive functions in early childhood. *Child Development*, *82*(6), 1970–1984. https://doi.org/10.1111/j.1467-8624.2011.01643.x

Blair, R. J. R. (2017). Emotion-based learning systems and the development of morality. *Cognition*, *167*, 38–45. https://doi.org/10.1016/j.cognition.2017.03.013

Bonalumi, F., Isella, M., & Michael, J. (2019). Cueing implicit commitment. *Review of Philosophy and Psychology*, *10*(4), 669–688. https://doi.org/10.1007/s13164-018-0425-0

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of Extension*, *50*(2), 1–5.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803–843. https://doi.org/10.1086/673885

Bouquet, C. A., Shipley, T. F., Capa, R. L., & Marshall, P. J. (2011). Motor contagion: Goal-directed actions are more contagious than non-goal-directed actions. *Experimental Psychology, 58(1)*, 71–78. https://doi.org/10.1027/1618-3169/a000069

Bower, G. H., & Rinck, M. (1999). Goals as generators of activation in narrative understanding. In S. Goldman, A. Graesser, & P. van den

Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 111–134).

Brandone, A. C., Horwitz, S. R., Aslin, R. N., & Wellman, H. M. (2014). Infants' goal anticipation during failed and successful reaching actions. *Developmental Science*, *17*(1), 23–34. https://doi.org/10.1111/desc.12095

Bratman, M. (1999). I Intend that we J. In *Faces of Intention: Selected Essays on Intention and Agency* (pp. 142–162). CUP.

Bratman, M. (2009). Intention, belief, practical, theoretical. In S. Robertson (Ed.), *Spheres of Reason: New Essays in the Philosophy of Normativity* (pp. 29–59). OUP.

Bratman, M. (2014). *Shared agency: A planning theory of acting together*. OUP.

Bridgers, S., & Gweon, H. (2018). Means-Inference as a source of variability in early helping. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01735

Brooks, J. L. (2012). Counterbalancing for serial order carryover effects in experimental condition orders. *Psychological Methods*, *17*(4), 600. https://doi.org/10.1037/a0029310

Broomell, A. P. R., Smith, C. L., Calkins, S. D., & Bell, M. A. (2020). Context of maternal intrusiveness during infancy and associations with preschool executive function. *Infant and Child Development*, *29*(1), e2162. https://doi.org/10.1002/icd.2162

Brownell, C. (2011). Early developments in joint action. *Review of Philosophy and Psychology*, *2*(2), 193–211. https://doi.org/10.1007/s13164-011-0056-1

Brownell, C. A. (2013). Early development of prosocial behavior: Current perspectives. *Infancy*, *18*(1), 1–9. https://doi.org/10.1111/infa.12004

Brownell, C. A. (2016). Prosocial behavior in infancy: The role of socialization. *Child Development Perspectives*, *10*(4), 222–227. https://doi.org/10.1111/cdep.12189

Brownell, C. A., Svetlova, M., Anderson, R., Nichols, S. R., & Drummond, J. (2013). Socialization of early prosocial behavior: Parents' talk about emotions is associated with sharing and helping in toddlers. *Infancy : The Official Journal of the International Society on Infant Studies*, *18*, 91–119. https://doi.org/10.1111/j.1532-7078.2012.00125.x

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics*, *2*(1), 671–698. https://doi.org/10.1146/annurev.economics.102308.124324

Bshary, R., & Raihani, N. J. (2017). Helping in humans and other animals: A fruitful interdisciplinary dialogue. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1863), 20170929. https://doi.org/10.1098/rspb.2017.0929

Burke, W. W. (2011). A perspective on the field of organization development and change: The Zeigarnik effect. *The Journal of Applied Behavioral Science*, *47*(2), 143–167. https://doi.org/10.1177/0021886310388161

Butler, L. P., & Walton, G. M. (2013). The opportunity to collaborate increases preschoolers' motivation for challenging tasks. *Journal of Experimental Child Psychology*, *116*(4), 953–961. https://doi.org/10.1016/j.jecp.2013.06.007

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

Butterfill, S. A. (2019). Goals and targets: A developmental puzzle about sensitivity to others' actions. *Synthese*. https://doi.org/10.1007/s11229-019-02214-9

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). 'Unwilling' versus 'unable': Chimpanzees' understanding of human intentional action. *Developmental Science, 7(4)*, 488–498. https://doi.org/10.1111/j.1467-7687.2004.00368.x

Callaghan, T., & Corbit, J. (2018). Early prosocial development across cultures. *Current Opinion in Psychology, 20*, 102–106. https://doi.org/10.1016/j.copsyc.2017.07.039

Campbell, S. B., Leezenbaum, N. B., Schmidt, E. N., Day, T. N., & Brownell, C. A. (2015). Concern for another's distress in toddlers at high and low genetic risk for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(11), 3594–3605. https://doi.org/10.1007/s10803-015-2505-7

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

Carlson, V. J., & Harwood, R. L. (2003). Attachment, culture, and the caregiving system: The cultural patterning of everyday experiences among Anglo and Puerto Rican mother–infant pairs. *Infant Mental Health Journal, 24(1),* 53–73. https://doi.org/10.1002/imhj.10043

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Carlson, S. M. (2009). Social origins of executive function development. *New Directions for Child and Adolescent Development*, *2009*(123), 87–98. https://doi.org/10.1002/cd.237

Carpendale, J. I., Hammond, S. I., & Atwood, S. (2013). Chapter Six—A Relational Developmental Systems Approach to Moral Development. In R. M. Lerner & J. B. Benson (Eds.), *Advances in Child Development and Behavior* (Vol. 45, pp. 125–153). JAI. https://doi.org/10.1016/B978-0-12-397946-9.00006-3

Carpendale, J. I., Kettner, V. A., & Audet, K. N. (2015). On the nature of toddlers' helping: Helping or interest in others' activity? *Social Development, 24*(2), 357–366. https://doi.org/10.1111/sode.12094

Carpendale, J. I., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within

social interaction. *Behavioral and Brain Sciences*, *27*(1), 79–96. https://doi.org/10.1017/S0140525X04000032

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., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315–330. https://doi.org/10.1016/S0163-6383(98)90009-1

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve-and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13–F20. https://doi.org/10.1111/j.1467-7687.2004.00385.x

Carver, L. J., & Vaccaro, B. G. (2007). 12-month-old infants allocate increased neural resources to stimuli associated with negative adult emotion. *Developmental Psychology*, 43(1), 54–69. https://doi.org/10.1037/0012-1649.43.1.54

Castro, V., & Pacherie, E. (2020). Joint actions, commitments and the need to belong. *Synthese*. https://doi.org/10.1007/s11229-020-02535-0

Chan, G. C., Magliano, J. P., & O'Brien, E. J. (2018). Processing the outcomes of characters' actions: The impacts of character goals and situational context. *Discourse Processes*, *55*(5–6), 497–509. https://doi.org/10.1080/0163853X.2018.1435147

Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, *71*(3), 464–478. https://doi.org/10.1037//0022-3514.71.3.464

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—Simulation and Computation®*, 39(4), 860–864. https://doi.org/10.1080/03610911003650383

Chen, S.-Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease*, *9*(6), 1725-1729–1729. https://doi.org/10.21037/jtd.2017.05.34

Cheng, Y., Lee, S.-Y., Chen, H.-Y., Wang, P.-Y., & Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *Journal of Cognitive Neuroscience*, *24*(6), 1411–1419. https://doi.org/10.1162/jocn_a_00214

Chennells, M., & Michael, J. (2018). Effort and performance in a cooperative activity are boosted by perception of a partner's effort. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-34096-1

Chiarella, S. S., & Poulin-Dubois, D. (2013). Cry babies and pollyannas: Infants can detect unjustified emotional reactions. *Infancy*, *18*, E81–E96. https://doi.org/10.1111/infa.12028

Chin, J.-C., & Lin, M.-H. (2018). Children's understanding of conditional promise contract violations. *Infant and Child Development*, *27*(3), e2082. https://doi.org/10.1002/icd.2082

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, 19(22), 3127–3131. https://doi.org/10.1002/1097-0258(20001130)19:22<3127::aid-sim784>3.0.co;2-m

Choi, Y., Mou, Y., & Luo, Y. (2018). How do 3-month-old infants attribute preferences to a human agent? *Journal of Experimental Child Psychology*, *172*, 96–106. https://doi.org/10.1016/j.jecp.2018.03.004

Christ, S. E., Holt, D. D., White, D. A., & Green, L. (2007). Inhibitory control in children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *37*(6), 1155–1165. https://doi.org/10.1007/s10803-006-0259-y

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011. https://doi.org/10.1111/desc.12193

Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2016). Social effects of movement synchrony: Increased infant helpfulness only transfers to affiliates of synchronously moving partners. *Infancy*, *21*(6), 807–821. https://doi.org/10.1111/infa.12140

Clark, H. (2006). Social actions, social commitments. In N. Enfield & S. Levinson (Eds.), *Roots of human sociality: Culture, cognition, and interaction* (pp. 126–150). Berg.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Cohen, A., Bayer, U., Jaudas, A., & Gollwitzer, P. (2007). Self-regulatory strategy and executive control: Implementation intentions modulate task switching and Simon task performance. *Psychological Research*, *72*(1), 12–26. https://doi.org/10.1007/s00426-006-0074-2

Cohn, N., & Bender, P. (2017). Drawing the line between constituent structure and coherence relations in visual narratives. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(2), 289–301. https://doi.org/10.1037/xlm0000290

Cook, P. F., Prichard, A., Spivak, M., & Berns, G. S. (2016). Awake canine fMRI predicts dogs' preference for praise vs food. *Social Cognitive and Affective Neuroscience*, *11*(12), 1853–1862. https://doi.org/10.1093/scan/nsw102

Correia-Caeiro, C., Guo, K., & Mills, D. S. (2020). Perception of dynamic facial expressions of emotion between dogs and humans. *Animal Cognition*, *23*(3), 465–476. https://doi.org/10.1007/s10071-020-01348-5

Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of*

Applied Behavior Analysis, *34*(1), 101–122. https://doi.org/10.1901/jaba.2001.34-101

Cross, L., Wilson, A. D., & Golonka, S. (2016). How moving together brings us together: When coordinated rhythmic movement affects cooperation. *Frontiers in Psychology*, *7*, 1983. https://doi.org/10.3389/fpsyg.2016.01983

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. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*(1), 60–78. https://doi.org/10.1016/j.actpsy.2006.09.007

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Cutting, A. L., & Dunn, J. (2006). Conversations with siblings and with friends: Links between relationship quality and social understanding. *British Journal of Developmental Psychology, 24(1),* 73–87. https://doi.org/10.1348/026151005X70337

Dahl, A. (2015). The developing social context of infant helping in two US samples. *Child Development*, *86*(4), 1080–1093. https://doi.org/10.1111/cdev.12361

Dahl, A. (2018). How, not whether: Contributions of others in the development of infant helping. *Current Opinion in Psychology, 20,* 72–76. https://doi.org/10.1016/j.copsyc.2017.07.038

Dahl, A. (2019). The science of early moral development: On defining, constructing, and studying morality from birth. In J. Benson (Ed.), *Advances in child development and behavior* (Vol. 56, pp. 2–24). Academic Press. https://doi.org/10.1016/bs.acdb.2018.11.001

Dahl, A., & Brownell, C. A. (2019). The social origins of human prosociality. *Current Directions in Psychological Science*, 0963721419830386. https://doi.org/10.1177/0963721419830386

Dahl, A., Campos, J. J., & Witherington, D. C. (2011). Emotional action and communication in early moral development. *Emotion Review, 3*(2), 147–157. https://doi.org/10.1177/1754073910387948

Dahl, A., Gross, R. L., & Siefert, C. (2020). Young children's judgments and reasoning about prosocial acts: Impermissible, suberogatory, obligatory, or supererogatory? *Cognitive Development, 55*, 100908. https://doi.org/10.1016/j.cogdev.2020.100908

Dahl, A., & Paulus, M. (2019). From interest to obligation: The gradual development of human altruism. *Child Development Perspectives*, *13*(1), 10–14. https://doi.org/10.1111/cdep.12298

Dahl, A., Satlof-Bedrick, E. S., Hammond, S. I., Drummond, J. K., Waugh, W. E., & Brownell, C. A. (2017). Explicit scaffolding increases simple helping in younger infants. *Developmental Psychology*, *53*(3), 407–416. https://doi.org/10.1037/dev0000244

Davidov, M., Paz, Y., Roth-Hanania, R., Uzefovsky, F., Orlitsky, T., Mankuta, D., & Zahn-Waxler, C. (2020). Caring babies: Concern for others in distress during infancy. *Developmental Science*, e13016. https://doi.org/10.1111/desc.13016

Decety, J., Bartal, I. B.-A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: Highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150077. https://doi.org/10.1098/rstb.2015.0077

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners. *Psychological Science*, *27*(6), 848–858. https://doi.org/10.1177/0956797616639301

Delin, C. R., & Baumeister, R. F. (1994). Praise: More than just social reinforcement. *Journal for the Theory of Social Behaviour*, *24*(3), 219–241. https://doi.org/10.1111/j.1468-5914.1994.tb00254.x

Derrick, B., & White, P. (2017). Comparing two samples from an individual Likert question. *International Journal of Mathematics and Statistics*, *18*(3).

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dickinson, A. (1980). *Contemporary animal learning theory* (Vol. 1). CUP Archive.

Dik, G., & Aarts, H. (2007). Behavioral cues to others' motivation and goal pursuits: The perception of effort facilitates goal inference and contagion. *Journal of Experimental Social Psychology*, *43*(5), 727–737. https://doi.org/10.1016/j.jesp.2006.09.002

Doebel, S. (2019). Rethinking executive function development. *Perspectives on Psychological Science*, *preprint*. https://doi.org/10.31234/osf.io/au9rn

Drummond, J., Paul, E. F., Waugh, W. E., Hammond, S. I., & Brownell, C. A. (2014). Here, there and everywhere: emotion and mental state talk in different social contexts predicts empathic helping in toddlers. *Frontiers in Psychology*, *5*, 361. https://doi.org/10.3389/fpsyg.2014.00361

Dunfield, K. A. (2014). A construct divided: Prosocial behavior as helping, sharing, and comforting subtypes. *Frontiers in Psychology*, *5*, 958. https://doi.org/10.3389/fpsyg.2014.00958

Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, *21*(4), 523–527. https://doi.org/10.1177/0956797610364119

Dunn, L., Magalhaes, L. C., & Mancini, M. C. (2014). Internal structure of the Children Helping Out: Responsibilities, Expectations, and Supports (CHORES) measure. *American Journal of Occupational Therapy*, *68*(3), 286–295. https://doi.org/10.5014/ajot.2014.010454

Eisenberg, M. L., Zacks, J. M., & Flores, S. (2018). Dynamic prediction during perception of everyday events. *Cognitive Research: Principles and Implications*, *3*(1), 53. https://doi.org/10.1186/s41235-018-0146-z

Eisenberg, N. (1982). The development of reasoning regarding prosocial behavior. In *The development of prosocial behavior* (pp. 219–249). Academic Press.

Eisenberg, N., VanSchyndel, S. K., & Spinrad, T. L. (2016). Prosocial motivation: Inferences from an opaque body of work. *Child Development, 87*(6), 1668–1678. https://doi.org/10.1111/cdev.12638

Elsner, B., & Adam, M. (2020). Infants' goal prediction for simple action events: The role of experience and agency cues. *Topics in Cognitive Science*. https://doi.org/10.1111/tops.12494

Engelmann, J., Haux, L. M., & Herrmann, E. (2019). Helping in young children and chimpanzees shows partiality towards friends. *Evolution and Human Behavior, 40*(3), 292–300. https://doi.org/10.1016/j.evolhumbehav.2019.01.003

Engelmann, J., Over, H., Herrmann, E., & Tomasello, M. (2013). Young children care more about their reputation with ingroup members and potential reciprocators. *Developmental Science, 16*(6), 952–958. https://doi.org/10.1111/desc.12086

Engelmann, J., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology, 20*, 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychological Science, 20*(4), 438–440. https://doi.org/10.1111/j.1467-9280.2009.02317.x

Esseily, R., Rat-Fischer, L., O'Regan, K., & Fagard, J. (2013). Understanding the experimenter's intention improves 16-month-olds' observational learning of the use of a novel tool. *Cognitive Development, 28*(1), 1–9. https://doi.org/10.1016/j.cogdev.2012.10.001

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: tests for correlation and regression

analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Fay-Stammbach, T., Hawes, D. J., & Meredith, P. (2014). Parenting influences on executive function in early childhood: A review. *Child Development Perspectives*, *8*(4), 258–264. https://doi.org/10.1111/cdep.12095

Feigenson, L. (2016). Surprise enhances early learning. In D. Barner & A. S. Baron (Eds.), *Core Knowledge and Conceptual Change* (pp. 89–102). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190467630.003.0006

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low Kappa: I. the problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543–549. https://doi.org/10.1016/0895-4356(90)90158-L

Ferguson, C. J. (2013). Spanking, corporal punishment and negative long-term outcomes: A meta-analytic review of longitudinal studies. *Clinical Psychology Review, 33(1),* 196–208. https://doi.org/10.1016/j.cpr.2012.11.002

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64(3),* 657–674. https://doi.org/10.1111/j.1467-8624.1993.tb02934.x

Fischer, A. H., Rodriguez Mosquera, P. M., van Vianen, A. E. M., & Manstead, A. S. R. (2004). Gender and culture differences in emotion. *Emotion, 4(1),* 87–94. https://doi.org/10.1037/1528-3542.4.1.87

Fischer, R., Callander, R., Reddish, P., & Bulbulia, J. (2013). How do rituals affect cooperation? *Human Nature*, *24*(2), 115–125. https://doi.org/10.1007/s12110-013-9167-y

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

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

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science, 1*(2), 292–304. https://doi.org/10.1111/j.1756-8765.2009.01020.x

Geangu, E., Benga, O., Stahl, D., & Striano, T. (2010). Contagious crying beyond the first days of life. *Infant Behavior and Development, 33*(3), 279–288. https://doi.org/10.1016/j.infbeh.2010.03.004

Geeraerts, S. B., Endendijk, J. J., Deković, M., Huijding, J., Deater-Deckard, K., & Mesman, J. (2021). Inhibitory control across the preschool years: Developmental changes and associations with parenting. *Child Development, 92*(1), 335–350. https://doi.org/10.1111/cdev.13426

Geller, D., & Bamberger, P. (2009). Bringing avoidance and anxiety to the job: Attachment style and instrumental helping behavior among co-workers. *Human Relations, 62*(12), 1803–1827. https://doi.org/10.1177/0018726709337524

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature, 415*(6873), 755–756. https://doi.org/10.1038/415755a

Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: imitative learning as a mechanism of human pedagogy. *Interaction Studies, 6*(3), 463–481. https://doi.org/10.1075/is.6.3.10ger

Gerson, S. A., & Woodward, A. L. (2014). 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

Gilbert, M. (2006). *A theory of political obligation: Membership, commitment, and the bonds of society*. Clarendon. http://0-dx.doi.org.pugwash.lib.warwick.ac.uk/10.1093/0199274959.001.0001

Gilbert, M. (2018). Rights and demands: A foundational inquiry. In *Rights and Demands*. Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/oso/9780198813767.001.0001/oso-9780198813767

Godman, M., Nagatsu, M., & Salmela, M. (2014). The social motivation hypothesis for prosocial behavior. *Philosophy of the Social Sciences*, *44*(5), 563–587. https://doi.org/10.1177/0048393114530841

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, *45*(5), 1430–1443. https://doi.org/10.1037/a0016122

Gräfenhain, M., Carpenter, M., & Tomasello, M. (2013). Three-year-olds' understanding of the consequences of joint commitments. *PloS One*, *8*(9), e73039. https://doi.org/10.1371/journal.pone.0073039

Gralinski, J. H., & Kopp, C. B. (1993). Everyday rules for behavior: Mothers' requests to young children. *Developmental Psychology, 29(3),* 573–584. https://doi.org/10.1037/0012-1649.29.3.573

Green, A., McEllin, L., & Michael, J. (2019). Does sensorimotor communication stabilize commitment in joint action? *Physics of Life Reviews*, *28*, 28–30. https://doi.org/10.1016/j.plrev.2019.01.009

Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792. https://doi.org/10.1037/0033-2909.130.5.769

Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210x.12504

Green, R. T. (1956). Surprise as a factor in the von Restorff effect. *Journal of Experimental Psychology*, *52*(5), 340–344.

https://doi.org/10.1037/h0047496

Grosse, G., Moll, H., & Tomasello, M. (2010). 21-month-olds understand the cooperative logic of requests. *Journal of Pragmatics*, *42*(12), 3377–3383. https://doi.org/10.1016/j.pragma.2010.05.005

Grossmann, T. (2018). How to build a helpful baby: A look at the roots of prosociality in infancy. *Current Opinion in Psychology*, *20*, 21–24. https://doi.org/10.1016/j.copsyc.2017.08.007

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development*, *83*(1), 137–145. https://doi.org/10.1111/j.1467-8624.2011.01695.x

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557. https://doi.org/10.1203/pdr.0b013e318168c6e5

Hammond, S. I. (2014). Children's early helping in action: Piagetian developmental theory and early prosocial behavior. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00759

Hammond, S. I., & Brownell, C. A. (2018). Happily unhelpful: Infants' everyday helping and its connections to early prosocial development. *Frontiers in Psychology*, *9*, 1770. https://doi.org/10.3389/fpsyg.2018.01770

Hammond, S. I., & Carpendale, J. I. M. (2015). Helping children help: The relation between maternal scaffolding and children's early help. *Social Development*, *24*(2), 367–383. https://doi.org/10.1111/sode.12104

Hammond, S. I., & Drummond, J. K. (2019). Rethinking emotions in the context of infants' prosocial behavior: The role of interest and positive emotions. *Developmental Psychology*, *55*(9), 1882–1888. https://doi.org/10.1037/dev0000685

Han, T. A. (2013). Intention recognition, commitment and the evolution of cooperation. In T. A. Han (Ed.), *Intention Recognition, Commitment and Their Roles in the Evolution of Cooperation: From Artificial*

Intelligence Techniques to Evolutionary Game Theory Models (pp. 123–132). Springer. https://doi.org/10.1007/978-3-642-37512-5_8

Hayden, B. Y. (2016). Time discounting and time preference in animals: A critical review. *Psychonomic Bulletin & Review*, *23*(1), 39–53. https://doi.org/10.3758/s13423-015-0879-3

Heintz, C., Celse, J., Giardini, F., & Max, S. (2015). Facing expectations: Those that we prefer to fulfil and those that we disregard. *Judgment & Decision Making*, *10*(5), 442–455.

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, *128*(5), 774–795. https://doi.org/10.1037/0033-2909.128.5.774

Henderson, A., Wang, Y., Matz, L., & Woodward, A. (2013). Active experience shapes 10-month-old infants' understanding of collaborative goals. *Infancy*, *18*(1), 10–39. https://doi.org/10.1111/j.1532-7078.2012.00126.x

Henderson, A., & Woodward, A. L. (2011). "Let's work together": What do infants understand about collaborative goals? *Cognition*, *121*(1), 12–21. https://doi.org/10.1016/j.cognition.2011.05.008

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, *53*(1), 3–35. https://doi.org/10.1016/s0167-2681(03)00110-0

Henrichs, I., Elsner, C., Elsner, B., & Gredebäck, G. (2012). Goal salience affects infants' goal-directed gaze shifts. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00391

Hepach, R. (2017). Prosocial arousal in children. *Child Development Perspectives*, *11*(1), 50–55. https://doi.org/10.1111/cdep.12209

Hepach, R., Benziad, L., & Tomasello, M. (2020). Chimpanzees help others with what they want; children help them with what they need. *Developmental Science*, *23*(3), e12922. https://doi.org/10.1111/desc.12922

Hepach, R., Haberl, K., Lambert, S., & Tomasello, M. (2017). Toddlers help anonymously. *Infancy*, *22*(1), 130–145. https://doi.org/10.1111/infa.12143

Hepach, R., Hedley, D., & Nuske, H. J. (2019). Prosocial attention in children with and without autism spectrum disorder: Dissociation between anticipatory gaze and internal arousal. *Journal of Abnormal Child Psychology*. https://doi.org/10.1007/s10802-019-00606-6

Hepach, R., Vaish, A., Grossmann, T., & Tomasello, M. (2016). Young children want to see others get the help they need. *Child Development*, *87*(6), 1703–1714. https://doi.org/10.1111/cdev.12633

Hepach, R., Vaish, A., Müller, K., & Tomasello, M. (2019). Toddlers' intrinsic motivation to return help to their benefactor. *Journal of Experimental Child Psychology*, *188*, 104658. https://doi.org/10.1016/j.jecp.2019.06.011

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, *23*(9), 967–972. https://doi.org/10.1177/0956797612440571

Hepach, R., Vaish, A., & Tomasello, M. (2013). Young children sympathize less in response to unjustified emotional distress. *Developmental Psychology*, *49*(6), 1132–1138. https://doi.org/10.1037/a0029501

Hepach, R., & Westermann, G. (2013). Infants' sensitivity to the congruence of others' emotions and actions. *Journal of Experimental Child Psychology*, *115*(1), 16–29. https://doi.org/10.1016/j.jecp.2012.12.013

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*(5843), 1360–1366. https://doi.org/10.1126/science.1146282

Herrmann, E., Engelmann, J. M., & Tomasello, M. (2019). Children engage in competitive altruism. *Journal of Experimental Child Psychology*, *179*, 176–189. https://doi.org/10.1016/j.jecp.2018.11.008

Hervé, M. (2020). *RVAideMemoire: Testing and plotting procedures for*

biostatistics. R package version 0.9-75. https://CRAN.R-project.org/package=RVAideMemoire

Hobbs, K., & Spelke, E. (2015). Goal attributions and instrumental helping at 14 and 24 months of age. *Cognition*, *142*, 44–59. https://doi.org/10.1016/j.cognition.2015.03.014

Hobson, R. P. (2002). *The cradle of thought*. Pan MacMillan.

Hofferth, S. L., & Sandberg, J. F. (2001). How american children spend their time. *Journal of Marriage and Family*, *63*(2), 295–308. https://doi.org/10.1111/j.1741-3737.2001.00295.x

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. https://doi.org/10.1016/j.tics.2012.01.006

Holmes, C. J., Kim-Spoon, J., & Deater-Deckard, K. (2016). Linking executive function and peer problems from early childhood through middle adolescence. *Journal of Abnormal Child Psychology*, *44*(1), 31–42. https://doi.org/10.1007/s10802-015-0044-5

Hornstein, H. A. (1972). Promotive tension: The basis of prosocial behavior from a lewinian perspective. *Journal of Social Issues*, *28*(3), 191–218. https://doi.org/10.1111/j.1540-4560.1972.tb00039.x

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development*, *20*(3), 251. https://doi.org/10.1002/icd.736

Hughes, C., & Devine, R. T. (2019). For better or for worse? Positive and negative parental influences on young children's executive function. *Child Development*, *90*(2), 593–609. https://doi.org/10.1111/cdev.12915

Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review*, *39*(1), 25–43. https://doi.org/10.1037/h0072640

Hussar, K., & Horvath, J. (2013). But you promised: Children's judgments of broken promises. *Psychology*, *4*(12), 1046–1050. https://doi.org/10.4236/psych.2013.412152

Ichikawa, H., & Yamaguchi, M. K. (2014). Infants' recognition of subtle anger facial expression: Infants' recognition of subtle facial expression. *Japanese Psychological Research*, *56*(1), 15–23. https://doi.org/10.1111/jpr.12025

Jafari, M., & Ansari-Pour, N. (2019). Why, When and How to Adjust Your P Values? *Cell Journal (Yakhteh)*, *20*(4), 604–607. https://doi.org/10.22074/cellj.2019.5992

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14–23. https://doi.org/10.1016/j.cognition.2015.03.006

Jensen, K. (2016). Prosociality. *Current Biology*, *26*(16), R748–R752. https://doi.org/10.1016/j.cub.2016.07.025

Jensen, K., Vaish, A., & Schmidt, M. (2014). The emergence of human prosociality: Aligning with others through feelings, concerns, and norms. *Frontiers in Psychology*, *5*, 822. https://doi.org/10.3389/fpsyg.2014.00822

Jin, K., & Song, H. (2017). You changed your mind! Infants interpret a change in word as signaling a change in an agent's goals. *Journal of Experimental Child Psychology*, *162*, 149–162. https://doi.org/10.1016/j.jecp.2017.05.001

Johnson, D. B. (1982). Altruistic behavior and the development of the self in infants. *Merrill-Palmer Quarterly*, *28*(3), 379–388.

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, *7*(4), 396–403. https://doi.org/10.9734/BJAST/2015/14975

Kachel, U., Svetlova, M., & Tomasello, M. (2017). Three-year-olds' reactions to a partner's failure to perform her role in a joint commitment. *Child Development*, *89*(5), 1691–1703. https://doi.org/10.1111/cdev.12816

Kachel, U., Svetlova, M., & Tomasello, M. (2019). Three- and 5-year-old children's understanding of how to dissolve a joint commitment.

Journal of Experimental Child Psychology, *184*, 34–47. https://doi.org/10.1016/j.jecp.2019.03.008

Kachel, U., & Tomasello, M. (2019). 3- and 5-year-old children's adherence to explicit and implicit joint commitments. *Developmental Psychology*, *55*(1), 80–88. https://doi.org/10.1037/dev0000632

Kahn, J. H., & Garrison, A. M. (2009). Emotional self-disclosure and emotional avoidance: Relations with symptoms of depression and anxiety. *Journal of Counseling Psychology*, *56*(4), 573–584. https://doi.org/10.1037/a0016574

Kanngiesser, P., Köymen, B., & Tomasello, M. (2017). Young children mostly keep, and expect others to keep, their promises. *Journal of Experimental Child Psychology*, *159*, 140–158. https://doi.org/10.1016/j.jecp.2017.02.004

Kärtner, J. (2018). Beyond dichotomies—(M)others' structuring and the development of toddlers' prosocial behavior across cultures. *Current Opinion in Psychology*, *20*, 6–10. https://doi.org/10.1016/j.copsyc.2017.07.040

Kärtner, J., Keller, H., & Chaudhary, N. (2010). Cognitive and social influences on early prosocial behavior in two sociocultural contexts. *Developmental Psychology, 46(4),* 905–914. https://doi.org/10.1037/a0019718

Kärtner, J., Torréns, M. G., & Schuhmacher, N. (2020). Parental structuring during shared chores and the development of helping across the second year. *Social Development, 0(0),* 1–22. https://doi.org/10.1111/sode.12490

Kenward, B., & Gredebäck, G. (2013). Infants help a non-human agent. *PLoS One*, *8*(9), e75130. https://doi.org/10.1371/journal.pone.0075130

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, 13(6), 522–525. https://doi.org/10.1016/S0960-9822(03)00165-9

Kim, E. Y., & Song, H. (2015). Six-month-olds actively predict others'

goal-directed actions. *Cognitive Development, 33,* 1–13. https://doi.org/10.1016/j.cogdev.2014.09.003

Kim, H., & Johnson, S. (2013). Do young infants prefer an infant-directed face or a happy face? *International Journal of Behavioral Development, 0(0),* 1–6. https://doi.org/10.1177/0165025413475972

Kim, H.-Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics, 42(2),* 152–155. https://doi.org/10.5395/rde.2017.42.2.152

Király, I., Oláh, K., Csibra, G., & Kovács, Á. M. (2018). Retrospective attribution of false beliefs in 3-year-old children. *Proceedings of the National Academy of Sciences*, *115*(45), 11477–11482. https://doi.org/10.1073/pnas.1803505115

Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research*, *43*(1), 39–58. https://doi.org/10.1509/jmkr.43.1.39

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In *Psychology of learning and motivation* (Vol. 54, pp. 59–101). Elsevier.

Knudsen, B., & Liszkowski, U. (2012). Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science, 15*(1), 113–122. https://doi.org/10.1111/j.1467-7687.2011.01098.x

Knudsen, B., & Liszkowski, U. (2013). One-year-olds warn others about negative action outcomes. *Journal of Cognition and Development*, *14*(3), 424–436. https://doi.org/10.1080/15248372.2012.689387

Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment-if the rhythm comes easily. *PLoS One, 6*(11), e27272. https://doi.org/10.1371/journal.pone.0027272

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint

tapping. *Quarterly Journal of Experimental Psychology, 63*(11), 2220–2230. https://doi.org/10.1080/17470218.2010.497843

Koomen, R., Grueneisen, S., & Herrmann, E. (2020). Children delay gratification for cooperative ends. *Psychological Science, 31*(2), 139–148. https://doi.org/10.1177/0956797619894205

Köster, M., Itakura, S., Omori, M., & Kärtner, J. (2019). From understanding others' needs to prosocial action: Motor and social abilities promote infants' helping. *Developmental Science,* e12804. https://doi.org/10.1111/desc.12804

Krogh-Jespersen, S., Henderson, A. M. E., & Woodward, A. L. (2020). Let's get it together: Infants generate visual predictions based on collaborative goals. *Infant Behavior and Development*, *59*, 101446. https://doi.org/10.1016/j.infbeh.2020.101446

Krogh-Jespersen, S., Liberman, Z., & Woodward, A. L. (2015). Think fast! The relationship between goal prediction speed and social competence in infants. *Developmental Science*, *18*(5), 815–823. https://doi.org/10.1111/desc.12249

Krogh-Jespersen, S., & Woodward, A. L. (2014). Making smart social judgments 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

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition*, *40*(5), 812–826. https://doi.org/10.3758/s13421-011-0179-8

Kurby, C. A., & Zacks, J. M. (2019). Age differences in the perception of goal structure in everyday activity. *Psychology and Aging*, *34*(2), 187–201. https://doi.org/10.1037/pag0000321

Laible, D., & Thompson, R. A. (2007). Early socialization: A relationship perspective. In J. Grusec & P. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 181–207). The Guilford Press.

Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology*, *127*, 191–197. https://doi.org/10.1016/j.biopsycho.2017.06.001

Laurin, K. (2016). Interpersonal influences on goals: Current and future directions for goal contagion research. *Social and Personality Psychology Compass*, *10*(11), 668–678. https://doi.org/10.1111/spc3.12289

Leimgruber, K. L. (2018). The developmental emergence of direct reciprocity and its influence on prosocial behavior. *Current Opinion in Psychology*, *20*, 122–126. https://doi.org/10.1016/j.copsyc.2018.01.006

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS ONE*, *7*(10), e48292. https://doi.org/10.1371/journal.pone.0048292

Leonard, J., Lee, Y., & Schulz, L. (2017). Infants make more attempts to achieve a goal when they see adults persist. *Cognitive Development*, *357*(6357), 1290–1294. https://doi.org/10.1126/science.aan2317

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers (Edited by Dorwin Cartwright.)*. Harpers.

Lewis, H. B. (1944). An experimental study of the role of the ego in work: I, the role of the ego in cooperative work. *Journal of Experimental Psychology*, *34*(2), 113–126. https://doi.org/10.1037/h0054102

Lewis, H. B., & Franklin, M. (1944). An experimental study of the role of the ego in work: II, the significance of task-orientation in work. *Journal of Experimental Psychology*, *34*(3), 195–215. https://doi.org/10.1037/h0053522

Lewis, C., & Carpendale, J. I. M. (2009). Introduction: Links between social interaction and executive function. *New Directions for Child and Adolescent Development*, *2009*(123), 1–15. https://doi.org/10.1002/cd.232

Liebal, K., Colombi, C., Rogers, S. J., Warneken, F., & Tomasello, M. (2008). Helping and cooperation in children with autism. *Journal of Autism and Developmental Disorders*, *38(2)*, 224–238. https://doi.org/10.1007/s10803-007-0381-5

Liepelt, R., Von Cramon, D. Y., & Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *NeuroImage*, *43*(4), 784–792. https://doi.org/10.1016/j.neuroimage.2008.08.007

Linderholm, T., Gernsbacher, M. A., van den Broek, P., Neninde, L., Robertson, R. R. W., & Sundermier, B. (2004). Suppression of story character goals during reading. *Discourse Processes*, *37*(1), 67–78. https://doi.org/10.1207/s15326950dp3701_4

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition*, *160*, 35–42. https://doi.org/10.1016/j.cognition.2016.12.007

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041. https://doi.org/10.1126/science.aag2132

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476–490. https://doi.org/10.3758/bf03210951

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*(2), 167–204. https://doi.org/10.1016/j.cognition.2004.12.009

Loucks, J., Mutschler, C., & Meltzoff, A. N. (2017). Children's representation and imitation of events: How goal organization

influences 3-year-old children's memory for action sequences. *Cognitive Science*, *41*(7), 1904–1933. https://doi.org/10.1111/cogs.12446

Lucca, K., Horton, R., & Sommerville, J. A. (2019). Keep trying!: Parental language predicts infants' persistence. *Cognition*, *193*, 104025. https://doi.org/10.1016/j.cognition.2019.104025

Ludwig, K. (2007). Collective intentional behavior from the standpoint of semantics. *Nous*, *41*(3), 355–393. https://doi.org/10.1111/j.1468-0068.2007.00652.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

Luo, Y., Hennefield, L., Mou, Y., van Marle, K., & Markson, L. (2017). Infants' understanding of preferences when agents make inconsistent choices. *Infancy*, *22*(6), 843–856. https://doi.org/10.1111/infa.12194

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology*, *58*(4), 441–486. https://doi.org/10.1016/j.cogpsych.2008.11.001

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. *Journal of Memory and Language*, *36*(2), 293–310. https://doi.org/10.1006/jmla.1996.2491

Magliano, J. P., & Radvansky, G. A. (2001). Goal coordination in narrative comprehension. *Psychonomic Bulletin & Review*, *8*(2), 372–376. https://doi.org/10.3758/BF03196175

Magliano, J. P., Radvansky, G. A., Forsythe, J. C., & Copeland, D. E. (2014). Event segmentation during first-person continuous events. *Journal of Cognitive Psychology*, *26*(6), 649–661.

https://doi.org/10.1080/20445911.2014.930042

Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The other half of the story: Effect size analysis in quantitative research. *CBE—Life Sciences Education*, 12(3), 345–351. https://doi.org/10.1187/cbe.13-04-0082

Mangiafico, S. (2016). Summary and analysis of extension program evaluation in R. *New Brunswick, NJ: Rutgers*, *442*. Retrieved from: http://rcompanion.org/handbook/A_02.html (21/08/2018).

Mangiafico, S. (2018). *rcompanion: Functions to support extension education program evaluation.* R package version 2.0.0. https://CRAN.R-project.org/package=rcompanion.

Mant, C., & Perner, J. (1988). The child's understanding of commitment. *Developmental Psychology*, *24*(3), 343–351. https://doi.org/10.1037/0012-1649.24.3.343

Marshall, D., Stuart, M., & Bell, R. (2006). Examining the relationship between product package colour and product selection in preschoolers. *Food Quality and Preference*, *17*(7–8), 615–621. https://doi.org/10.1016/j.foodqual.2006.05.007

Martin, A., & Olson, K. R. (2013). When kids know better: Paternalistic helping in 3-year-old children. *Developmental Psychology*, *49*(11), 2071–2081. https://doi.org/10.1037/a0031715

Martin, A., & Olson, K. R. (2015). Beyond good and evil: What motivations underlie children's prosocial behavior? *Perspectives on Psychological Science*, *10*(2), 159–175. https://doi.org/10.1177/1745691615568998

Martin, A., Shelton, C. C., & Sommerville, J. A. (2017). Once a frog-lover, always a frog-lover? Infants' goal generalization is influenced by the nature of accompanying speech. *Journal of Experimental Psychology: General*, *146*(6), 859–871. https://doi.org/10.1037/xge0000268

Martin, D. U., Perry, C., MacIntyre, M. I., Varcoe, L., Pedell, S., & Kaufman, J. (2020). Investigating the nature of children's altruism

using a social humanoid robot. *Computers in Human Behavior*, *104*, 106149. https://doi.org/10.1016/j.chb.2019.09.025

Matsumoto, D., & Hwang, H. S. C. (2019). Culture, Emotion, and Expression. In *Cross-Cultural* Psychology (pp. 501–515). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119519348.ch24

McAlister, A. R., & Peterson, C. C. (2013). Siblings, theory of mind, and executive functioning in children aged 3–6 years: New longitudinal evidence. *Child Development*, *84*(4), 1442–1458. https://doi.org/10.1111/cdev.12043

McCulloch, K. C., Fitzsimons, G. M., Chua, S. N., & Albarracín, D. (2011). Vicarious goal satiation. *Journal of Experimental Social Psychology*, *47*(3), 685–688. https://doi.org/10.1016/j.jesp.2010.12.019

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 51, pp. 297–384). Elsevier Science.

Melis, A. P. (2018). The evolutionary roots of prosociality: The case of instrumental helping. *Current Opinion in Psychology*, *20*, 82–86. https://doi.org/10.1016/j.copsyc.2017.08.019

Melis, A. P., & Semmann, D. (2010). How is human cooperation different? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1553), 2663–2674. https://doi.org/10.1098/rstb.2010.0157

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*(5), 838–850. https://doi.org/10.1037//0012-1649.31.5.838

Mermelshtine, R. (2017). Parent–child learning interactions: A review of the literature on scaffolding. *British Journal of Educational Psychology*, *87*(2), 241–254. https://doi.org/10.1111/bjep.12147

Meyer, M., Bekkering, H., Haartsen, R., Stapel, J. C., & Hunnius, S. (2015). The role of action prediction and inhibitory control for joint action

coordination in toddlers. *Journal of Experimental Child Psychology, 139*, 203–220. https://doi.org/10.1016/j.jecp.2015.06.005

Meyer, M., van der Wel, R. P., & Hunnius, S. (2016). Planning my actions to accommodate yours: Joint action development during early childhood. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1693), 20150371. https://doi.org/10.1098/rstb.2015.0371

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

Michael, J., Knoblich, G., & Sebanz, N. (2016). The sense of commitment: A minimal approach. *Frontiers in Psychology, 6*(1968), 1–11. https://doi.org/10.3389/fpsyg.2015.01968

Michael, J., & Pacherie, E. (2015). On commitments and other uncertainty reduction tools in joint action. *Journal of Social Ontology, 1*(1), 89–120. https://doi.org/10.1515/jso-2014-0021

Michael, J., & Salice, A. (2017). The sense of commitment in human–robot interaction. *International Journal of Social Robotics, 9*(5), 755–763. https://doi.org/10.1007/s12369-016-0376-5

Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition, 157*, 106–113. https://doi.org/10.1016/j.cognition.2016.08.024

Michael, J., & Székely, M. (2018). The developmental origins of commitment. *Journal of Social Philosophy, 49*(1), 106–123. https://doi.org/10.1111/josp.12220

Michael, J., & Székely, M. (2019). Goal slippage: A mechanism for spontaneous instrumental helping in infancy? *Topoi, 38*(1), 173–183. https://doi.org/10.1007/s11245-017-9485-5

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review, 126*(2), 292–311. https://doi.org/10.1037/rev0000120

Milward, S. J., Kita, S., & Apperly, I. A. (2014). The development of co-representation effects in a joint task: Do children represent a co-actor? *Cognition, 132(3),* 269–279. https://doi.org/10.1016/j.cognition.2014.04.008

Miyake, A., Friedman, N., Emerson, M., Witzki, A., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Moore, C., & Barresi, J. (2017). The role of second-person information in the development of social understanding. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01667

Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Reason*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Morgan, G. A., Harmon, R. J., & Maslin-Cole, C. A. (1990). Mastery motivation: Definition and measurement. *Early Education and Development*, *1*(5), 318–339. https://doi.org/10.1207/s15566935eed0105_1

Moriguchi, Y. (2014). The early development of executive function and its relation to social interaction: A brief review. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00388

Moss, J., Kotovsky, K., & Cagan, J. (2007). The influence of open goals on the acquisition of problem-relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 876–891. https://doi.org/10.1037/0278-7393.33.5.876

Müller, C. A., Schmitt, K., Barber, A. L. A., & Huber, L. (2015). Dogs can discriminate emotional expressions of human faces. *Current Biology*, *25*(5), 601–605. https://doi.org/10.1016/j.cub.2014.12.055

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science, 31(6),* 678–701. https://doi.org/10.1177/0956797620916782

Myowa-Yamakoshi, M., Scola, C., & Hirata, S. (2012). Humans and chimpanzees attend differently to goal-directed actions. *Nature Communications, 3(1),* 693. https://doi.org/10.1038/ncomms1695

Nakazawa, M. (2019). *fmsb: Functions for medical statistics book with some demographic data*. R package version 0.7.0. https://CRAN.R-project.org/package=fmsb

Neale, D., & Whitebread, D. (2019). Maternal scaffolding during play with 12- to 24-month-old infants: Stability over time and relations with emerging effortful control. *Metacognition and Learning*, *14*(3), 265–289. https://doi.org/10.1007/s11409-019-09196-6

Nelson, C. A., & Dolgin, K. G. (1985). The generalized discrimination of facial expressions by seven-month-old infants. *Child Development, 56*(1), 58–61. JSTOR. https://doi.org/10.2307/1130173

Nelson, C. A., Morse, P. A., & Leavitt, L. A. (1979). Recognition of facial expressions by seven-month-old infants. *Child Development*, *50*(4), 1239–1242. JSTOR. https://doi.org/10.2307/1129358

Newton, E. K., Goodman, M., & Thompson, R. A. (2014). Why do some toddlers help a stranger? Origins of individual differences in prosocial behavior. *Infancy*, *19*(2), 214–226. https://doi.org/10.1111/infa.12043

Nichols, S. R., Svetlova, M., & Brownell, C. A. (2015). Toddlers' responses to infants' negative emotions. *Infancy*, *20*(1), 70–97. https://doi.org/10.1111/infa.12066

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*(4), 361–383. https://doi.org/10.1111/jcpp.12675

Odum, A. L., & Rainaud, C. P. (2003). Discounting of delayed hypothetical money, alcohol, and food. *Behavioural Processes*, *64*(3), 305–313. https://doi.org/10.1016/S0376-6357(03)00145-1

Ogren, M., Burling, J. M., & Johnson, S. P. (2018). Family expressiveness relates to happy emotion matching among 9-month-old infants. *Journal of Experimental Child Psychology*, *174*, 29–40. https://doi.org/10.1016/j.jecp.2018.05.003

Ogren, M., & Johnson, S. P. (2020). Factors facilitating early emotion understanding development: Contributions to individual differences. *Human Development, 64(3)*, 108–118. https://doi.org/10.1159/000511628

Ogren, M., & Johnson, S. P. (2021). Primary caregiver emotional expressiveness relates to toddler emotion understanding. *Infant Behavior and Development, 62*, 101508. https://doi.org/10.1016/j.infbeh.2020.101508

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130(3), 466–478. https://doi.org/10.1037//0096-3445.130.3.466

Over, H. (2016). The origins of belonging: Social motivation in infants and young children. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150072. https://doi.org/10.1098/rstb.2015.0072

Over, H., & Carpenter, M. (2009). Eighteen-month-old infants show increased helping following priming with affiliation. *Psychological Science*, *20*(10), 1189–1193. https://doi.org/10.1111/j.1467-9280.2009.02419.x

Ovsiankina, M. (1928). Die Wiederaufnahme von untersbrochenen Handlungen. *Psychologische Forschung, 11*, 302–379.

Paulus, M. (2014). The emergence of prosocial behavior: Why do infants and toddlers help, comfort, and share? *Child Development Perspectives*, *8*(2), 77–81. https://doi.org/10.1111/cdep.12066

Paulus, M. (2018). The multidimensional nature of early prosocial behavior: A motivational perspective. *Current Opinion in Psychology, 20*, 111–116. https://doi.org/10.1016/j.copsyc.2017.09.003

Paulus, M. (2019). Is young children's helping affected by helpees' need? Preschoolers, but not infants selectively help needy others. *Psychological Research*, 1–11. https://doi.org/10.1007/s00426-019-01148-8

Paulus, M., & Moore, C. (2012). Producing and understanding prosocial actions in early childhood. In J. B. Benson (Ed.), *Advances in Child Development and Behavior* (Vol. 42, pp. 271–305). JAI. https://doi.org/10.1016/B978-0-12-394388-0.00008-3

Paulus, M., Wörle, M., & Christner, N. (2020). The emergence of human altruism: Preschool children develop a norm for empathy-based comforting. *Journal of Cognition and Development*, *21*(1), 104–124. https://doi.org/10.1080/15248372.2019.1693375

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Pérez-Manrique, A., & Gomila, A. (2019). Bottlenose dolphins do not behave prosocially in an instrumental helping task. *Behavioural Processes*, *164*, 54–58. https://doi.org/10.1016/j.beproc.2019.04.014

Perry, R. E., Braren, S. H., Rincón-Cortés, M., Brandes-Aitken, A. N., Chopra, D., Opendak, M., Alberini, C. M., Sullivan, R. M., & Blair, C. (2019). Enhancing executive functions through social interactions: Causal evidence using a cross-species model. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02472

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166. https://doi.org/10.1016/j.tics.2009.01.006

Pettit, P., & Schweikard, D. (2006). Joint actions and group agents. *Philosophy of the Social Sciences*, *36*(18), 18–39. https://doi.org/10.1177/0048393105284169

Pettygrove, D. M., Hammond, S. I., Karahuta, E. L., Waugh, W. E., & Brownell, C. A. (2013). From cleaning up to helping out: Parental socialization and children's early prosocial behavior. *Infant Behavior and Development*, *36*(4), 843–846. https://doi.org/10.1016/j.infbeh.2013.09.005

Pezzulo, G., & Dindo, H. (2011). What should I do next? Using shared representations to solve interaction problems. *Experimental Brain Research*, *211*(3–4), 613–630. https://doi.org/10.1007/s00221-011-2712-1

Pezzulo, G., Donnarumma, F., Dindo, H., D'Ausilio, A., Konvalinka, I., & Castelfranchi, C. (2018). The body talks: Sensorimotor communication and its brain and kinematic signatures. *Physics of Life Reviews*. https://doi.org/10.1016/j.plrev.2018.06.014

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*(1), 53–78. https://doi.org/10.1016/s0010-0277(02)00073-2

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). Young children show the bystander effect in helping situations. *Psychological Science*, *26*(4), 499–506. https://doi.org/10.1177/0956797615569579

Posada, G., Jacobs, A., Richmond, M. K., Carbonell, O. A., Alzate, G., Bustamante, M. R., & Quiceno, J. (2002). Maternal caregiving and infant security in two cultures. *Developmental Psychology, 38(1),* 67–78. https://doi.org/10.1037/0012-1649.38.1.67

Powell, H., & Michael, J. (2019). Feeling committed to a robot: Why, what, when and how? *Philosophical Transactions of the Royal Society B, 374*(1771), 20180039. https://doi.org/10.1098/rstb.2018.0039

Preston, S. D., & De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/s0140525x02000018

Qu, L. (2011). Two is better than one, but mine is better than ours: Preschoolers' executive function during co-play. *Journal of Experimental Child Psychology*, *108*(3), 549–566. https://doi.org/10.1016/j.jecp.2010.08.010

R Development Core Team. (2018). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rachlin, H. (2016). Self-control based on soft commitment. *The Behavior Analyst*, *39*(2), 259–268. https://doi.org/10.1007/s40614-016-0054-9

Radvansky, G. A., & Curiel, J. M. (1998). Narrative comprehension and aging: The fate of completed goal information. *Psychology and Aging*, *13*(1), 69. https://doi.org/10.1037/0882-7974.13.1.69

Recchia, H. E., & Howe, N. (2009). Associations between social understanding, sibling relationship quality, and siblings' conflict strategies and outcomes. *Child Development, 80(5),* 1564–1578. https://doi.org/10.1111/j.1467-8624.2009.01351.x

Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized prosociality? *Religion, Brain & Behavior*, *4*(1), 3–19. https://doi.org/10.1080/2153599x.2013.764545

Reddy, V. (2008). *How infants know minds*. Harvard University Press.

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science*, *11*(1), 50–74. https://doi.org/10.1111/tops.12292

Reschke, P. J., Walle, E. A., Flom, R., & Guenther, D. (2017). Twelve-month-old infants' sensitivity to others' emotions following positive

and negative events. *Infancy*, *22*(6), 874–881. https://doi.org/10.1111/infa.12193

Revelle, W. (2018). *psych: Procedures for personality and psychological research.* Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.4.

Reynolds, B. (2006). A review of delay-discounting research with humans: Relations to drug use and gambling. *Behavioural Pharmacology*, *17*(8), 651. https://doi.org/10.1097/FBP.0b013e3280115f99

Rhoades, B. L., Greenberg, M. T., Lanza, S. T., & Blair, C. (2011). Demographic and familial predictors of early executive function development: Contribution of a person-centered perspective. *Journal of Experimental Child Psychology*, *108*(3), 638–662. https://doi.org/10.1016/j.jecp.2010.08.004

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: A basic principle of brain function. *Nature Reviews Neuroscience*, 17(12), 757–765. https://doi.org/10.1038/nrn.2016.135

Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, *70*(4), 901–908. https://doi.org/10.1016/j.anbehav.2005.02.006

Rogoff, B. (2003). *The cultural nature of human development* (Vol. 42). Oxford University Press.

Rolan, E. P., Schmitt, S. A., Purpura, D. J., & Nichols, D. L. (2018). Sibling presence, executive function, and the role of parenting. *Infant and Child Development*, *27*(4), e2091. https://doi.org/10.1002/icd.2091

Rooney, N. J., & Cowan, S. (2011). Training methods and owner–dog interactions: Links with dog behaviour and learning ability. *Applied Animal Behaviour Science*, *132*(3), 169–177. https://doi.org/10.1016/j.applanim.2011.03.007

Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, *34*(3), 447–458. https://doi.org/10.1016/j.infbeh.2011.04.007

Rothermund, K. (2003). Automatic vigilance for task-related information: Perseverance after failure and inhibition after success. *Memory & Cognition*, *31*(3), 343–352. https://doi.org/10.3758/BF03194392

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2019). How do you feel? Preverbal infants match negative emotions to events. *Developmental Psychology*, *55*(6), 1138–1149. https://doi.org/10.1037/dev0000711

Ruffman, T., Then, R., Cheng, C., & Imuta, K. (2019). Lifespan differences in emotional contagion while watching emotion-eliciting videos. *PLOS ONE, 14(1),* e0209253. https://doi.org/10.1371/journal.pone.0209253

Rusch, H., & Luetge, C. (2016). Spillovers from coordination to cooperation: Evidence for the interdependence hypothesis? *Evolutionary Behavioral Sciences*, *10*(4), 284–296. https://doi.org/10.1037/ebs0000066

Saby, J. N., Bouquet, C. A., & Marshall, P. J. (2014). Young children co-represent a partner's task: Evidence for a joint Simon effect in five-year-olds. *Cognitive Development, 32,* 38–45. https://doi.org/10.1016/j.cogdev.2014.08.001

Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., & Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Experimental Brain Research*, *226*(4), 473–486. https://doi.org/10.1007/s00221-013-3459-7

Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task constraints in joint action. *Journal of Experimental Psychology: Human Perception and Performance, 43(8),* 1480–1493. https://doi.org/10.1037/xhp0000403

Schrift, R. Y., & Parker, J. R. (2014). Staying the course: The option of doing nothing and its impact on postchoice persistence. *Psychological Science*, *25*(3), 772–780. https://doi.org/10.1177/0956797613516801

Schuhmacher, N., Köster, M., & Kärtner, J. (2018). Modeling prosocial behavior increases helping in 16-month-olds. *Child Development, 0*(0). https://doi.org/10.1111/cdev.13054

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88(3),* B11–B21. https://doi.org/10.1016/S0010-0277(03)00043-X

Senehi, N., & Brophy-Herb, H. E. (2020). Role of maternal affect and regulatory strategies in toddlers' emotion and behavior regulation. *Infant Behavior and Development, 60,* 101472. https://doi.org/10.1016/j.infbeh.2020.101472

Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4- to 6-month-old infants. *Developmental Psychobiology, 25(6),* 411–425. https://doi.org/10.1002/dev.420250603

Serrano, J. M., Iglesias, J., & Loeches, A. (1995). Infants' responses to adult static facial expressions. *Infant Behavior and Development, 18(4),* 477–482. https://doi.org/10.1016/0163-6383(95)90036-5

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology*, *38*(4), 368–383. https://doi.org/10.1016/s0022-1031(02)00005-7

Siegel, E., & Rachlin, H. (1995). Soft commitment: Self-control achieved by response persistencee. *Journal of the Experimental Analysis of Behavior*, *64*(2), 117–128. https://doi.org/10.1901/jeab.1995.64-117

Simons, D. A., & Wurtele, S. K. (2010). Relationships between parents' use of corporal punishment and their children's endorsement of spanking and hitting other children. *Child Abuse & Neglect, 34(9),* 639–646. https://doi.org/10.1016/j.chiabu.2010.01.012

Singh, L., Morgan, J., & Best, C. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3(3),* 365–394. https://doi.org/10.1207/S15327078IN0303_5

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of factorial experiments*. R package version 0.27-2. https://CRAN.R-project.org/package=afex

Singmann, H., & Kellen, D. (2019). An Introduction to Mixed Models for Experimental Psychology. In D. Spieler & E. Schumacher (Eds.), *New Methods in Cognitive Psychology* (1st ed., pp. 4–31). Routledge. https://doi.org/10.4324/9780429318405-2

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, *189*, 260–274. https://doi.org/10.1016/j.cognition.2019.03.019

Siposova, B., Tomasello, M., & Carpenter, M. (2018). Communicative eye contact signals a commitment to cooperate for young children. *Cognition*, *179*, 192–201. https://doi.org/10.1016/j.cognition.2018.06.010

Skerry, A. E., Sheskin, M., & Santos, L. R. (2011). Capuchin monkeys are not prosocial in an instrumental helping task. *Animal Cognition*, *14*(5), 647. https://doi.org/10.1007/s10071-011-0399-0

Skerry, A. E., & Spelke, E. S. (2014). Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition*, *130*(2), 204–216. https://doi.org/10.1016/j.cognition.2013.11.002

Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to goal-directed action. *Journal of Cognition and Development*, *5*(3), 289–316. https://doi.org/10.1207/s15327647jcd0503_1

Sommerville, J. A., & Crane, C. C. (2009). Ten-month-old infants use prior information to identify an actor's goal. *Developmental Science*, *12*(2), 314–325. https://doi.org/10.1111/j.1467-7687.2008.00787.x

Sommerville, J. A., Enright, E. A., Horton, R. O., Lucca, K., Sitch, M. J., & Kirchner-Adelhart, S. (2018). Infants' prosocial behavior is governed by cost-benefit analyses. *Cognition*, *177*, 12–20. https://doi.org/10.1016/j.cognition.2018.03.021

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

Song, H., Baillargeon, R., & Fisher, C. (2014). The development of infants' use of novel verbal information when reasoning about others' actions. *PLOS ONE*, *9*(3), e92387. https://doi.org/10.1371/journal.pone.0092387

Sorenson, E. R. (1975). Culture and the Expression of Emotion. In T. R. Williams (Ed.), *Psychological Anthropology* (pp. 361–372). De Gruyter Mouton. https://www.degruyter.com/document/doi/10.1515/9783110802818.361/html

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

Southgate, V., & Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition*, *130*(1), 1–10. https://doi.org/10.1016/j.cognition.2013.08.008

Spaepen, E., & Spelke, E. (2007). Will any doll do? 12-month-olds' reasoning about goal objects. *Cognitive Psychology*, *54*(2), 133–154. https://doi.org/10.1016/j.cogpsych.2006.06.001

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, *18*(5), 449–455. https://doi.org/10.1111/j.1467-9280.2007.01920.x

Spinrad, T. L., & Gal, D. E. (2018). Fostering prosocial behavior and empathy in young children. *Current Opinion in Psychology*, *20*, 40–44. https://doi.org/10.1016/j.copsyc.2017.08.004

Stajković, A., & Sergent, K. (2019). *Cognitive automation and organizational psychology: Priming goals as a new source of competitive advantage*. Routledge. https://doi.org/10.4324/9780429295881

Staub, E. (1978). Predicting prosocial behavior: A model for specifying the nature of personality-situation interaction. In *Perspectives in interactional psychology* (pp. 87–110). Springer.

Stevenson, M. K. (1986). A discounting model for decisions with delayed positive or negative outcomes. *Journal of Experimental Psychology: General*, *115*(2), 131–154. https://doi.org/10.1037/0096-3445.115.2.131

Steel, P., & König, C. J. (2006). Integrating theories of motivation. *Academy of Management Review*, *31*(4), 889–913. https://doi.org/10.5465/amr.2006.22527462

Steel, P., Svartdal, F., Thundiyil, T., & Brothen, T. (2018). Examining procrastination across multiple goal stages: A longitudinal study of temporal motivation theory. *Frontiers in Psychology*, *9*, 327. https://doi.org/10.3389/fpsyg.2018.00327

Story, G., Vlaev, I., Seymour, B., Darzi, A., & Dolan, R. (2014). Does temporal discounting explain unhealthy behavior? A systematic review and reinforcement learning perspective. *Frontiers in Behavioral Neuroscience*, *8*. https://doi.org/10.3389/fnbeh.2014.00076

Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language*, *32*(3), 279–300. https://doi.org/10.1006/jmla.1993.1015

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from likert-type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18

Svetlova, M., Nichols, S., & Brownell, C. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child*

Development, *81*(6), 1814–1827. https://doi.org/10.1111/j.1467-8624.2010.01512.x

Székely, M., & Michael, J. (2018). Investing in commitment: Persistence in a joint action is enhanced by the perception of a partner's effort. *Cognition*, *174*, 37–42. https://doi.org/10.1016/j.cognition.2018.01.012

Székely, M., Powell, H., Vannucci, F., Rea, F., Sciutti, A., & Michael, J. (2019). The perception of a robot partner's effort elicits a sense of commitment to human-robot interaction. *Interaction Studies*, *20*(2), 234–255. https://doi.org/10.1075/is.18001.sze

Sznycer, D., Delton, A. W., Robertson, T. E., Cosmides, L., & Tooby, J. (2019). The ecological rationality of helping others: Potential helpers integrate cues of recipients' need and willingness to sacrifice. *Evolution and Human Behavior, 40(1),* 34–45. https://doi.org/10.1016/j.evolhumbehav.2018.07.005

Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific Reports*, *8*(1), 9534. https://doi.org/10.1038/s41598-018-27804-4

Tennie, C., Jensen, K., & Call, J. (2016). The nature of prosociality in chimpanzees. *Nature Communications*, *7*, 13915. https://doi.org/10.1038/ncomms13915

Thompson, R. A., & Newton, E. K. (2013). Baby altruists? Examining the complexity of prosocial motivation in young children. *Infancy*, *18*(1), 120–133. https://doi.org/10.1111/j.1532-7078.2012.00139.x

Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.

Tomasello, M. (2019). The moral psychology of obligation. *The Behavioral and Brain Sciences*, 1–33. https://doi.org/10.1017/S0140525X19001742

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, *64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, *16*(1–2), 3–34. https://doi.org/10.1080/01638539309544827

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Developmental Psychology*, *45*(2), 534–543. https://doi.org/10.1037/a0014322

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Vaish, A., & Hepach, R. (2019). The development of prosocial emotions. *Emotion Review*, 175407391988501. https://doi.org/10.1177/1754073919885014

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science*, *7*(3), 261–269. https://doi.org/10.1111/j.1467-7687.2004.00344.x

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

van Buren, B., & Scholl, B. J. (2017). Minds in motion in memory: Enhanced spatial memory driven by the perceived animacy of simple shapes. *Cognition*, *163*, 87–92. https://doi.org/10.1016/j.cognition.2017.02.006

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, *80*(1), 234–248. https://www.jstor.org/stable/2006745

van Lier, P. A. C., & Deater-Deckard, K. (2016). Children's elementary school social experience and executive functions development:

Introduction to a special section. *Journal of Abnormal Child Psychology*, *44*(1), 1–6. https://doi.org/10.1007/s10802-015-0113-9

Velleman, J. (1997). How to share an intention. *Philosophy and Phenomenological Research*, *57*(1), 29–50. https://doi.org/10.2307/2953776

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, *23*(8–9), 998–1003. https://doi.org/10.1016/j.neunet.2010.06.002

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37(5), 360–363.

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

Walle, E. A., Reschke, P. J., Camras, L. A., & Campos, J. J. (2017). Infant differential behavioral responding to discrete emotions. *Emotion (Washington, D.C.)*, *17*(7), 1078–1091. https://doi.org/10.1037/emo0000307

Wan, Y., Fu, H., & Tanenhaus, M. K. (2018). Effects of coordination and gender on prosocial behavior in 4-year-old Chinese children. *Psychonomic Bulletin & Review*, *26*(2), 685–692. https://doi.org/10.3758/s13423-018-1549-z

Warneken, F. (2013). Young children proactively remedy unnoticed accidents. *Cognition*, *126*(1), 101–108. https://doi.org/10.1016/j.cognition.2012.09.011

Warneken, F. (2015). Precocious prosociality: Why do young children help? *Child Development Perspectives*, *9*(1), 1–6. https://doi.org/10.1111/cdep.12101

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental Science*, *15*(1), 54–61. https://doi.org/10.1111/j.1467-7687.2011.01107.x

Warneken, F., Hare, B., Melis, A., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biology, 5*(7), e184. https://doi.org/10.1371/journal.pbio.0050184

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

Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology, 44*(6), 1785–1788. https://doi.org/10.1037/a0013860

Warneken, F., & Tomasello, M. (2009). The roots of human altruism. *British Journal of Psychology*, *100*(3), 455–471. https://doi.org/10.1348/000712608X379061

Warneken, F., & Tomasello, M. (2013). Parental presence and encouragement do not influence helping in young children. *Infancy*, *18*(3), 345–368. https://doi.org/10.1111/j.1532-7078.2012.00120.x

Wasserman, E. A., & Castro, L. (2005). Surprise and change: Variations in the strength of present and absent cues in causal learning. *Animal Learning & Behavior*, *33*(2), 131–146. https://doi.org/10.3758/BF03196058

Waugh, W. E., & Brownell, C. A. (2017). "Help yourself!" What can toddlers' helping failures tell us about the development of prosocial behavior? *Infancy*, *22*(5), 665–680. https://doi.org/10.1111/infa.12189

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Willits, F., Theodori, G., & Luloff, A. (2016). Another look at Likert Scales. *Journal of Rural Social Sciences*, *31*(3). https://egrove.olemiss.edu/jrss/vol31/iss3/6

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.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. (2009). Infants' grasp of others' intentions. *Current directions in Psychological Science, 18(1),* 53-57. https://doi.org/10.1111/j.1467-8721.2009.01605.x

Wu, H., & Leung, S.-O. (2017). Can Likert scales be treated as interval scales?—A simulation study. *Journal of Social Service Research, 43*(4), 527–532. https://doi.org/10.1080/01488376.2017.1329775

Yamamoto, S., Humle, T., & Tanaka, M. (2009). Chimpanzees help each other upon request. *PLOS ONE*, *4*(10), e7416. https://doi.org/10.1371/journal.pone.0007416

Yamamoto, S., Humle, T., & Tanaka, M. (2012). Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals. *Proceedings of the National Academy of Sciences*, *109*(9), 3588–3592. https://doi.org/10.1073/pnas.1108517109

Yarnold, P. R. (2016). ODA vs. π and κ: Paradoxes of kappa. *Optimal Data Analysis*, 5, 160–161.

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*(1), 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28*(1), 126–136. https://doi.org/10.1037/0012-1649.28.1.126

Zeigarnik, B. (1927). Uber den Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung*, *9*(1), 1–8.

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*(4), 1028–1034. https://doi.org/10.3758/s13423-015-

0864-x

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123(2),* 162-185. https://doi.org/10.1037//0033-2909.123.2.162