

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


The Young Observer: Children's Third-Party Inferences About Social Relationships

Citation

Afshordi, Narges. 2018. The Young Observer: Children's Third-Party Inferences About Social Relationships. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:41121217

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at
https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

The Young Observer:

Children's Third-party Inferences about Social Relationships

A dissertation presented

by

Narges Afshordi

to

The Department of Psychology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Psychology

Harvard University

Cambridge, Massachusetts

June 2018

© 2018 Narges Afshordi

All rights reserved.

Dissertation advisor: Professor Susan Carey Narges Afshordi

The Young Observer:

Children's Third-party Inferences about Social Relationships

Abstract

Human social life is dependent on and structured by the many social relationships between individuals. Young and old, we are engaged in relationships both as first-party participants and as third-party observers. The developmental study of how we think about relationships is the topic of the current dissertation. The three papers provide insight into children's conceptualization and reasoning about affiliative and hierarchical relationships from different angles.

Paper 1 examined children's third-party inferences about friendship and preferences from reported dyadic information. Four-year-olds inferred friendship between individuals based on reports of joint activities, and reports of similarity, but not based on arbitrary links. Children also privileged joint activities over similarity when asked to adjudicate between the two. Further, four-year-olds expected individuals who had engaged in a joint activity to be playmates as well as friends, and to also share preferences for novel games, but not novel foods. These findings shed crucial light on preschoolers' ability to infer friendship between others, and add to our knowledge of their concept of friendship.

Paper 2 tested preschoolers' inferences about affiliation and continued imitation of the same target based on observed imitation. We found that even three-year-olds have an explicit concept of imitation, but that the abilities of three- to five-year-olds in making spontaneous inferences from imitation are limited. However, performance improves significantly when

imitation is highlighted for children, either through drawing their attention to the fact that the imitator is copying the target (propositional scaffolding) or by priming the concept of imitation ahead of the task. These findings provide some answers to the puzzling discontinuity between previously reported success in infancy on the one hand and failure at age four on the other.

Paper 3 changes focus to hierarchical relationships, tackling children's inferential abilities with regard to the cues and consequences of dominance and prestige. We compare participants in the UK with those in China, where ethnographic evidence reveals strong norms about prestigious individuals yielding to others, even those lower in status. We found that children around age six distinguished between cues to dominance and prestige in both countries. Importantly, we also found empirical evidence for a cultural difference. The norm surrounding prestige in Chinese culture manifested in adults' and children's inferences about who would win a conflict over a desired resource between low- and high-status parties.

Table of Contents

Introduction **1**
Relationships are challenging *3*
The current state of affairs *5*
The need for further research *11*
The questions that matter *14*
The current dissertation *16*

Paper 1: Children's inferences about friendship and shared preferences based on reported information **19**
Introduction *20*
Experiment 1 *30*
Experiment 2 *35*
Experiment 3 *37*
Experiment 4 *40*
General discussion *47*

Paper 2: Preschoolers' recognition of and inferences from observed imitation **51**
Introduction *52*
Experiment 1 *57*
Experiment 2 *66*
Experiment 3 *70*
Experiment 4 *82*
General Discussion *87*

Paper 3: Inferences about yielding in hierarchical relationships: A cross-cultural comparison **91**
Introduction *92*
Experiment 1 *99*
Experiment 2 *112*
Experiment 3 *118*
General Discussion *124*

Conclusion **130**

References **138**

Appendix **149**
Appendix A – Paper 1 *150*
Appendix B – Paper 1 *153*
Appendix C – Paper 1 *155*

Acknowledgements

First and foremost, I wish to thank my advisor, Susan Carey, who is a true model to live up to in science, in mentorship, and in life. Working with Susan has been a true privilege, and I am deeply indebted to her for making my graduate school experience better in every way. Susan has been consistently supportive of my attempts to answer questions that I find interesting, while deftly course-correcting me along the way. She has seamlessly alchemized my half-baked ideas into empirical questions actually worth testing, and shown me how to tackle the challenges that start once data collection stops. I sincerely hope some of it has rubbed off on me in the end. It is not lightly that I say this, Susan, but I wish I could do three more PhDs with you!

I would also like to thank my committee members. First, my gratitude goes to Jesse Snedeker, for always taking the work seriously, for questioning it vigorously, and for caring. I have never come across more on-point and at the same time out-of-the-box questions than those that Jesse has sent my way, and I do not expect to going forward either. I thank Joe Henrich for his valuable feedback on the work. Joe's approach to the study of culture made me realize there is a big beautiful world of fresh ideas outside of William James Hall, and that those ideas can inform the research I do and bear on it in unexpected ways. I also would like to sincerely thank Paul Harris, my outside examiner, and I look forward to his feedback.

Part of this dissertation was conducted in collaboration with Anni Kajanus and Felix Warneken. I am grateful to Anni for introducing me to her fascinating ethnographic work, and for many enjoyable and illuminating conversations about culture, cross-disciplinary research, and much else besides. I am also thankful to Felix for welcoming me to his lab and supporting the shared project.

I am deeply indebted to the wonderful Lori Markson, first a mentor--now also a friend--for her enthusiastic support, both personal and professional. Lori was the person who solidified my interest in social cognition, and I can happily say that I have not looked back since.

Carey Lab is a reflection of Susan's approach to research and as such, I have consistently found it a to be a collaborative, productive, and happy environment. I am grateful to my many labmates Alexandra Was, Jonathan Kominsky, Shilpa Mody, Roman Feiman, Nancy Soja, Rebecca Zhu, Paul Haward, Ivan Kroupin, and Brian Leahy, for their help, thoughtful discussions on the work, great conversations on all else, and the precious sense of camaraderie. I am also grateful to many great LDS friends, including but not limited to Alia Martin, Lindsey Powell, and Kate Hobbs.

To the many research assistants across the years, the biggest thank you! This work would not have been done without all your efforts, and I have learned more from working with you than the other way around. The list is long, but I would like to especially mention Sona Kumar, Esme Trahair, Megan Dempster, and Jessica Pu for their indispensable contributions. The other unsung heroes of developmental research are the amazing children, and their families who generously share their time, and I am deeply grateful to all of them.

I would like to sincerely thank the department's graduate assistant, Celia Raia. Celia is a fairy godmother to graduate students, knowing the answer to every question and never failing to smooth the path in critical moments with grace and compassion.

I am immeasurably grateful to my parents. They have given up so much to see me succeed, not the least of which has been a bitter and difficult separation across continents and time zones, exacerbated by political roadblocks that robbed us of many moments that should have been shared. My mother, Parvin, has always been and remains the example I still strive

towards, in integrity, strength, and humor. My father, Mohammad Naghi, is the most persevering and dedicated person I know, and this coupled with his unfailing optimism never cease to inspire me. I am also deeply grateful for the love and support of my extended family, some kin through affection, others also through biology.

A huge token of gratitude to the tiny people in my life: Sara, who is almost three, and her soon-to-arrive sister, who has been kicking up a storm while I wrote this dissertation. Sara's presence has deeply marked half of my graduate career with joy (and exasperation!), and she has served as living proof of the beauty, madness, and still deeply unknown adventure that is social cognitive development. Finally, the greatest debt is owed my husband, Mohsen, who has been the best of friends and at the center of everything since we were barely more than teenagers. Thank you for the unending love and laughter, for patiently walking besides me in the darkness, for rejoicing with me in the sunshine, for being the toughest editor anyone could possibly have, and the gentlest of companions in all else.

Introduction

Social relationships pervade human life. We grow up nurtured by parents and family, make friends and foes along the way, spend much of our adult life heavily invested in seeking and maintaining romantic bonds, and care for our offspring. We develop ties with our neighbors, bosses, coworkers, and fellow citizens. We navigate rivalries, negotiate conflicts, and battle our enemies. Our well-being, happiness, and perhaps even the survival of our species, is dependent on our ability to form and maintain these bonds (for a review, see Roberts, Arrow, Gowlett, Lehmann, & Dunbar, 2014). Thus, it is not surprising that we are highly interested and engaged in our relationships, but that is not all. We are not just protagonists in the drama of our own lives; we are also hooked on as front-seat observers to the lives of others. We pay attention to what happens between others because it also matters for us. The knowledge and information we gather about those around us allows us to know and understand them better, and importantly to predict how they will act in the future. Through an understanding of the social network we can make more expedient decisions about whom to befriend, support, rely on, learn from, and avoid. The reliability of such predictions makes the social environment more stable and consequently more navigable.

The prominence of gossip and the importance of reputation in human communities highlight the significance of information about others in our own deliberations and decisions. Adults spend roughly two-thirds of their conversations gossiping, and the mental cost associated with the social relationships around us (both those we are a part of and those we observe) may be at the root of the group sizes that we are most facile with (Dunbar, 1993). Given our processing and cognitive capacities, there seems to be a limit on the numbers in a group (e.g. Dunbar has suggested 150) such that we can recognize, monitor, remember, maintain, and work with the multitude of relationships and coalitions involved.

Conceptualizing and recognizing relationships between others is important not only for adults, but also for youngsters. For infants and young children, it is vital in helping them understand why people act the way that they do, for instance why they talk with some individuals more than others, or why they respond to one person's entrance into the elevator with happiness and another's with disinterest. In order to answer questions like these, children need to unlock the information embedded in social behaviors, meaning that they need to understand what specific relationships entail. Importantly, they need to infer relationships from their associated cues when viewing other people, and understand how relationships influence behaviors.

Relationships are challenging

The task of uncovering relationships between others is a formidable one for young observers. Relationships are not visually accessible, meaning that features of individuals' appearance do not disambiguate the relationship between them (there are exceptions to this such as striking family resemblance, but even that requires having an internally represented expectation of facial resemblance among kin). The cues that guide an observer towards the nature of a relationship are embedded in specific information that has to do with the dyad, most often behaviors. In other words, friends—for instance—are people who act towards each other in ways that friends typically do. Therefore, the challenge is understanding how behaviors and relationships are associated with each other, and therein lies the problem. There are a multitude of relationship types. Two people can be friends, enemies, neighbors, coworkers, siblings, cousins, spouses, teacher and student, boss and employee, parent and child, aunt and niece—the list goes on forever. To complicate matters further, the mapping between behaviors and relationships is not one-to-one. Observing an episode of helping could mean that the two people are kin, friends, coworkers, neighbors, or even complete strangers. How can children have any

hope of learning how to dissociate and recognize any of the myriad possible relationships between two individuals from any of the myriad possible observable behaviors? The answer is they may not need to. Although the number of possible social relationships is sizable, there may exist a very small set of fundamental relationships that serve as the basis for all else.

In his seminal relational models theory, Fiske (1992) used extensive anthropological evidence to posit four elementary building blocks to relationships and claimed that all relationships are in essence combinations of these elementary models, which are universally available across different cultures to boot. These four forms are communal sharing (e.g. parent and child, close friends), authority ranking (e.g. military hierarchies), equality matching (e.g. coworkers who socially interact from time to time), and market pricing (e.g. tenant and landlord). The first three types are most relevant to young children. Communal sharing relations are ones in which people are seemingly part of the same metaphorical body. They are each other's equals in many ways, do not keep tabs on shared resources or responsibilities, and act towards each other in an altruistic manner. Authority ranking relationships rise out of contrast along a linear dimension of social value, resulting in (and from) differences in power, control, or resources. Unlike communal sharing, authority ranking relationships are not symmetrical and involve differing privileges and responsibilities for the two parties. Equality matching relations are governed by close consideration for even distribution of work and resources among partners, and balanced rendering of favors from one party to the other. While these basic forms exhaustively define some relationships, other relationships encompass some or all of them, depending on the case. Along the same lines, Clark and Mills (1993) proposed a similar but simpler framework, distinguishing between communal relationships, in which parties act altruistically without considering reciprocity, and exchange relationships, which are driven by balanced benefits. In

similar fashion, Bugental (2000) suggested five social domains that pose distinct challenges. These domains were attachment, hierarchical power, coalitional groups, reciprocity, and mating. These three theoretical frameworks draw on evidence from different approaches, namely anthropology, social psychology, and hormonal mechanisms, and they carve up the social world in slightly different ways. Nonetheless, they affirm two principles. First, that there exists a small set of foundations that all social relationships are built upon. Second, they interestingly underline the centrality of affiliative relationships on the one hand (e.g. close family and friends) and hierarchical ones on the other. Therefore, these relationships may be the ones that are essential to infants and young children as well.

Infants and young children may receive another leg up on the task of conceptualizing relationships from built-in representations. Fiske's (1992) view of elementary relation forms is that, "children impose them on their social world before they learn the implementation rules for realizing them in a culturally appropriate manner" (p. 717). Thomsen and Carey (2013) similarly proposed that infants have an innate set of conceptual primitives when it comes to elementary social relationships. These include a concept for basic relationships and at least one cue to recognize each one by, for instance associating love with being held. Further cues to relationships are subsequently learned through experience and through their association with the initial cue. While the specific set of innate primitives is still very much an open question, their possible existence further highlights the importance of the need to form and recognize social relationships from early on in life.

The current state of affairs

Infancy. Research in developmental psychology has uncovered a host of social proclivities and preferences in infancy. Even as newborns, for instance, infants attend to direct

eye contact (Farroni, Csibra, Simion, & Johnson, 2002) and imitate facial gestures (Meltzoff & Moore, 1977). Later on, they show a wide range of social behaviors, for example pointing to objects around them (Tomasello, Carpenter, & Liszkowski, 2007), preferring speakers of their own language (Kinzler, Dupoux, & Spelke, 2007), and helping others in need (Warneken & Tomasello, 2006). Beyond showing social facility in engaging with others, infants also demonstrate a burgeoning understanding of social behaviors as observers. Just in the context of communicative interactions, for instance, infants expect a person to speak to another person and not to an object (Molina, Van de Walle, Condry, & Spelke, 2004), expect an individual to convey information through speech and not coughs (Martin, Onishi, & Vouloumanos, 2014), and use gaze direction as a cue to finding a speaker's intended audience (Beier & Spelke, 2012). These are some examples of what young infants can do and understand, demonstrating that infants are quite sophisticated in their social abilities and cognition.

Relevant to the current discussion, I now turn to work that has directly addressed infants' understanding of social relationships, predominantly using looking time paradigms. Regarding kin or communal relationships, infants between 12 and 16 months with secure attachment to their caregivers were surprised to see a cartoon caregiver move away from a crying baby, but infants with insecure attachment did not exhibit any particular expectation (Johnson, Dweck, & Chen, 2007). This study shows that securely attached infants, at least, associate caregiving with the intention to comfort a baby when in distress. Building on this finding, Spokes and Spelke (2017) found that 15- to 18-month-old infants expected two infants who had been comforted by the same adult to affiliate with each other, and two adults who had comforted the same baby to also affiliate. Infants did not make these inferences if the same scenes involved same-sized peers and one approaching another following laughing, rather than crying. In other words, what infants

found compelling was a bigger agent's act of coming to the aid of a smaller agent who was upset. This experiment provides some initial evidence of infants' representation of families or communal groups as bound by caring for youngsters.

Moreover, infants have expectations about or based on affiliation. Seven-month-olds expected members of a group, established through an affiliative dance, to perform the same actions as each other, and not those of another group (Powell & Spelke, 2013). At nine months, infants expected two individuals who had expressed similar food preferences to greet each other in a friendly manner, and those who had dissimilar ones to interact negatively (Liberman, Kinzler, & Woodward, 2013). In reverse, slightly older infants expected individuals to share food preferences, but not if they had disengaged and reacted negatively to each other beforehand (Liberman, Woodward, Sullivan, & Kinzler, 2016). Infants also expected an individual to affiliate with someone who had helped her over someone who has hindered her (Hamlin, Wynn, & Bloom, 2007), and showed an expectation for an imitator to approach and affiliate with the target of her imitation (Powell & Spelke, 2017). This impressive array of findings showcases the promise of a largely unexplored landscape of cues and behaviors that infants can decipher to make inferences about affiliative relationships.

Beyond this, infants exhibit understanding of hierarchical relationships. They expect a bigger party to win when in competition with a smaller one, recognizing size as a cue to dominance by ten months (Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011). Infants also expect a member of a larger group to win against a member of a smaller one (Pun, Birch, & Baron, 2016), and for a dominant party to be allocated more resources than a subordinate one (Enright, Gweon, & Sommerville, 2017). Furthermore, their representation of dominance operates over a pair of individuals rather than being tied to each person, meaning that they expect

the party who has previously overcome another in a zero-sum game to win again in new situations, but harbor no expectation of the same party winning against a newcomer (Mascaro & Csibra, 2012). They do, however, expect transitivity of dominance roles under some circumstances (Gazes, Hampton, & Lourenco, 2015; also see Mascaro & Csibra, 2014).

These studies with infants make worthwhile contributions by probing third-party expectations. The promise of such work is amplified by the difficulty of exploring such questions from a first-person perspective with this age range. Given that infants' behavioral repertoire is relatively limited, there is only so much that they can demonstrate in terms of their own behavior in relationships. In other words, there is a better chance of finding out what they know through third-person tasks, and there is every reason to think that what they know and understand extends far beyond what they can do. Further, they provide a glimpse of the initial state (or very close to initial state) of cognition with regard to social relationships, and of the types of behaviors and relationships that infants represent.

Early childhood. With respect to early childhood, a large literature exists on how children conceptualize different relationships, but there is much left unexplored. On the topic of friends, for instance, there is a considerable line of work from the 70s and 80s, which mainly characterized how children described the concept of friendship (Bigelow & La Gaipa, 1975; Damon, 1979; Selman, 1981; Bigelow, 1977; Youniss & Volpe, 1978; Berndt, 1981). Out of these descriptions came several Piagetian-like stage theories, which proposed growing dimensions to children's understanding over time (for a brief review, see Hruschka, 2010). Furman and Bierman (1983) took the work further and conducted an experiment which showed that the stage theories underestimated preschoolers' knowledge of friendship. Their findings revealed that although preschoolers' concept of friendship was more likely than that of older

children to include superficial features (e.g. two people with black hair are likely to be friends), they were nonetheless aware of the crucial and central hallmarks of the relationship in childhood, namely shared play and companionship, affection, and providing support. Thus, even preschoolers seem to know and understand more about friendship than apparent from their descriptions. More recent investigations into children's concepts of friendship are rare, hence the need for more research on this topic, as well as on how the attribution of friendship influences further inferences.

The collection of work focusing on kin understanding is also considerable. Starting with Piaget, there has long been an interest in how children acquire kinship terms. In one of Piaget's classic tasks (1928), children had to find the contradiction in the statement "I have three brothers: Paul, Ernest, and myself." It was not until about ten years that children succeeded at this task, even though they could answer simpler questions, and list their own and other people's siblings easily at much earlier ages. Piaget was more interested in what responses revealed about children's grasp of logic, rather than kinship understanding per se, and theorized that the reason for the failure was children's perspective-taking limitations. While this finding has been replicated with children elsewhere (e.g. in the US by Danziger, 1957), it has also been criticized as not giving a true representation of kinship concepts (Hirschfeld, 1989). Nonetheless, preschool-aged children's grasp of what necessitates kinship relationships is still shaky in the preschool period. Keil and Batterman (1984) found that five-year-olds endorsed someone who had the characteristic features of an uncle, but not the defining ones, as an uncle ("loves you and your parents and loves to visit and bring presents, but he's not related to your parents at all"). In the opposite case, they refused to endorse a person who had the defining but not characteristic feature of an uncle, and thought that a mother's brother who was only two years old could not be

an uncle. Similarly, Landau (1982) found that five-year-olds, unlike older children and adults, were more likely to think that an elderly person was a grandmother even if she had no children compared to a younger-looking person who had children and grandchildren. These findings are added to by a more recent study that asked three- to five-year-old children a number of questions to see if they distinguish between friends, siblings, and strangers, and to examine how they would share with each of these individuals (Spokes & Spelke, 2016). Although three- and four-year-olds differentiated between strangers and friends, only five-year-olds understood the difference between siblings and friends (e.g. siblings, but not friends, have the same mom and share the same grandparents). Further, children were for the most part likely to equally share with friends or siblings (both more than strangers), not privileging kin over non-kin. In short, it appears that conceptions of kin relations may be even more difficult than friendship to master, perhaps not surprising given the many subtypes within them. Furthermore, distinguishing friendship from kinship, both affiliative relationships, may be particularly difficult for young children.

In contrast to friendship and kinship, there has been much recent interest in children's grasp of hierarchical relationships, and efforts on this front have been particularly fruitful. Children in the preschool age can use a variety of cues to identify the high-status person in a pair. According to their judgments, a person is higher status if they enforce their decisions on the other party, if they are older, if they win in play fights, if they possess more of a desired resource (Charafeddine et al., 2015), if they are able to gain access to more resources, if they achieve their goal in a zero-sum situation where the two parties' goals conflict, if they are asked for permission by others (Gulgoz & Gelman, 2016), and if they are imitated (Over & Carpenter, 2014). Children also expect higher-status individuals to win competitive games (Charafeddine et

al., 2015), allocate them larger shares of resources (Charafeddine et al., 2015), and think they are more knowledgeable about the location of a lost item or the name of a novel object (Bernard et al., 2016). To sum, valuable steps have been taken in understanding children's concept of hierarchical relationships.

The need for further research

While these results are of great import and provide insight into how young children think about relationships, they are far from exhaustive. In fact, this topic is still vastly underdeveloped and in need of further research. One important direction is to comprehensively chart out how young children incorporate an understanding of relationships into their predictions about people's behaviors. Predicting future behaviors is itself a crucial social cognitive skill, but that is not all. One strategy for diagnosing concepts is to test the inferences that they support, including inferences about future behaviors. The theory of mind literature is an illuminating example of this. In the Sally-Anne task (Baron-Cohen, Leslie, & Frith, 1985), for instance, children are presented with the following situation. Sally has a marble that she places in her basket before going out. While she is out, Anne hides the marble in a box next to the basket. The question the child needs to answer is where Sally will look for her marble after coming back: in the basket where she left it, or in the box Anne hid it in when Sally wasn't there. Asking children to predict Sally's future behavior gives us an effective tool for discovering how they think about people's knowledge states, which speaks to the broader topic of how mentalizing abilities develop. While the theory of mind literature is not particularly informative regarding social concepts, research has probed children's predictions in social contexts. For instance, preschoolers' reasoning about social categories manifests in their inference that a member of one category is more likely to act negatively towards a member of different category out of allegiance to the group (Chalik &

Rhodes, 2014). In time, older preschoolers come to discount category-based information when in conflict with individual-level information (e.g. a person may act negatively towards another group member they are mad at, rather than a member of another group they are not mad at), and appreciate that a person's mental state is a more telling sign of how they will act than general group-based sentiments (Chalik, Rivera, & Rhodes, 2014). Moreover, five-year-olds think norms are better predictors of future behavior than preferences, but eight-year-olds and adults privilege preferences over social norms (Kalish & Shiverick, 2004). These studies demonstrate that children's reasoning about group-relevant motivations are systematic and that they start off quite rigid, but gain flexibility and nuance with age. What is sorely lacking, nonetheless, is an investigation into how representations of relationships impact predictions about an individual's social behaviors, and how children falter at or master them. In the case of social relationships too, predictions of future behavior can act as gateways into conceptual content.

Another reason to pursue this research program is that the characterization offered by the older literature is in need of revival and improvement. Some of the older work utilized approaches that may have underestimated children's abilities or understanding, such as the clinical interview method (e.g. Bigelow & La Gaipa, 1975; Damon, 1979; Selman, 1981). In many studies using this method, children were asked open-ended questions about a particular concept, e.g. what does friendship mean. Although this method is sophisticated in terms of the depth it can provide, it can put younger participants at a disadvantage as the onus of drawing out and expressing the concept is more on them than the structure and opportunities afforded by the experimental situation. Further, although recent work has succeeded at providing a clearer representation of what children understand by reducing task demands, some studies are couched in frameworks that focus on somewhat different questions. For instance, Shutts and colleagues

found that preschool-aged children prefer befriending other children of the same gender and race as themselves, and four-year-olds even have similar expectations for others in third-person situations (Shutts, Roben, & Spelke, 2013). Although this finding bears directly on how young children think about friendship, the main goal of the research was to examine how they think about the social categories of gender and race.

Finally, it is worth pointing out that even in cases where infancy research has reported positive findings, it does not trivially follow that young children will readily demonstrate success. In other words, what comes out of looking time paradigms with babies does not always generalize to tasks with older children. The most prominent case of this centers on theory of mind, the ability to understand that individuals can have different or false beliefs about the state of the world. Children younger than four years typically fail on the classic test of theory of mind tasks like the Sally-Anne scenario described above, as well as on its many different versions (Wellman, Cross, & Watson, 2001). The mistake that young preschoolers make is to assume that that a person will act in accordance with what the child knows, rather than the person's false belief about the state of the world. In contrast to three-year-olds who fail, a number of looking time studies have shown that infants correctly expect a person to respond according to their own false mental state, instead of the child's true one (e.g. Onishi & Baillergon, 2005; Southgate, Senju, & Csibra, 2007). This discontinuity is baffling, and currently in need of satisfying explanations (for a discussion, see Saxe, 2013). Another instance of discontinuity is the case of object solidity, where an observer needs to infer where an object will end up when coming up against a barrier. Even at two months, infants understand that the object cannot pass through a solid barrier and expect the object to have stopped just short of it (Spelke, Breinlinger, Macomber, & Jacobson, 1992). A number of studies have tested toddlers' understanding of the

same situation, uncovering surprising difficulty. In one study, for instance, two-year-old children watched a ball roll down a ramp behind an occluder (Berthier, DeBlois, Poirier, Novak, & Clifton, 2000). Children were aware of a barrier placed on the ramp, and had to find the ball by reaching into the right door in the occluder, either just before the barrier or at another position further down from it. Toddlers failed at this task, and success only manifested at three years of age, even with training and feedback. Once again, this puzzle is in want of a satisfying explanatory account (for a discussion, see Carey, 2009). To sum, it is not at all clear what precipitates these inconsistent findings about abilities, or indeed if the different cases are instances of the same phenomenon or unrelated yet seemingly similar. Consequently, more evidence is required about the conditions and boundaries of such discontinuities. Due to the importance of social understanding and the range of interesting successes demonstrated by infants, working out what young children understand about social relationships can significantly contribute to this endeavor. In fact, the apparent disconnect between the result that very young infants expect an imitator to affiliate with its target (Powell & Spelke, 2017), and the failure of four-year-olds on making the same kind of inference (Over & Carpenter, 2014) is a relevant example of this.

The questions that matter

In brief, children's conceptual framework of different types of relationships--dyadic ones especially--and what inferences they make about and from relationships are in need of further investigation. A research program designed to implement such an investigation should approach the topic directly, utilizing age-appropriate methods that are most likely to yield a true reflection of children's knowledge. Where possible, the work should also elucidate the continuities or lack

thereof in comparison to infant cognition. The critical questions that need to be answered include, but are certainly not limited to, the following:

- What cues (whether in the form of observed behaviors or heard reports) can children use to infer relationships between others? Is there a canonical set of behaviors and displays that children associate with each type of relationship? Are there cues and behaviors that have one-to-one mappings to certain relationships?
- Are children simply passive processors of such inferential input, or do they also proactively seek it out?
- Is there a categorical difference between how children think of their own relationships compared to those they observe between others (i.e. first-person vs. third-person)? Does first-person engagement in a particular type of relationship inform children's concepts about that relationship type? Are there cases in which third-person observations enrich, inform, or even precede first-person engagement?
- When do children distinguish between similar but distinct relationships, for instance between friendship and kinship, both subtypes of affiliative relationships? In this vein, when do they differentiate between different types of hierarchical relationships, e.g. dominance vs. prestige?
- When do children understand the rules that govern relationships? For instance, when do they develop an understanding of the responsibilities and privileges relationships afford their parties, or how relationships form and break?

- What role does the cultural context play in the above questions? What aspects of social relationship understanding are more prone to cultural differences, and what aspects are more or less universal?

Questions like these and the multitude of questions that follow from them make for an extensive research program in need of focused strategies to gain the best traction. One simple yet important strategy is to first focus on affiliative and hierarchical relationships, because different frameworks of social relationships have indicated both as being central (Fiske, 1992; Bugental, 2000). Indeed, these two families of relationships encompass so much of the relationships that young children experience, engage in, encounter, and find important early on in life.

The current dissertation

The trio of papers presented in this dissertation takes small steps towards exploring the topic of children's social relationship understanding, and probes a mix of questions from different angles. In line with the strategy mentioned above, all experiments deal with affiliative (Papers 1 and 2) or hierarchical relationships (Paper 3).

Paper 1 asks whether four-year-olds privilege information that is relevant to friendship over information that is not when inferring who is friends with whom. Further, it examines whether they make any inferences about shared preferences between individuals based on the same kinds of friendship-relevant information. By asking about the types of information that children can use to make inferences about friendship, the study gains insight into whether those types of information are represented in the concept, and even explores whether some are prioritized over others. Further, this paper investigates whether the clinical interview methods used in older work on friendship (e.g. Selman, 1981) diagnosed the concept correctly, or whether they underestimated what children understand in early childhood. Another general question

pursued here is whether young children can infer a relationship when they are provided with reports of individuals' behaviors towards each other.

Paper 2 explores at what age and under what conditions three- to five-year-old children use observed imitation as a basis for inferring affiliation between an imitator and her target. This paper seeks to provide more clarity on the puzzling discontinuity between affiliation inferences from imitation in infants (Powell & Spelke, 2017) and four-year-old children (Over & Carpenter, 2014). Although five-year-olds succeed at Over and Carpenter's task, it is not clear why four-year-olds do not. The paper tests whether children have an explicit representation of imitation, and whether highlighting imitation for them can lead to not just four-year-olds', but even perhaps three-year-olds' success. Additionally, it examines whether children use the observation to infer continued imitation of the same target in the future. This aspect of the study tests whether children's conceptions of the reasons behind why someone may imitate another person are stable enough to merit them continuing to imitate the person in the future.

Paper 3 changes gears from the other papers in two main ways. First, it deals with hierarchical, rather than affiliative relationships. Second, it inspects inferences through both developmental and cross-cultural perspectives. As mentioned above, previous work on hierarchical relationships has shown that preschoolers can infer status from a host of cues, but it is an open question whether they differentiate between dominance and prestige--two forms of hierarchical relationships, the first rooted in force or threat of force, and the second rooted in merit and influence. In order to examine this, children were recruited at different ages (5-7 years vs. 9-11 years). Furthermore, children were tested in the UK as well as China. Testing children outside of Euro-American cultures provides the first chance of examining whether their inferences about status are in line with previous research, predominantly undertaken in the US

and Europe. The other essential question this paper asks is whether cultural norms and values might influence expectations about outcomes of conflict between parties differing in status. Drawing on ethnographic evidence from China on the value of prestigious individuals yielding to others, even those lower than themselves, the paper investigates a possible cross-cultural difference between children in the UK and China. This allows us to test the possible cultural variation in one aspect of understanding and reasoning about hierarchical relationships.

Paper 1:

Children's inferences about friendship and shared preferences based on reported information

Narges Afshordi

Introduction

Recognizing social relationships between people and understanding how relationships regulate behaviors between individuals poses a formidable task for young children. Relationships are not features of an individual but rather of two or more partners. As such, an observer whose goal is to understand the nature of the relationships between others needs to attend not only to each party, but also—and perhaps more importantly—to any connections between them. Such connections are sometimes perceptually accessible. For instance, preverbal infants expect a relatively bigger individual to dominate a smaller one (Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011). More often, however, cues to relationships are embedded in detailed and specific dyadic information, such as individuals' actions and behaviors towards each other, and making correct inferences about relationships depends on successful interpretation of such information. The current studies examine whether four-year-olds infer friendship between peers by relying on reported information about the connections between individuals. This method of soliciting children to identify friendships between others can be highly effective not only in gauging their abilities in uncovering friendships, itself a socially important skill, but also in providing insight into their concept of this particular relationship.

If children are to infer friendship between peers, not only do childhood friendships need to exist, but also children need to possess a concept of friendship to rely on when making inferences. With regard to childhood friendships, they do in fact exist, and are highly important to children's social lives, providing them with one of their first experiences of non-kin relationships. Children make and maintain friends as early as one year of age (Howes, 1983), and improve in social competence and modes of play with friends with age (Howes, 1987). Children talk differently when in conversation with their friends, and patterns of gossip, self-disclosure,

and joking appear in speech between preschooler-aged friends (Dunn, 1993). In contrast to initial skepticism in the field, early friendships are typically stable over time, with pairs of four-year-olds in one study having been friends for two years (Howes, 1987). Friendship also has broad implications for children's lives; it provides them with emotional support (Howes & Farver, 1987; Schwarz, 1972) and later on predicts school adjustment and performance (Ladd & Kochenderfer, 1996). Thus, children's friendships appear to be stable and meaningful.

In addition to studying early friendships, researchers have examined the development of children's thinking about the concept of friendship, mainly with the use of clinical interviews (Bigelow & La Gaipa, 1975; Damon, 1979; Selman, 1981). In these studies, children were typically asked to describe verbally (and sometimes in writing) what a friend is or what they value in their friends compared to others. Mostly concerned with explicit theories of friendship, the majority of studies in this literature did not include preschoolers, and the youngest groups were typically either in kindergarten or first grade. In a seminal study, Bigelow (1977; also see Bigelow & LaGaipa, 1975) categorized children's essays about their best friends along a number of dimensions. This work reported a Piagetian-inspired stage-like progression, and talk of common activities, evaluation and propinquity dominated descriptions in the first stage by children younger than ten years. Admiration for a friend followed in the next stage, in turn followed by more sophisticated notions such as acceptance and common interests. In another stage theory, Selman (1981) proposed that 'Stage 0' spans from three to seven years of age, and that children's notions of friendship during this time are superficial and situational, such that a friend is someone who is proximate and playing with the child at a particular moment. Selman claimed that in the subsequent 'Stage 1', children define friendship in the context of receiving help and support from their friends, and knowing them at a deeper level. Youniss and Volpe

(1978) described friendship for children around age six to seven as centered on rules of behavior around situations like playing together. Berndt (1981) asked kindergarteners how they would know someone is their friend, and they similarly responded with common activities, as well as what he viewed as defining features of friendship to them such as knowing someone, liking them, and whether they said they were the child's friend. Finally, using a different approach, Corsaro's (1981) ethnographic study adds another dimension by highlighting when and how children explicitly refer to friendship within peer groups. His work, the only one in this set to focus on three- and four-year-olds, showed that preschoolers' foremost objectives when in peer settings was joining ongoing play episodes, and barring access to others once involved. In such situations, children often used friendship as a pretext for inclusion or exclusion, and thus two children engaged in play would label themselves as friends, while a child who wanted to exclude another from joining in would assert that they were not friends. The overall pattern from this literature suggests that if preschoolers, or in fact children in the early school years, have any notion of friendship at all, it is one of a relationship that is superficial, transient, and restricted to specific play episodes.

Despite the picture drawn by these studies, there are reasons to doubt that early understanding of friendship is as superficial as suggested. First, as mentioned above, children's actual friendships seem to involve noteworthy depth, which may be reflected in their understanding of the relationship. Second, the methods used in the majority of such studies require significant explicit knowledge of the concept, which may have impeded younger children from revealing all that they knew. In general, open-ended questions elicit spontaneous responses that readily come to mind, and as such they may not be good indicators of the presence or absence of children's understanding of a particular aspect of friendship. In a way, the difference

between children's responses to these questions and their potential responses to less demanding questions is akin to the distinction between recall and recognition in memory research. Given this concern, it is plausible that young children's concept of friendship was underestimated by the studies mentioned above. Evidence for this comes from Furman and Bierman (1983), who asked two groups of children (4-5 years and 6-7 years) to describe what friendship meant to them in an open-ended interview (e.g. what a friend is, why people need friends). Crucially, they included several versions of the question and additional prompts to encourage more responses, and added other components to the study. In a second task, they showed children triads of line drawings depicting different features (affection, support, common activities, propinquity, and physical characteristics) and asked them to choose ones that indicated friendship. In the third task, subjects saw the same drawings in all possible pairings and were asked to choose the one more important to friendship. In the open-ended interview, common activities such as playing together were the most frequently mentioned feature by almost all children. Contrasting with other studies (e.g. Selman, 1981), most of the younger children talked about affection (63%), around half even mentioned support (53%), either as helping or sharing, and recognition of these topics in the other two tasks significantly increased with age. Thus, when asked to rank features important to friendship, the older group thought affection and support were more important than common activities, while younger participants found all three (i.e. common activities, affection, and support) comparable. These results reveal that preschoolers' knowledge about friendship is deeper than posited by earlier work, and while still undergoing development, already includes core notions of companionship, affection, and support. Nonetheless, there were also indications of more shallow aspects to the concept, especially for the four- and five-year-old children. The second most frequent feature talked about by the younger group was physical characteristics,

describing particular details about appearance, which younger children did at a significantly higher rate than their older counterparts (69% compared to 38%). In the tasks involving the line drawings, depictions of physical characteristics showed two children wearing blue hats and having black hair respectively. In both tasks, physical appearance responses were significantly higher for younger participants. In a similar vein, younger children talked more about propinquity than the older group (50% compared to 31%), although this difference was not significant. Overall, this study demonstrates that the early concept of friendship is anchored in meaningful features, but at the same time is not completely free of superficial associations. Crucially, however, Furman and Bierman's study does not clarify whether four- and five-year-olds can infer friendships between others, and if so whether they view any of the features they acknowledged with regard to friendship as a basis for such inferences.

Recent work has shown that preschoolers make systematic choices when asked to infer friendship, both for themselves and for others. When presented with pictures of two children of different genders and asked whom they would like to be friends with, three-year-olds preferred same-gender peers (Shutts, Roben, & Spelke, 2013). Similarly, five-year-old children preferred to be friends with other children of their own race, others who spoke in their native language, and in their native accent (Kinzler, Shutts, DeJesus, & Spelke, 2009). In third-person tasks, four-year-olds expected a child to be friends with another child of the same gender and race (Shutts et al., 2013). Interestingly, these responses reflect reality, and studies on children's actual friendships reveal that young children are indeed predominantly friends with others of their own gender (e.g. Maccoby, 1990), race (e.g. Aboud, Mendelson, & Purdy, 2003), and even age (Guralnick & Groom, 1988). Overall, children's responses to questions in these experimental tasks seem to hint at coherent inferences about friendship. However, given that these studies deal

with social category markers and as such, they do not shed light on the question of what friendship means to preschoolers, and whether they understand the term to denote anything beyond the liking and affiliation typically associated with fellow group members. Additionally, they do not reveal whether preschoolers are able to infer friendships based on more nuanced information. While the convergence across experimental and observational studies highlights the importance of social categories such as race and gender with regard to early friendships, children do not treat all in-group members as friends. In fact, they are often surrounded by many possible partners who match them on broad categories such as gender and race. Nevertheless, that is not to say that children are friends with all of them, or that friendships across social category boundaries are nonexistent. Thus, if children rely only on category markers to infer the friendships around them, they would wrongly assume all group members to be friends and miss any relationships that did not fit those criteria. Consequently, the ability to infer friendships between others often necessitates attention to more specific, dyadic information. Given all of this, the question of whether children are able to consider such dyadic information—evidence about behaviors and associations between individuals—when reasoning about friendship is a highly important and understudied one. Not only would the answer to this question further elucidate how much children know about friendship as a relationship, it would also clarify the role of that knowledge in interpreting the social scene. This paper tackles this issue across four experiments: Experiments 1-3 assessed preschoolers' ability to infer friendship from reported dyadic information, and Experiment 4 examines inferences about shared preferences.

Given Furman and Bierman's (1983) findings, the current studies tested four-year-olds who are most likely to possess both deep and shallow associations with regard to friendship, rendering them an interesting age in which to examine inferential capacities. In each experiment,

participants viewed pictures of three children and heard facts of two different types linking a central character to each of the other two individuals. They were then asked to identify the central character's friend, in effect by adjudicating between the two types of facts that were reported to them. Three fact types were used across the experiments: joint activities, similarity, and arbitrary links.

The first type of fact was about *joint activities*, and similar to the category of common activities that almost all children mentioned when asked to talk about friendship (Furman & Bierman, 1983). Given the special significance of playing together in early friendships (Furman & Bierman, 1983; Howes, 2009), the facts used here avoided describing play episodes, in order to focus more generally on shared activities. Instead, the joint activities described in the facts were initiated by either helping or invitation. In the case of helping, the reports described the central character being helped by another child, followed by the two of them completing a task together (e.g. "Matthew's room was messy and this boy helped him clean it up. They worked together and got the room tidy.") Early work suggests that children do not have an appreciation of the support involved in friendship until much later than age four (Selman, 1981), and even in Furman and Bierman's study (1983), only half of a sample of four- and five-year-olds readily mentioned support of any kind (including helping) when asked about friendship. However, more recent research has revealed that even infants have rudimentary expectations of affiliation between helper and helped (Hamlin, Wynn, & Bloom, 2007). Further, children are more likely to help friends than non-friends (Dunn, Cutting, & Fisher, 2002), and three-year-olds think others should share more with friends than strangers (Olson & Spelke, 2008). Therefore, it is possible that four-year-olds do in fact associate helping with friendship, and could rely on it (especially when paired with a joint activity) to identify friendships. In the case of invitations, the report

described the central character inviting another child to a social occasion, followed by their joint activity (e.g. "Maya invited this girl to the zoo. Last week, they went to the zoo together and they saw all sorts of different animals.") Although inviting or being invited has not been reported in in children's spontaneous talk about friendship until middle childhood (Youniss & Volpe, 1978), children's own experiences of invitations in the context of friendship, and the fact that invitations indicate an intentional goal to socially engage with the invitee, may help them recognize friendships based on it.

The second type of fact described *similarity* between two individuals. The link between similarity and attraction has long been a topic of study (e.g. Newcomb, 1956), and applies to childhood friendships as well. Apart from being similar to friends in terms of group markers such as age, race, and gender (Guralnick & Groom, 1988; Aboud et al., 2003; Maccoby, 1990), young friends are similar in other ways such as temperament (Dunn & Cutting, 1999). In experimental work, Fawcett and Markson (2010a) found that three-year-old children preferred to play with puppets that were similar to themselves in terms of hair color and preferences for foods and toys, but they did not show the preference when the puppet's similarity to them consisted only of having received the same sticker that was given out arbitrarily. While this work does not directly probe friendship choices, it has revealed that young children prefer similar others and can be discerning when it comes to the dimension of similarity. In short, children's choices and patterns of friendship appear to suggest that they value similarity in making and retaining friends. Given this, the question arises whether similarity only affects children's own choices, or whether it also informs their concept of friendship more generally, thus extending to their judgments about others' relationships in third-party situations. The similarities that were reported to children here were of two subtypes. Facts about similarity in ability described two characters as excelling at a

particular skill, such as singing. This dimension was chosen as a possible signal of friendship on the assumption that individuals who possess a skill have an implied shared interest that could lead them to enjoy conversing and spending time with one another. The second type of facts described similarity in experience such that two characters separately had the same experience, for instance having visited the aquarium recently. The specific experiences that were included were chosen for their possible appeal to young children as interesting occurrences that may imply deeper similarities and possible opportunity for bonding. Overall, use of the similarity facts examined whether dyadic similarities would be sufficient for children to infer friendship. Some support for this possibility comes from the fact that four-year-olds expect friendship between two children of the same gender or race (Shutts et al., 2013). Additionally, elementary school children use similarity in academic orientation and athletic ability to infer friendships between their actual classmates (Neal, Neal, & Capella, 2014). As a result, it is possible that four-year-olds would also expect friendship between two individuals who are reported as being similar. On the other hand, children's responses in Shutts et al. (2013) may have been driven by a general expectation of liking among social group members, rather than any appreciation of the role of similarity in friendship. Thus, it is alternatively possible that similarity only influences children's own liking of others at this age, but not their expectations about others.

The final type of fact reported to participants was about *arbitrary links*. These facts were of the flavor of the more shallow features that four- and five-year-olds seemed to associate with friendship (Furman & Bierman, 1983), namely propinquity and physical characteristics. Importantly, the facts were generated in order to be utterly coincidental in nature, so as to offer a foil against the first two types of facts in a forced-choice task. The arbitrary links consisted of the subtypes coincidental proximity (e.g. one character was eating in a café that the other character

passed while riding in a taxi) and coincidental commonality (e.g. it had snowed on both characters' birthdays). Young children associate proximity (or propinquity) with friendship (Bigelow, 1977; Furman & Bierman, 1983), thinking of friends as being physically near each other. However, it is not clear how deep their grasp of proximity as a signal of friendship is. If children's understanding of the role of proximity is shallow and not context-dependent, they may view any kind of proximity as a cue to friendship. If, however, children understand that proximity is indicative of friendship only in particular situations (e.g. choosing to always sit next to each other at lunchtime), they should not treat it as the basis of inference in the current task, since the proximity occurs completely by happenstance. Similarly, children may erroneously view coincidental commonalities as meaningfully highlighting the affinity between two characters. The four- and five-year-olds in Furman and Bierman's (1983) study were more likely than their older counterparts to associate shared physical characteristics (e.g. having black hair) with friendship. If their notion of shared non-physical features is equally shallow, then they would be open to interpreting any statement that talks about a shared feature between two individuals, however meaningless, as a sign of friendship. Alternatively, children may be more cautious in relying on such information when reasoning about friendship. Previous work shows that children *can* disregard arbitrary commonalities. Three-year-olds prefer similar others, but not when the similarity is limited to arbitrarily having received the same sticker (Fawcett & Markson, 2010a). Impressively, even infants who show a comparable liking for others who are similar to them do not prefer an individual just because they have both received the same-colored mitten by chance (Mahajan & Wynn, 2012). Five-year-olds do not rely on having the same proper name, a coincidental commonality, when inferring shared preferences between

individuals (Diesendruck & haLevi, 2006). Thus, children in this task too may discount the coincidental commonalities.

Experiment 1

The two facts types presented to participants in this experiment were joint activities and arbitrary links. As mentioned above, preschoolers' concepts of friendship identify core features of the relationship, but are also susceptible to more shallow associations (Furman & Bierman, 1983). Further, it is not known whether children can rely on the dimensions they associate with friendship when asked to infer it between third-party characters. Therefore, one possibility is that children may think that any kind of link between two people, meaningful or arbitrary, can be taken as a sign of friendship between them. If so, children in this experiment would not weigh evidence of joint activities more heavily than mere hints of an association such as those in the arbitrary links, and would not respond systematically. On the other hand, it is possible that four-year-olds' notion of friendship is rich enough to prioritize the more meaningful dimensions when making inferences. If so, four-year-olds would privilege joint activities over arbitrary links when inferring friendship between others.

Method

Participants. Sixteen typically-developing four-year-old children (7 female, M_{age}=54.7 months, range=4;0-4;11) from a metropolitan area participated in this experiment. Children were recruited from a laboratory database of families interested in participating in experiments. The majority of families in the database are middle class and Caucasian (79% European American, 15% Asian American, 6% African American). Three additional children were excluded from analysis due to failure to complete task (2), and experimenter error (1). Parents received travel reimbursement and children chose a small prize for participation.

Materials. The task was presented using Keynote on a 13-inch MacBook laptop. Eight sets of pictures were prepared, each containing three child faces. Within each set, one was chosen as the central character and the other two as possible targets. All pictures were rated by a group of adults and pairs of possible targets were chosen to be similar on perceived age, happiness, and attractiveness. All children in a given set were of the same gender, with half the sets containing pictures of girls. On each trial, the central character's picture was centered at the top of the screen (see Figure 1.1 for a sample layout). Possible targets were presented in the lower half of the screen, equidistant from the central character. The central character was referred to using a common name (e.g. Maya, Jake), while the other two were referred to as "this girl/boy" or "this one" throughout. Each trial also involved a pair of hint images to help children encode the information facts that they heard. Hint images were colorful clip-art depictions of items thematically related to each fact, e.g. a picture of a card to help participants remember a fact about one character helping another make a card for her mom. Hint pictures appeared between the central character and each of the possible targets.

Sets of facts about joint activities and arbitrary links were prepared (see Appendix A for a full list). There were two subtypes of joint activities facts: helping (e.g. one character helped the other character make a card) and invitations (e.g. one character invited the other character to a sleepover). There were also two subtypes of arbitrary links facts: proximity and commonality. As in the example above, proximity facts described situations in which the two characters were physically close by without actively having chosen to be or having knowledge of the other's location. Commonality facts on the other hand described shared features between characters that were not meaningful or indicative of any other shared features.

a)



b)



Figure 1.1. Sample trial layout from Experiment 1 (top), and sample trial layout from Experiment 4 with a novel food item (bottom).

Procedure. The participant and the experimenter sat at a child-sized table facing the laptop screen. At the start of each trial, pictures of the three characters and hint images appeared on the screen. The experimenter began by pointing to and naming the central character (e.g. "This is Maya.") She then told the subject about a joint activity between the central character and one of the characters, and an arbitrary link between the central character and the third character. Order of presenting the two facts was counterbalanced across trials and they were matched for

length. At each point, the experimenter drew the subject's attention to the relevant character by pointing.

After explaining the facts, hint images disappeared and the experimenter asked questions to ensure that children remembered the information (memory checks). She asked which character each fact referred to, and encouraged the subject to point. If a subject answered incorrectly, hint images were shown again and the fact was repeated. Following this, the experimenter asked the test question while pointing to the central character, for instance saying, "Maya is friends with one of these girls. Which one is her friend?" Children responded by pointing to one of the possible targets. If a child hesitated to respond, they were prompted to do so, "It's okay to guess." Sessions were recorded for later offline coding.

Design. Each subject viewed eight trials, consisting of four female- and four male-character trials, which alternated within each four-trial block. Helping and invitation facts were interspersed across the eight trials, as were proximity and commonality facts, with the constraint that two instances of each fact type appeared in the first four-trial block. Side of target (i.e. the character who engaged in the joint activity with the central character) and order of presentation of the joint activity fact were counterbalanced within and across participants.

Results

Participants chose the correct character in response to memory checks 95% of the time, demonstrating that they remembered the facts they heard almost perfectly. In response to the test question, which asked who the central character was friends with, children picked the target, i.e. the character involved in a joint activity, 71% of the time (see Figure 1.2). Analyses were carried out using the lme4 package in the R statistical computing software (2013). A mixed-effects binary logistic regression model containing only the intercept as a fixed effect and subject as a

random effect (glmer(Response ~ 1 + (1|Subject)) showed this performance to be significantly better than chance (b = .98, SE = .24, z = 4.15, p < .001). (A one-sample t-test against chance (set at 50%) using proportion of correct responses for each child provided the same finding of performance being significantly better than chance, $t(15) = 4.06$, p = .001. Henceforth, I will not report results using data aggregated for each participant, but will rather use regression models as they take into account that each participant provided responses for several trials.) Subjects chose the target character 69% of the time on invitation fact trials and 75% on helping fact trials. A second binary logistic regression including target subtype as a fixed effect and subject as a random effect but no intercept (glmer(Response ~ FactSubType + (1|Subject)) revealed that performance on helping and invitation trials were not significantly different from each other, (b = -.32, SE = .40, z = -.80, p = .42).

Discussion

Four-year-olds inferred that an individual's friend was a person they had engaged in a joint activity with, either because the person helped them or because they had invited the individual to a social occasion, but not friends with someone who happened to be in the same location or who happened to have a minor commonality with them. This result reveals children's theories of friendship to be richer and more sophisticated than suggested by some earlier work. Importantly, their concept is strong enough to merit inferences about the existence of friendship between two people the child knows nothing else about. It is also the first evidence of children inferring friendship on the basis of reported behaviors. Nonetheless, the study is limited due to the nature of the forced choice that was used. This experiment does not clarify whether children's choices were driven by reliance on joint activity facts, by avoidance of arbitrary links, or by a combination of the two.

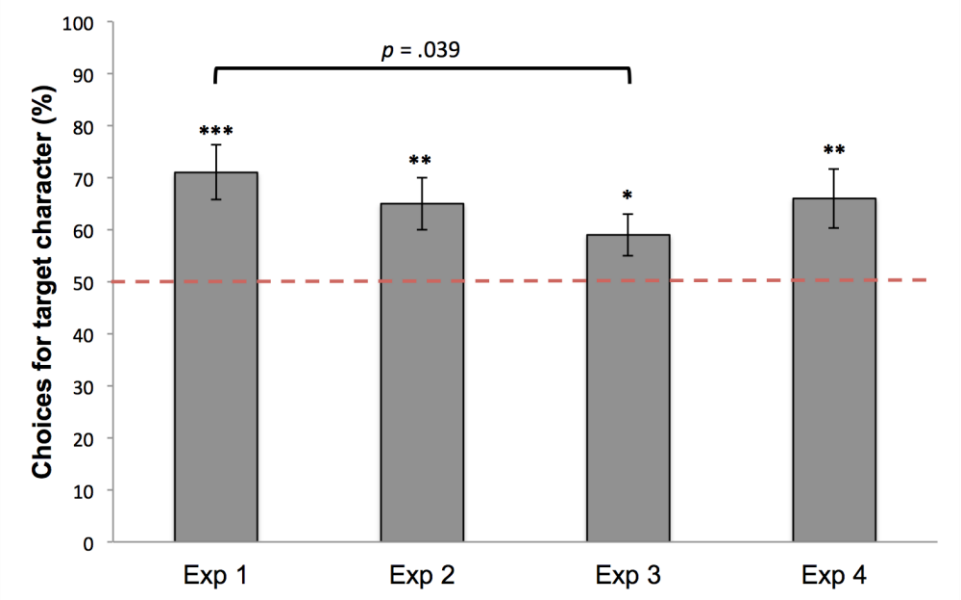


Figure 1.2. Results for friendship choices from Experiments 1-4. Bars reflect standard error, and asterisks denote that performance was significantly different from chance (* $p < .05$, ** $p < .01$, *** $p < .001$). The bracket above shows that performance in Experiment 1 was significantly different from Experiment 3.

Experiment 2

This experiment tested whether similarity affects four-year-olds' inferences about friendships between others. Therefore, children heard reports about similarity and arbitrary links. As mentioned above, the similarity facts comprised of two subtypes: similarity in ability and similarity in experience. These two subtypes of similarity were pitted against the arbitrary link facts used in Experiment 1 describing either coincidental proximity or coincidental commonality. It is possible that despite similarity's effect on first-person liking, children do not view it as relevant when asked to uncover friendships between others. Alternatively, similarity may act as a possible signal of friendship between others to young children.

Method

Participants. Sixteen four-year-old children (10 female, M_{age}= 53.3 months, range=4;0-4;11) took part in this study. Children were recruited from the same database as before, and none had taken part in the previous experiment.

Materials. A set of eight facts about similarities were prepared which described two characters as similar in terms of excelling in a particular activity (e.g. both being good at drawing) or similar in terms of an experience (e.g. both having gone to the beach for holidays). Similarity facts were length-matched to arbitrary link facts. Eight hint images were chosen to correspond with the similarity facts, e.g. a clip-art picture of a fish tank to go with the fact about a visit to the aquarium. The set of arbitrary link facts and their corresponding hint images from Experiment 1 were used, which described coincidental proximity and coincidental commonality. The eight sets of child face pictures were identical to Experiment 1.

Procedure. The procedure was the same as Experiment 1.

Results

Performance on memory checks was high at 86% correct, showing that once again, subjects remembered the reported facts. In response to the test question, participants chose the target, i.e. the similar individual, as the friend 65% of the time. A regression model like that used in Experiment 1 was run on the data, such that overall performance against chance was tested by a mixed-effects binary logistic regression with only the intercept as a fixed effect and subject as a random effect (glmer(Response ~ 1 + (1|Subject)). Participants' responses were significantly above chance (b = .64, SE = .22, z = 2.86, p = .004). Children chose the target character correctly 67% for similarity in ability and 63% for similarity in experience. Responses to questions involving the two subtypes of similarity were not significantly different from each other (b = -.22, SE = .38, z = -.57, p = .57) according to a mixed-effects binary logistic regression including

fact subtype as a fixed effect and subject as a random effect (glmer(Response ~ FactSubType + (1|Subject)).

Discussion

Four-year-olds in this study tended to infer friendship between two individuals who were described as having similar abilities or experiences, but not between two individuals who shared a coincidental proximity or commonality. This finding suggests that (at least) in the context of reported information, four-year-olds can use similarity as a basis for inferring friendship. Thus, the role of similarity is not limited to children's own friendships, and extends to their reasoning about other people's as well. Once again, it is not clear whether children's choices reflect reliance on similarity, avoidance of arbitrary links, or a combination of the two. The next experiment sheds some light onto this issue.

Experiment 3

The goal of this experiment was to examine the relative priority, if any, of joint activities or similarity in children's third-party friendship inferences, and in their concept of friendship by extension. Once again, the task consisted of being presented with one central character and two potential targets. The link between the central character and one of the individuals was a joint activity (initiated through helping or invitation). The link to the other character was similarity between the two individuals (in terms of ability or experiences). Knowing these two pieces of information, and knowing that only one of the two possible targets was the central character's friend, children had to identify the friend by deciding which piece of information, if either, was more informative in drawing inferences about friendship.

On the one hand, interactions and behaviors are readily available and a reliable source of information about others' relationships, and in fact many of children's descriptions about

friendships center on joint activities (e.g. Youniss & Volpe, 1978; Furman & Bierman, 1983). Further, voluntary actions such as helping or invitations used here imply the intentional goal of the individuals involved to affiliate with each other in a way that being similar to another person does not. As such, children may value behaviors directed from one person to another as stronger evidence of their friendship, and choose the character that interacted with the central character as his or her friend. On the other hand, many children's social experiences may entail often seeing or hearing about multiple joint activities between others, but not many observations or reports of people as being similar to each other in a specific regard. Thus, it is alternatively possible that similarity is more informative precisely because it is a less frequently available piece of dyadic information, thereby having more information value. In other words, reported similarity may be a more reliable basis on which to infer friendship. If this is the case, children should choose the similar character as the friend. Finally, it is possible for young children to view joint activities and similarity as equally informative and equally indicative of friendship between others, in which case their overall responses would not favor either feature.

Method

Participants. Sixteen new four-year-old children (7 female, M_{age}=55 months, range=4;1-4;11) from the same population as before were tested.

Materials. The facts about joint activities from Experiment 1, consisting of four facts about helping and four facts about invitations, were used. These statements were matched against the similarity facts from Experiment 2, which include four facts about similar abilities and four facts about similar experiences. The hint images associated with the facts were the same as those used before. The child face pictures were the same as the previous experiments.

Procedure. The procedure was the same as previous experiments.

Results

Children remembered facts with very high accuracy: 97% correct. When asked to identify the central character's friend, participants chose the character that had performed a joint activity with him or her 59% of the time. This performance was significantly above chance (b = .38, SE = .18, z = 2.11, p = .035) according to a regression model like the ones used in the previous experiments (glmer(Response ~ 1 + (1|Subject)). As the joint activity facts were used in both Experiment 1 (joint activities vs. arbitrary links) and the current experiment (joint activities vs. similarity), a regression model with pooled data was run in order to compare children's performance across the two studies. This model contained experiment (Experiment 1 vs. Experiment 3) as a fixed effect and subject as a random effect (glmer(Response ~ Experiment + (1|Subject)), and found a significant difference between the experiments (b = .56, SE = .27, z = 2.07, p = .039), such that children chose the character that had engaged in a joint activity with the central character significantly more in Experiment 1 than in Experiment 3 (71% compared to 59%).

Discussion

Four-year-olds as a group inferred the central character to be friends with the person he or she engaged in a joint activity with, and not the individual similar in terms of abilities or experiences. This result suggests that four-year-olds prioritize information about joint activities over similarity in inferences about others' friendships. By extension, it is possible that joint activities are of higher value in children's friendship concepts at this age. The comparison between Experiments 1 and 3 found that children chose the character with a joint activity significantly more when the other candidate had an arbitrary link (Experiment 1) rather than when was the other candidate was similar (Experiment 3). In other words, four-year-olds found it

easier to disregard arbitrary information as opposed to similarity when choosing joint activities as a basis for friendship inference, further reflecting that similarity itself is indicative of friendship to them. Together with the result from Experiment 2 (similarity vs. arbitrary links), these findings demonstrate that children view similarity as informative about friendship, but not to the same degree that they find joint activities informative.

Given the forced choice structure of the experiments, the issue was raised earlier that children's performance in the previous experiments could be due to either seeking the target character, avoiding the non-target character, or a combination of the two. The current experiment clarifies this ambiguity to a great extent. The repeated reliance on joint activities in Experiments 1 and 3 suggests that responses in Experiment 1 were not solely a result of avoiding the character with arbitrary links, because children recognized the informational value of joint activities in two separate studies. Additionally, given the significant difference between Experiments 1 and 3, children could not have been responding by solely relying on joint activities. In fact, the difference in responses between Experiments 1 and 3 suggests that children were attending to both types of information and weighing them against each other when making a choice.

Given children's repeated reliance on joint activities in reasoning about friendship, the next experiment takes a step towards examining what other inferences, if any, they draw from this information.

Experiment 4

This experiment served three purposes. First, it aimed to replicate the effect reported in Experiment 1. Therefore, new instances of facts about joint activities and arbitrary links were used and children were once again asked to infer a character's friend from reported information. Second, the experiment asked children to also identify the central character's playmate based on

the same information. Children at this age have a strong association between friendship and playing together (Furman & Bierman, 1983; Howes, 2009), so their ability to infer friendship between two individuals should reasonably extend to inferring them to be playmates, and they should regard evidence for friendship (i.e. joint activities initiated by helping and invitations) as also evidence for playing together.

Finally, and perhaps most importantly, this experiment explored whether children can use the same information they viewed as indicative of friendship in Experiments 1 and 3, namely joint activities, as a basis to infer shared preferences between individuals. Preferences are an interesting case of properties that are individual on the one hand, and open to social influence on the other. They are specific to each person, but they can also match those of socially-connected others, and children base their own preferences on those of other group members. For instance, 12-month-old infants show preference for a food endorsed by a native speaker of their language (Shutts, Kinzler, McKee, & Spelke, 2009), and three-year-olds prefer objects and activities favored by others of their own gender and age (Shutts, Banaji, & Spelke, 2010; Shutts et al., 2013). Children can also reason about preferences in third-person situations, although the manner in which information is presented seems to influence their responses. Shutts and colleagues (2013) found that four-year-olds did not readily expect children of the same gender or race to prefer the same activities when the social categories were not explicitly pointed out. On the other hand, Diesendruck and haLevi (2006) found that five-year-old children expected individuals described to be of the same gender, social status, religion or ethnicity to share preferences (also see Heyman & Gelman, 2000, for similar findings).

This experiment involved preferences for games and food, two domains that have been explored in previous work about preferences. Given that young children's conceptions of

friendship center on playing together and shared activities (Furman & Bierman, 1983; Howes, 2009), it would be reasonable for them to think that friends like the same games, as this would facilitate shared play. With regard to food, there is evidence in support of the social influence of others on first-person food choices in early childhood (Shutts, Kinzler, & DeJesus, 2013). Infants like food that is preferred by someone speaking their language rather than a foreign one (Shutts et al., 2009), children prefer to try a food endorsed by a peer rather than one endorsed by an adult (Frazier, Gelman, Kaciroti, Russell, Lumeng, 2012), and they prefer food (as well as toys, games, and clothing items) liked by someone of the same age and gender as themselves (Shutts et al., 2010). While there is not as much evidence for third-party preference inferences about food, children may infer individuals who have socially interacted to like eating the same things. Thus, it is possible that four-year-olds would expect shared preferences for both games and food based on information about positive interactions.

Another possibility is that children may view the joint activities as informative about one type of preferences but not the other. Previous research has demonstrated that children sometimes treat food and objects (including toys and games) as separate domains to which the same rules need not apply. Two-year-old children preferred an unseen book or toy—but not food—endorsed by a person with whom they had previously shared toy preferences (Fawcett & Markson, 2010b), preschoolers preferred to direct questions about foods to adults and about toys to peers (VanderBorght & Jaswal, 2009), and children preferred to learn about food, but not artifacts, from a model whose prestige was established for food and vice versa (Chudek, Heller, Birch, & Henrich, 2012). Finally, it is possible that the information about joint activities between two individuals only merits a general inference about their relation to one another, and does not

extend further into generalizing a preference from one party to the other, neither for foods nor for games.

Method

Participants. Twenty-four four-year-old children (12 female, M_{age} = 54.5 months, range = 4;0-4;11 months) participated in this experiment. Three additional children were excluded from analysis due to a side bias in responses, which was defined as 15 or 16 out of 16 choices on the same side. The sample size for Experiment 4 was increased from 16 to 24 participants before testing began, in order to compensate for the fewer number of trials that subjects viewed for each question compared to the previous experiments (four trials instead of eight).

Materials. The pictures of the child faces used in the previous experiments were used in a new arrangement: The central character was presented centered in the lower half of the screen, and the two possible targets appeared in the top right and left corners. Additionally, an item—a picture of a novel food or game—was presented below the central character's picture on each trial (see Figure 1.1). A new set of facts about joint activities and arbitrary links was prepared that were similar to Experiment 1, but shorter in length (see Appendix B). As before, joint activity facts were comprised of facts about helping and invitations and arbitrary link facts were either about proximity or commonality.

Procedure. Each subject completed the following blocks: 1) *friendship* (4 trials), to infer who the central character's friend was; 2) *playmate* (4 trials), to infer who the central character liked to play with; 3) *game preference* (4 trials), to infer who else liked the game preferred by the central character; 4) *food preference* (4 trials), to infer who else liked the food preferred by the central character. To prevent distraction and keep subjects on task, the experiment was conducted across two mini-sessions (friendship and playmate blocks in one session, and game

and food preferences in the other), separated by a five-minute play break outside the testing room. Order of sessions was counterbalanced across participants. (See Appendix C for more details about the design and procedure.)

Trial procedure. The structure of all trials, regardless of block type, was as follows:

1. The central character was shown together with an item (either a game or food) he or she liked. The character was named, but the item was not.
2. In order to invite the participant to attend to the item, the experimenter asked a question about it, e.g. "Do you think it's a breakfast food or a dinner food?" The experimenter reacted positively to the child's response.
3. Pictures of the other two characters and the hint images appeared one by one, and the experimenter told the subject facts about the central character and each of the possible targets (i.e. joint activity and arbitrary link facts). Manipulation questions were asked only for the first trials in each four-trial block, mainly to ensure children's attention.
4. The experimenter asked the test question with pictures of the three characters and the (food or game) item visible and hint images absent. The test question differed depending on the block. For instance, on a food preference trial, the experimenter said, "One of these girls also likes to eat this food like Maya. Which one do you think likes to eat it?"

Results

When asked to identify the central character's friend in the friendship block, subjects chose the character that had engaged in a joint activity with the central character 66% of the time. According to a binary logistic regression model containing a negative intercept, block as a fixed effect and subject as a random effect (glmer(Response ~ -1 + Block + (1|Subject)), this

performance is significantly better than chance (b = 1.09, SE = .37, z = 2.92, p = .003). The same analysis showed that when asked who the central character's preferred playmate was, participants chose the target more often than chance (M = 65%, b = 1.01, SE = .37, z = 2.71, p = .006). Regarding preferences, responses were above chance for games (M = 61%, b = .86, SE = .37, z = 2.35, p = .018), but at chance for foods (M = 46%, b = .19, SE = .36, z = .53, p = .59*)*.

A second mixed-effects binary logistic regression model including block type as a fixed effect and subject as a random effect was fitted to test several planned contrasts between blocks (glmer(Response ~ Block + (1|Subject)). First, I compared responses for friendship and playmate inferences; they were not significantly different from each other (b = .08, SE = .31, z = .26, p = .79). Second, game and food preference choices were compared to friendship inferences, and to each other. Responses for game preferences were not significantly different from friendship responses (b = .22, SE = .31, z = .72, p = .46), but responses for food preferences and friendship were different (b = -.90, SE = .31, z = -2.91, p = .004). Finally, game and food preferences were significantly different from each other (b = -.67, SE = .30, z = -2.22, p = .026).

Discussion

Replicating Experiment 1, the results from the friendship block once again demonstrate that four-year-olds infer friendship from reports of joint activities, but not arbitrary links. This finding is robust in that it extends to their inferences about a character's preferred playmate as well. Regarding preferences, children's responses differed depending on the kind of item they were asked about. For games, participants tended to think that the character that had engaged in a joint activity with the central character was likely to also like his or her preferred game. Consistent and comparable choices for the target in response to friendship, playmate, and game preference questions may reflect the link between these dimensions in four-year-olds' concept of

friendship. Regarding food, children did not choose systematically between the two possible targets, suggesting that they did not expect two characters that had preformed a joint activity in the past to share food preferences. This result, together with the lack of previous work pointing at a connection between food and friendship suggests that children's notions of friendship may not include any strong expectations when it comes to food. It is also possible that they would only infer shared food preferences between friends if the friendship is directly pointed out, rather than implied through behavioral information, as was the case here. In the current study, participants may have relied on general features of eating behaviors when asked to reason about the food preferences of peers. For instance, it is possible that preschoolers observe divergence in food preferences of social partners more frequently than in game preferences. In fact, they may even witness discord about what people like or dislike to eat within their own family. Further, while toys are more likely to be shared and parents typically endorse sharing of toys, food items are treated as having more exclusive owners, particularly in situations outside the home. Even at home, children may have noticed the individual manner in which food is divided such that each person is provided with a separate eating vessel and utensils. Considerations such as these may have influenced children and their unwillingness to ascribe shared food preferences between the individuals.

Interestingly, the lack of consistent choices for the target character on the food block rules out a possible alternative explanation for children's success when inferring friendship in the friendship block, as well as Experiments 1 and 3, and for their success on the game and playmate blocks. One low-level account of four-year-olds' pattern of responses in all these cases is the following: Joint activities such as those involving helping used here are inherently appealing to children, drawing their attention and subsequent preference for the character associated with it,

causing children to choose said character regardless of the actual question being asked. However, the lack of systematic responses for inferences about food preferences provides evidence against this deflationary account. Had children been simply picking that character regardless of the question they were asked, they should have done so for the food preference questions as well, which was not the case.

General discussion

The current set of studies show that at age four, children can use reports of joint activities or similarity between two individuals to infer friendship between them, and that they privilege joint activities over similarity. Four-year-olds can also use information about joint activities to make inferences about shared preferences, but do so only for preferences for games and not for foods. This result, together with their inference about playmateship based on the same information, further highlights the centrality of play to early friendships.

The current results raise a number of points. Children's high accuracy at remembering facts that they were told about the characters shows that they easily attend to and effortlessly remember reported behaviors between others—a necessary first step in using such information for drawing social inferences. Further, it is impressive that young children demonstrate the ability to infer friendship based on very little information in a third-person task, both here and elsewhere (e.g. Shutts et al., 2013). This emphasizes the significance of friendship as a canonical relationship that children can reason about not only when they are one of the involved parties, but also when they are not. It is important to note, however, that children did not find this task trivially easy. While performance was significantly above chance, it was by no means perfect. So while children succeed at this problem in its current form at four years, their abilities are still undergoing substantial improvement.

In addition to explicitly inferring who a character's friend was in Experiments 1 and 3, preschoolers in Experiment 4 also made similar inferences about a playmate. Jointly, these results paint an interesting picture of children's conceptions about friendship at this age. Joint activities initiated through either helping or invitation were behaviors that children associated with both friends and playmates. This collection both confirms and contrasts with previous work on children's theory of friendship. As numerous studies have shown, playing together is strongly associated with friendship, similar to previous studies (Furman & Bierman, 1983, Howes, 2009; Damon, 1979). Furthermore, the shared game preference inference in Experiment 4 is consistent with this. Children think of friends as being similar in their taste for games, perhaps because it would facilitate playing together. In contrast, descriptions about helping and invitations, as well as similarities in Experiment 2, were also associated with friendship, which extends previous research that doubted preschoolers' knowledge about these particular aspects of friendship (e.g. Selman, 1981).

The current findings provide the first evidence of young children using behavioral information about pairs of individuals to infer shared preferences. Previous work has demonstrated that children make shared preference inferences between individuals who are in the same social group (Diesendruck & haLevi, 2006), of the same kind (Diesendruck, Salzer, Kushnir, & Xu, 2013), and who have the same traits (Heyman & Gelman, 2000). In all these cases, children may have based preference inductions on essentialist notions by assuming that individuals from the same group or with the same internal characteristics share a deep, causal essence, and then generalized other characteristics, such as a preference for something, from one individual to another. In this task, however, a different mechanism must be at work. One possibility is that children relied on the dyadic relationship. They may have inferred the

friendship between the two individuals and further inferred shared game preferences based on that. In other words, children may have thought that two characters who had interacted in a joint activity were friends; friends like the same games, and so these individuals too must like the same games. If this is the case, children may have responded unsystematically to the food preference trials because they do not think of friends as sharing food preferences. In short, children may be unwilling to ascribe shared preferences to individuals based on information that is indicative of friendship, unless the domain in question is relevant to their notion of friendship. As current data do not direct test this, this question is left for future work.

The current findings demonstrate that children expect social interactions to have consequences for game, but not food, preferences. One possible reason may be that children expect similar food preferences only between group members, but not between individuals connected in subtler or more specific ways. Taste for food varies between different cultures and ethnicities, and children may have adapted to be sensitive to food preferences in line with group membership (Cashdan, 1998). Thus, children may be especially sensitive to any kind of group information with regard to food choices, even for groups such as gender that do not necessarily affect food preferences. In contrast to groups, they may see dyadic relationships as irrelevant to shared food preferences, especially in the current situation where the two characters' friendship is presented indirectly through behavior. Even so, this is not to say that inferences about shared food preferences are always immune to dyadic information. A recent looking-time study found contrasting results such that infants expected two individuals who had affiliated to like the same food, but not the same objects (Liberman, Woodward, Sullivan, & Kinzler, 2016). Given the many differences between the two studies (i.e. age of participants, making inferences about adults vs. peers, observing the interaction rather than hearing about it), it is not possible to know

what drives the divergence in findings, and further research is needed to explore different possibilities.

The current studies are limited in their ability to tell us about children's responses in more realistic situations. For instance, how would children at this age respond when asked to draw inferences about friendship based on observed rather than reported behaviors? And would they even notice similarities between individuals through observation? It is also unknown whether children are simply passive observers or whether, and if so under what conditions, they actively attempt to gather information about others' behaviors in order to solve the puzzle of figuring out the relationships around them. Additionally, the current study is limited in providing insight into children's thinking about friendship in other cultural contexts, especially ones in which the opportunities for non-kin relationships are less available. More generally, there is much we do not yet know about how young children map behaviors onto social relationships. Future work should explore whether children view different social behaviors as more or less indicative of specific relationships, and whether they view any behavior-relationship mappings as one-to-one. Answers to such questions would be highly informative because the problem under study is a consequential one. The social world is complex and nuanced, and individuals are connected to each other in myriad ways. In time, uncovering relationships and understanding their implications is one of the key achievements that allow children understand their social world.

Paper 2:

Preschoolers' recognition of and inferences from observed imitation

Narges Afshordi and Susan Carey

Introduction

Imitation, repetition of an observed action after a delay, is ubiquitous in human life. An important social behavior, imitation has early roots. Neonates imitate facial and manual gestures soon after birth (Meltzoff & Moore, 1977; although for opposing accounts see Jones, 2009; and Oostenbroek et al., 2016). While this proclivity for imitation is limited, copying abilities operate with stronger force in the second year of life, and by 14-18 months infants reliably imitate a host of behaviors including, but not limited to, vocal utterances and gestures (e.g. Jones, 2007). Imitation is significant because it allows infants and young children to learn about the world around them, both physical and social. Many cultural artifacts such as a simple spoon can be opaque to a novice, and imitation is one method through which children learn how to use such objects in daily life, and how to perform culturally determined behaviors (e.g. Gergely & Csibra, 2006). Within the boundary of imitative learning, children's choices show an interesting range. From early on, infants are efficient in their imitation when learning skills. They copy selectively, choosing to forgo repeating actions that are accidental or mistakes, and opting for efficient ones instead (Meltzoff, 1995; Carpenter, Akhtar, & Tomasello, 1998; Gergely, Bekkering, & Kiraly, 2002). However, when it comes to imitation in a more social context and involving a number of steps, children's choice of what to copy shows divergence from their selectivity elsewhere. Children are more likely to imitate a socially responsive partner than one who is not available and cannot communicate with them (Nielson, Simcock, & Jenkins, 2008), and they are more likely to imitate specific steps towards an action when pedagogical cues are used (Kiraly, Csibra, & Gergely, 2004). Further, children can be so committed to correctly imitating demonstrated actions that they sometimes engage in imitation of causally unnecessary ones (Horner & Whiten, 2005). Overimitation, this excessive loyalty to a model's actions, is so extreme that children do

so even when it hinders their chance of winning a game against a competitor by slowing them down (Lyons, Damrosch, Macris & Keil, 2011). One explanation for this phenomenon is children's motivations to be like others and to demonstrate that they are willing to do things in a culturally-sanctioned way (Over & Carpenter, 2013). Thus, even though imitation goes a long way towards helping infants and children learn, it is often deeply steeped in social goals.

Critically, imitation has social implications over and above learning. Mimicry, the subtle and often unconscious imitation of others' behaviors, has been theoretically touted as the 'social glue' that has been consequential in human evolution by fostering affiliation (Lakin, Jeffries, Cheng, & Chartrand, 2003). In experimental work, Chartrand and Bargh (1999) demonstrated that adults unconsciously mimic the mannerisms of a stranger they are partnered up with for a task. Further, following an interaction involving subtle mimicry by a confederate, adults report the interaction as having been more positive and gone more smoothly, and even prefer the interaction partner more, when compared to an interaction where no imitation took place (Chartrand & Bargh, 1999; also see Stel & Vonk, 2010). In addition to dyadic interactions, imitation has consequences for conformity and group cohesion, even bringing about more prosocial behavior. Adults were more likely to help an experimenter when the experimenter had previously mimicked them than not (van Baaren, Holland, Kawakami, & van Knippenberg, 2004). Further, participants who had been mimicked were also more helpful to a new experimenter who had not been the mimicker, and more likely to engage in a general prosocial act in support of a charity. Interestingly, the links between affiliation and imitation are not limited to adulthood, but rather appear in early childhood as well. An analog of van Baaren et al.'s (2004) study conducted with eighteen-month-old toddlers found that they were more likely to help someone after having been imitated (Carpenter, Uebel, & Tomasello, 2013), even if the

person needing help had not been their imitator. This finding demonstrates that imitation carries positive and prosocial connotations for young children. Other evidence is provided by a study showing that five-year-old children copy with more fidelity after being primed with social exclusion (Over & Carpenter, 2009), suggesting that a drive towards social engagement is eased by copying others.

Beyond first-person imitation, infants are interested in being imitated. By 14-18 months, they reliably notice when others are copying them (Agnetta & Rochat, 2004), and explicitly carry out 'testing' behaviors by performing unlikely sequences of actions, and monitoring the potential imitator in order to ensure that they are being imitated (Meltzoff & Moore, 1999). Furthermore, third-party observations of imitation are also informative about the social world. Observed imitation provides the observer with cues to the relationship existing between the involved parties such as their relative status, or inferences about psychological traits. For instance, although unconscious of the mimicry itself, adults viewed an individual who mimicked a condescending interviewer as less competent than both a non-mimicker of the same interviewer and someone who mimicked a cordial interviewer (Kavanagh, Suhler, Churchland, & Winkielman, 2011). Although the range of sophisticated inferences adults make on the basis of observed mimicry is out of reach for younger viewers, they are nonetheless capable of drawing impressive social inferences from viewing explicit and clear imitation. Using a looking-time paradigm, Powell and Spelke (2017) showed infants aged around four months displays of cartoon characters involving imitators and non-imitators. Infants' looking times showed that they expected an imitator to approach the target of its imitation, but they were surprised to see a third character who had not copied the target approach it. In a separate condition, they found that infants held no expectations about targets of imitation approaching their imitators compared to

approaching a third, non-imitator character. Further, infants preferred an imitator over a non-imitator when offered the chance to choose between physical replicas of the characters. These findings indicate that even very young infants see imitation as a positive act and associate it with affiliation from the imitator towards the target.

In work with older children, Over and Carpenter (2014) examined the link between imitation and affiliation in a preschool-aged group. They showed children videos of three individuals, one of whom consistently copied four actions performed by one of the other individuals (sitting positions, object choice, action on an object). Children aged four and five years were asked whom the imitator liked more. Five-year-olds, but not four-year-olds, responded by choosing the target of imitation significantly more than expected by chance, and the majority of those responding correctly provided explanations related to copying. In a second experiment, children of the same ages saw a video involving only two people with one imitating the other's actions four times. Here, children were asked to identify the 'boss'. Again, five-year-old children succeeded at choosing the target of imitation, while four-year-olds did not. As Over and Carpenter suggest, making explicit social inferences from observed imitation may require more experience and learning than is available to four-year-olds. However, this does not seem like a satisfying account of four-year-olds' failure, given that infants expect an imitator to affiliate with the target of its imitation (Powell & Spelke, 2017). Thus, the question arises why infants did and four-year-olds did not infer affiliation from observed imitation.

While puzzling, this seeming discontinuity between infants and preschoolers is not an isolated event. A body of work has revealed that infants demonstrate correct expectations in looking-time experiments that require false belief understanding (e.g. Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007). In contrast, preschoolers do not pass the classic theory

of mind task, which also requires false belief understanding, until age four (Wellman, Cross, & Watson, 2001). Another case has to do with naive physics and understanding that objects cannot pass through solid barriers. Once again, looking time experiments show very young infants harboring the right expectations (Spelke, Breinlinger, Macomber, & Jacobson, 1992), but children facing a similar conceptual task fail to reach in the correct location for an object faced with a barrier until three years of age (Berthier, DeBlois, Poirier, Novak, & Clifton, 2000). Both of these contradictory sets of findings are in need of explanatory accounts (for a discussion on the former case, see Saxe, 2013; for a discussion on the latter, see Carey, 2009). Further, it is unclear whether they are two occurrences of the same phenomenon or seemingly similar, yet completely unrelated cases.

In light of all this, the current paper explores the apparent discontinuity between infants and preschoolers in inferring affiliation from observed imitation. It is possible that this discontinuity is robust, like the two cases mentioned above. Alternatively, it is possible that preschoolers, even three-year-olds perhaps, could succeed at making the affiliation inference from imitation under the right circumstances. In Over and Carpenter's experiment, children were told upfront that they would watch a video and would then be asked who the central character liked more. This heads-up information highlighted the social inference of interest and could have prompted participants to watch the scenes with an eye towards extracting a basis for the inference. Given the results, five-year-olds seem to have been more successful at doing this than their younger counterparts. Therefore, even if it did help, it was not equally helpful to four- and five-year-olds. What the present experiments explore, however, is whether directly highlighting the basis of the inference--imitation--can lead children to succeed at a possibly even earlier age. Experiment 1 tests whether three-year-olds have an explicit concept of imitation and whether

they can access it to identify the target of imitation in scenes they view. This is important to know since highlighting imitation would not be useful, or even possible, unless children already possess the concept. Experiment 2 tests three- to five-year-olds (and adults) on their spontaneous inferences without any emphasis on imitation. Experiment 3 and 4 take separate routes towards highlighting imitation and test their effects on inferences. Experiment 3 does this through propositional scaffolding in the midst of the imitative actions by drawing attention to the proposition "the copier is copying the target". Experiment 4 does so by priming the concept of imitation ahead of the task. Overall, these experiments can shed light on whether the failure reported regarding preschoolers is robust.

Finally, the current studies extend previous work in two ways. The first is by asking children to predict an imitator's future action, testing whether they expect the imitator to continue imitating her previous target as opposed to her previous non-target. This inference also relies on encoding the imitation and the parties involved, and may therefore pattern alongside affiliation inferences. Alternatively, it is possible that this inference would be easier to make and children might succeed on this, even if they do not infer affiliation successfully. The second extension comes in the form of testing whether the type of action (bodily movements, vocalizations, object preferences) that an individual imitates affects children's performance.

Experiment 1

The goal of this experiment was to examine whether young preschoolers at age three have an explicit concept of imitation. We tested three-year-olds in order to find out whether further manipulations highlighting imitation would be worthwhile with this age group. In order to test this, we asked children to identify the target of an imitator whose actions they had watched. Given infants' and children's own proclivity towards imitation and their precocious

expectations about imitators in third-party contexts (Powell & Spelke, 2017), one hypothesis is that three-year-olds have an explicit concept and would have no trouble encoding imitation as observers. An alternative hypothesis is that four-year-olds' failed at making inferences from observed imitation (Over & Carpenter, 2014) because they failed to encode the imitation as a result of not having an explicit concept for it. If this is true, three-year-olds in this experiment would also fail. Additionally, this experiment asked if children could subsequently answer whether the imitator preferred the target of her imitation, and whether she would continue to imitate the same target in the future. If three-year-olds do encode the imitation, they may possibly make inferences based on it too. Alternatively, they may not notice the imitation, or notice it and yet fail to make further inferences. Following the sample size used by Over and Carpenter (2014), this experiment and all subsequent ones included forty participants (in each age range, where applicable).

Method

Participants. Forty three-year-olds (M = 42.7 months, 21 female) participated in this experiment. An additional six participants were excluded due to not completing the task (4), or being extremely distracted and not watching the stimuli consistently (2). Children were predominantly recruited from a database of interested families at the Harvard Lab for Developmental Studies. The families in the database are mostly White and middle-class (79% European American, 15% Asian American, 6% African American). Parents received travel reimbursement and children chose a small toy as a prize after their participation. A small number of participants (n=3) took part in the study at the Boston Children's Museum. In accordance with museum guidelines, these children and parents did not receive a prize or travel reimbursement.

Materials. The experiment was conducted using a 13-inch Macbook laptop, and stimuli were presented in Keynote. The keyboard was covered with a binder to prevent distraction. Slides were advanced using a wireless mouse that was held out of the child's view by the experimenter.

A number of video stimuli involving three characters were created using the software package Poser. The characters involved were three naturalistic-looking, computer-generated women, who were shown standing equidistant from each other in a horizontal line (see Figure 2.1). They were introduced to children with proper names (from the left, Jenny, Alice, and Zoe). If the participant's name was the same of any of the three characters, that character was given an alternative name. Throughout the experiment, the characters performed different bodily actions consisting of moving their head, arms, leg, or torso in different ways (e.g. bending sideways, knee lift). At the start of each action, a bell sound effect played and a green bar appeared beneath the character's feet, where it remained for the duration of the action. The individual performed her movement following this attention-getter sound and repeated it once more, e.g. lifting her knee twice. These measures were taken to ensure that children attended to the correct character at each point and did not miss any of the actions, especially more subtle ones such as head movements. Whenever all three characters were present in a trial, the central character turned her head to watch each of the other two individuals as they performed an action.



Figure 2.1. The three characters shown in the stimuli. The central character is the imitator, and the target and non-target are on the sides.

Procedure. Children were brought into a testing room (or testing area at the museum) and seated at a table. Parents were allowed to be present during testing if the child preferred, but were instructed to remain silent. The experimenter showed children a slide with a picture of a cartoon dog and flower, and asked the subject to point to each one. This was done as a warm-up and to make children feel comfortable pointing to the screen. Participants then viewed a number of videos in the following order:

Introduction. The central character, who would later on turn out to be the imitator, appeared and was introduced, e.g. "This is Jenny. Let's see what she does." She then performed all the actions she would have the chance to imitate later on. The purpose of this phase was to demonstrate to children that the imitator was able to perform all the actions, and that her future imitations were by choice rather than lack of ability. The experimenter narrated the actions as they were performed (e.g. "Jenny moved her arms."), in order to familiarize children with the wording that would be used to refer to the actions in future questions.

Side-character action trials. The two side characters, who would later become the target and non-target of imitation, appeared and were introduced with their names. Children then viewed two warm-up trials. On each trial, the two characters performed different actions. After each trial, the child was asked who had performed each action. For instance, the experimenter asked "Who moved her leg? Was it Alice or was it Zoe?" while pointing to the relevant characters at she spoke. Children could respond by either pointing to one of the characters or using their names. Positive feedback was provided for correct answers. If the child answered incorrectly for either of the two actions, both actions were shown again and the questions were asked a second time. Starting at this phase, a rotating star with a sound effect was used to separate trials.

Imitator action trials. Children saw two trials in which all three characters were present. The central character first turned her head toward each of the side characters when they performed their actions, and then performed one of those actions identically. The imitator imitated the same character on both trials. Children's memory of the copier's actions were probed, e.g. "Did Jenny move her head or did she move her leg?" Once again, positive feedback was provided for correct answers, and the actions and question were repeated following incorrect responses.

Imitation. Participants viewed another three trials involving all characters in which the central character, aka the imitator, watched the other characters act, then imitated the same target of imitation as before. These trials were in essence the same as the imitation action trials, but no memory question was asked after them. Importantly, the specific actions performed by the characters across the five trials were unique and not repeated.

Test questions. Children were asked three test questions:

Imitation. At the end of the last Imitation trial, children were asked who the central character had imitated. The experimenter said, "Who was Jenny copying?" and provided neutral feedback to their response.

Affiliation. The experimenter asked the child to infer who the copier preferred, "Who do you think Jenny likes best?" Children were once again provided with neutral feedback.

Future target prediction. The experimenter showed the participant an incomplete trial in which the side characters performed new actions while the imitator was watching, but the imitator did not move after the green bar appeared beneath her, i.e. when it was her turn to act. At this point, the experimenter asked the child to predict what the imitator would do, for instance saying, "What happened? Can you help me figure out what Jenny is going to do? Do you think she's going to move her arms like Alice, or bend sideways like Zoe?" The experimenter responded to the child's answer with neutral feedback.

The imitation question--"Who was Jenny copying?"--was always asked first. The order of the affiliation and prediction questions, however, was counterbalanced across participants. For half the participants, the affiliation question was asked immediately after the copying question. These children then saw the incomplete trial and answered the prediction question. For the rest of the participants, following the imitation question the experimenter first showed the incomplete trial and asked the prediction question, then asked the affiliation question immediately after.

Design. Each participant viewed one of eight counterbalanced orders, for which the identity of target of imitation, and the order of the specific actions performed during the imitation and test trials were counterbalanced across children. The order of action by the side characters (e.g. target acting first or non-target acting first) was counterbalanced across trials for each participant.

Coding. On side-character action trials, responses were coded as correct only if the child answered correctly for both characters' actions the first time. On imitator action trials, responses were similarly counted as correct if the child answered correctly the first time they were asked. Regarding all test questions, a choice for the target of imitation was counted as correct.

Results and discussion

Analyses in this experiment and all of the following were carried out using the lme4 package in the software R (2017). Figure 2.2 shows responses for all four experiments.

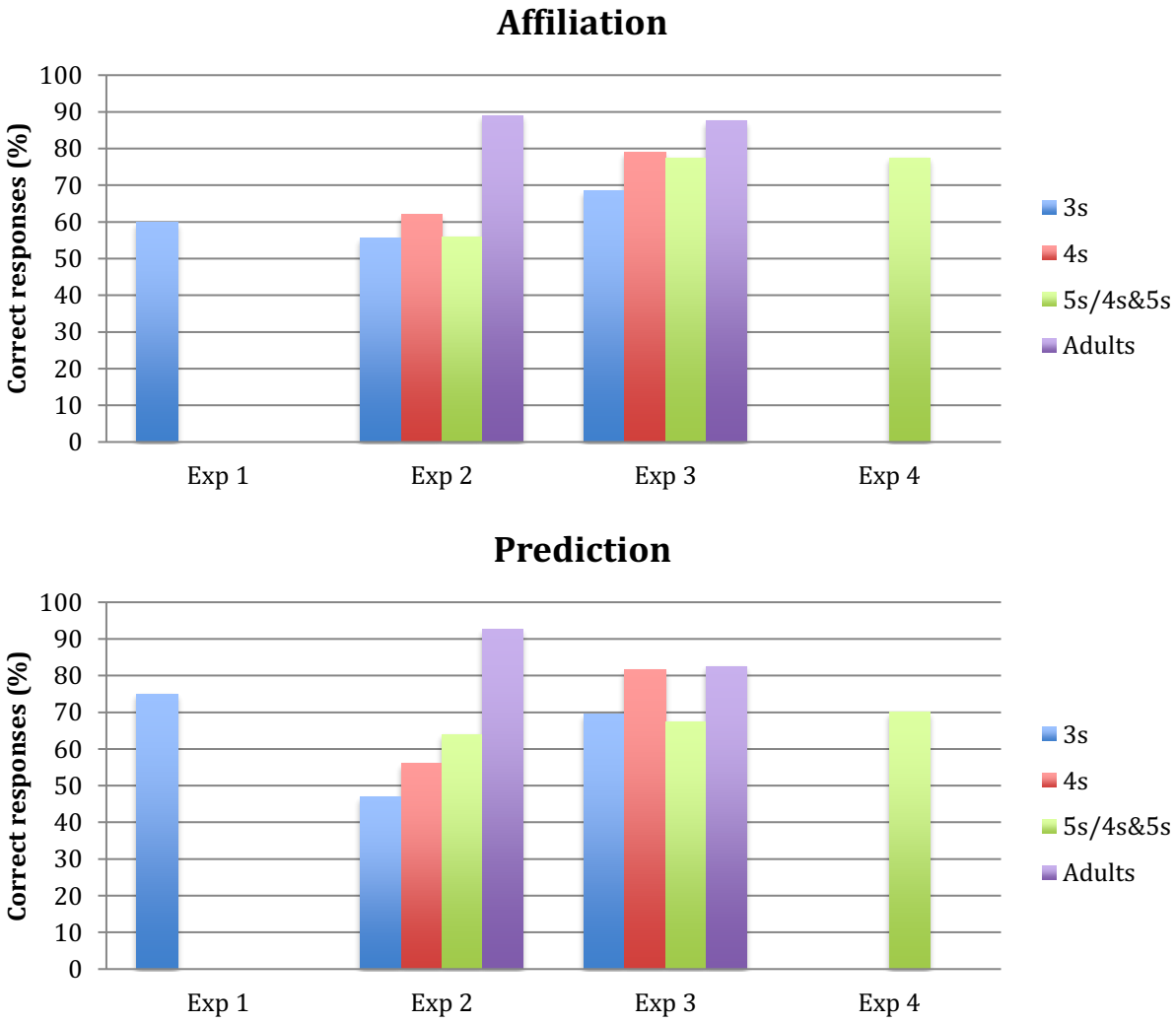


Figure 2.2. Affiliation and prediction responses across all four experiments separated by age. The data from Exp 4 corresponds to a mixed sample of four- and five-year-olds.

Participants responded correctly to the side-character action trials with 85% accuracy, and to the imitator action trials with 87.5% accuracy. To compare these performances against chance, mixed-effects binary logistic regression models including the intercept as the only fixed effect and subject as a random effect were run (glmer(Response ~ 1 + (1|Subject))). Both sets of responses were significantly higher than expected by chance performance (side-character action trials: $b = 1.73$, $SE = .31$, $z = 5.54$, $p = 3.03e\text{-}8$; imitator action trials: $b = 2.45$, $SE = .86$, $z = 2.86$, $p = .004$). In other words, three-year-olds easily remembered the actions that each character had performed, and indicated this accurately for the imitator, in addition to the (future) target and non-target characters.

In response to the imitation question (i.e. "Who was Jenny copying?"), 77.5% of participants chose the target character. A binomial test revealed this performance to be significantly better than chance, (31/40 correct, $p = 0.0007$). In response to the affiliation question (i.e. "Who does Jenny likes best?"), 60% of children responded correctly, which was not distinguishable from chance, (24/40 correct, $p = .27$). Finally, 75% of the sample responded correctly to the future target prediction question (i.e. who Jenny would perform the same action as), which was significantly better than expected by chance, (30/40 correct, $p = .002$).

Since the order of affiliation and prediction questions was counterbalanced, a binary logistic regression model including question order (affiliation first or second) was run on the data for both questions. While there was no effect of question order on prediction responses, there was one for affiliation responses ($b = 2.42$, $SE = 1.21$, $z = 1.99$, $p = .046$), such that performance was higher when affiliation was asked about first. Further inspection revealed that affiliation responses were significantly better than chance when affiliation was asked about first (15/20 correct, $p = .041$), but not when asked about second (9/20 correct, $p = .82$).

Participants' success in answering the imitation question indicates that three-year-olds have an explicit concept of "copying", and can recognize the target of copying behavior when viewing instances of imitation as third-party observers. Their correct pattern of responses for the prediction question, and their incorrect responses as a group for affiliation, are also noteworthy. The failure on affiliation is in line with Over and Carpenter's findings. Nonetheless, it is not clear whether inferring continued imitation is generally easier than inferring affiliation. The next experiment tests this possibility with a wider age range (three- to five-year-olds, and adults).

Additionally, it is unclear whether responding to the imitation question ("Who was Jenny copying?") could have influenced children's responses to the prediction and affiliation questions that followed it. This is possible, and in fact, the imitation question can be thought of as a manipulation that highlights both the general concept of imitation, and the specific proposition that the central character imitated the target. That is to say it is not clear what mechanism allowed children to correctly predict who Jenny would copy in the future. They may have done so due to having encoded the imitative behavior while watching the actions. Alternatively, they could have recovered the information from memory of the events when they were asked the imitation question, and hence changed their construal of the scene from a general repetition of actions to a stable pattern of imitation. Thus, the question arises whether children's spontaneous encoding of the events is robust enough for them to make inferences when imitation is not highlighted as it was here by the imitation question. In order to answer this, the following experiments refrain from asking the imitation question.

Finally, although three-year-olds in this study failed on the whole to infer affiliation from the scenes they viewed, the ordering of the questions mattered. Thus, the next experiment varies

the order of the questions across two separate conditions to give children the best chance of succeeding on each one.

Experiment 2

This experiment asked if preschoolers (and adults) can use observed imitation to spontaneously infer that an imitator prefers the target of her imitation, and that she will continue to imitate her target compared to another, never-imitated individual. Three-year-olds' failure in the previous experiment and four-year-olds' failure in previous work (Over & Carpenter, 2014) suggests that three- and four-year-olds may once again fail to answer the affiliation question correctly, but that five-year-old children may succeed given their previously reported success (Over & Carpenter, 2014). Both children and adults could succeed in predicting continued imitation in light of three-year-olds' correct response to this question in the previous experiment. Alternatively, it is possible that the imitation question in the previous experiment influenced their predictions, and that children would not succeed in responding correctly without it. Regardless, older children may be more successful at the same inference, even without the imitation question.

Given the question order effect found for affiliation in the previous experiment, participants were run in two conditions varying the order of the test questions. As a result, this experiment included 80 subjects in each of four age ranges, three-, four-, five-year-olds, and adults. The doubled sample size at each age was ensured that enough power was available to find an effect in each of the conditions, and any possible differences between them. Adults were tested to ensure that the inferences that were asked about were reasonable given the stimuli, and to test for possible change across development. As the imitation in the animations is explicit, adults should succeed at answering both questions.

Method

Participants. Eighty three-year-olds (M = 41.1 months, 38 female), 80 four-year-olds (M = 53.4 months, 45 female), and 80 five-year-olds (M = 66.1 months, 34 female) from the same populations as before participated in this experiment. Nineteen participants took part in the study at Boston's Children's Museum and the remainder came into the lab. An additional six children were excluded for being extremely distracted (4), and parental interference (1), and not completing the task (1). Additionally, 80 adults were recruited through Amazon's Mechanical Turk and participated in the study online (M = 32.4 years, 33 female).

Materials. The same set of animated stimuli from the previous experiment was used.

Procedure. The procedure up to the test phase was the same as Experiment 1, meaning that participants were first introduced to the central character, Jenny, and watched her perform a number of actions. They then viewed two trials in which the side characters performed actions, and participants were asked memory questions about the actions (i.e. side-character action trials). Next, they viewed two trials in which Jenny imitated one of the other characters and answered questions about what action Jenny had performed (i.e. imitator action trials). Finally, they saw three other instances of Jenny imitating the target. In the test phase, children were asked the prediction and affiliation questions. The order of the test questions was counterbalanced, meaning that half the participants answered the prediction question first, while the other half answered the affiliation question first.

Adult participants taking part in the study online were informed at the start that they would watch a number of animations designed for children and answer questions about them. They were also told that the task was straightforward and aimed at discovering their intuitions. The general procedure was the same for the online version of the study as the one the children

completed. Everything that the experimenter said out loud for child participants was displayed at the bottom of the screen. Questions were asked using forced-choice options.

Coding. Responses were coded as before. Missing data points were allowed for the second question. In other words, a participant's data were included in the analyses if they had a valid response for the first question they were asked, but not the second one. Overall, eight participants did not have responses available for their second test question (four refused to answer and four due to experimenter error).

Results and discussion

Adults. Performance on the side-character action trials and imitator action trials were respectively 92.5% and 99% accurate, both significantly above chance according to mixed-effects binary logistic regressions including only the intercept as a fixed effect and subject as a random effect (side-character action trials: $b = 2.5$, $SE = .3$; $z = 8.4$, $p = 2e\text{-}16$; imitator action trials: $b = 5.1$, $SE = 1$, $z = 5$, $p = 4.35e\text{-}07$). Thus, adults remembered the actions each character performed with high accuracy, demonstrating that they were attentive to the task.

Adults' performance on the prediction question was 93% accurate and significantly higher than chance according to a binomial test (74/80 correct, $p = 5.4e\text{-}16$). Their responses to the affiliation question were 89% accurate and also better than chance (71/80 correct, $p = 4.4e\text{-}13$). Binary logistic regression models were run to test for effects of question order (affiliation asked first or second) on affiliation and prediction responses. Neither found an effect of question order. In short, adult participants answered both test questions overwhelmingly correctly. This confirms that the imitation scenes were interpretable to adults, and that the affiliation and prediction inferences were easy for them to make. Further, it demonstrates that adults could

answer the questions spontaneously, even without the support that the imitation question might have provided children in Experiment 1.

Children. Children's responses to the side-character action trials were 90.1% accurate, which was significantly better than chance performance according to a binary logistic regression model including only the intercept and the subject as a random effect (b = 2.2, SE = .15, z = 14.2, p = 2e-16). Responses were similarly high for imitator action trials at 94.8%, and better than chance (b = 7.8, SE = 1, z = 7.8, p = 6.7e-15).

Responding to the prediction question, 56% of all children provided the correct response, which was not significantly different from chance as tested by a binomial test (130/234 correct, p = .10). With regard to the affiliation question, 58% of total responses were correct, which was significantly better than expected by chance (138/238 correct, p = .017). Binary logistic regression models including age (in months as a continuous variable) and question order (affiliation first vs. second) were run on both sets of responses, but failed to find any main effects or interaction for either prediction or affiliation. To summarize, preschoolers inferred that the imitator liked the target of her imitation, but not that she would go on to perform another action that was the same as hers in the future.

Children's failure on prediction and success on affiliation here was in contrast to Experiment 1, where three-year-olds succeeded on prediction and failed on affiliation. To get a better sense of the three-year-olds' performance here, their responses were analyzed separately. This revealed that they were at chance on both affiliation (55.7% accurate, 44/79 correct, p = .37) and prediction questions (46.8% accurate, 36/77 correct, p = .65). In order to compare three-year-olds from the two experiments directly, binary logistic regression analyses were conducted. Regarding affiliation, there was no difference between responses in the two experiments,

suggesting that three-year-olds were equally unsuccessful in both experiments. In other words, although the whole 3- to 5-year-old sample was better than chance in the current experiment, three-year-olds separately were not. Even for the whole sample, the effect was significant, but small in magnitude. This was hinted at by the confidence interval from the overall binomial test CI = [.51 .64], and the fact that the lower bound of the interval was close to 0.5. It is furthermore supported by a comparison between three-year-olds in Experiment 1 and all the children in Experiment 2, which found no difference between experiments.

With regard to the prediction question, the analysis found a significant difference (p = .004) such that children were more successful in the first experiment than here. In other words, three-year-olds were above chance on the prediction question in Experiment 1, and significantly better than three-year-olds here, who responded at chance.

In summary, preschoolers inferred affiliation, but the magnitude of success was small, and they did not predict the target of the imitator's next imitation. Given that there was not an age effect, this failure was the case for all ages, including five-year-old children. Further, not responding correctly to the prediction question stands in contrast to three-year-olds' success on the same question in Experiment 1. This finding suggests that the imitation question in Experiment 1 did in fact influence children's responses to the prediction question. Further, the current results demonstrated that unlike Experiment 1, inferring affiliation is not necessarily more difficult than predicting the target of future imitation.

Experiment 3

The goal of this experiment was to test whether highlighting imitation would lead to improvements children's inferences about affiliation and continued imitation of the same target. The manipulation in this experiment takes place during the first two instances of imitative

behavior, and comes in the form of a wording change to the memory questions asked in the imitator action trials in the previous experiments. Instead of asking, for instance, "Did Jenny move her arms or her leg?", we now ask, "Did Jenny move her arms *like Alice* or her leg *like Zoe*?" This propositional scaffolding highlights imitation without explicitly referring to copying by guiding children towards the proposition, "Jenny copied Alice." As the scaffolding takes place early on, children have the opportunity to watch a number of imitative actions with the proposition in mind afterwards. It is possible that this manipulation is sufficient for children to better succeed at the inferences. Alternatively, it is possible that the proposition is not sufficient on its own and that children need explicit priming of the concept, or of both that and the proposition to succeed. A further possibility is that no amount of emphasis on imitation would lead preschoolers to robustly succeed on both inferences.

Another main purpose of this experiment was to test whether children's inferences would be influenced by the type of action they viewed being imitated. One hypothesis is that inferring affiliation from observed imitation relies on a general understanding of imitation, and does not differ across various forms of actions. An alternative hypothesis is that children's ability to understand and make inferences about imitation might depend on the action type, and that the imitation might be more transparent for some actions types over others. If this is the case, it is possible that preschoolers could demonstrate differentiated performance on scenes involving different type of actions. The actions used here were bodily movements, which were used in the previous experiments, as well as vocalizations and object preferences.

Method

Participants. Forty three-year-olds (M = 43.0 months, 20 female), 40 four-year-olds (M = 53.3 months, 20 female), and 40 five-year-olds (M = 65.5 months, 18 female) from the same

laboratory population as before participated in this experiment. An additional five children were excluded due to not completing the task. Forty adults participated in the experiment online (M = 37.2 years, 20 female).

Materials. Three different action types were used in three separate sequences: bodily movements, vocalizations, and object preferences. The stimuli from Experiments 1-3 involving the three characters were used for movements. Two additional sets of animations were prepared for each of the remaining action types:

Vocalizations. This sequence involved three cartoon characters with different voices. Utterances were recorded by three female adults, who were chosen from an initial pool of six speakers because their voices were relatively distinct, and therefore distinguishable from each other. Each person recorded sixteen disyllabic utterances (e.g. shay-shay). Consonants were used only once across the set, but there were two utterances using each of four vowels: /ä/ (la-la, ba-ba), /ü/ (soo-soo, roo-roo), /ā/ (tay-tay, shay-shay), or /ē/ (ghee-ghee, chee-chee), /aɪ/ (jie-jie, kie-kie), /ɔ/ (faw-faw, paw-paw), /oʊ/ (doh-doh, zoh-zoh), and /aʊ/ (vow-vow, mow-mow). Visually, the cartoon characters were circular faces with eyes and a smiling mouth. The characters were identical except for their face colors (red, yellow, and green). They were presented in a horizontal line equidistant from each other (see Figure 2.3). When 'speaking', each character jumped up and down twice, and the vocalization played as she jumped. The central character's eyes moved to the left and right to watch the movements of the side characters at relevant points during the animation. Given that their voices were female, the characters were assigned female names (Hannah, Emily, and Katie). The animations were created in Keynote.

Object preferences. Three cartoon characters were involved in this sequence. The characters were presented with heads and arms. Their facial features were identical to each other,

but their faces varied in color (pink, orange, and teal) and shape (oval, square, and diamond). Their formation on the screen was as a triangle with the central character above the other two (see Figure 2.3). Each action in this sequence involved making a choice between two objects. The objects were of the same category but distinct in their appearance (e.g. top hat and cowboy hat). For each character's action, the objects moved into the screen from the side and landed below his arms. Then, the objects jiggled one by one accompanied by a ringing bell sound effect. The character looked at each object as it was moving, then picked up one of the objects and acted on it (e.g. wore the glasses, held the umbrella above their head). Following this, the character looked straight ahead briefly, before both the chosen and unchosen object disappeared in anticipation of the next character's action. In scenes with all three characters present, the central character's eyes looked toward the other character who was acting at the relevant point. These characters were assigned male names (from the left, Tommy, Ben, and Jason). The animations were created in Keynote.

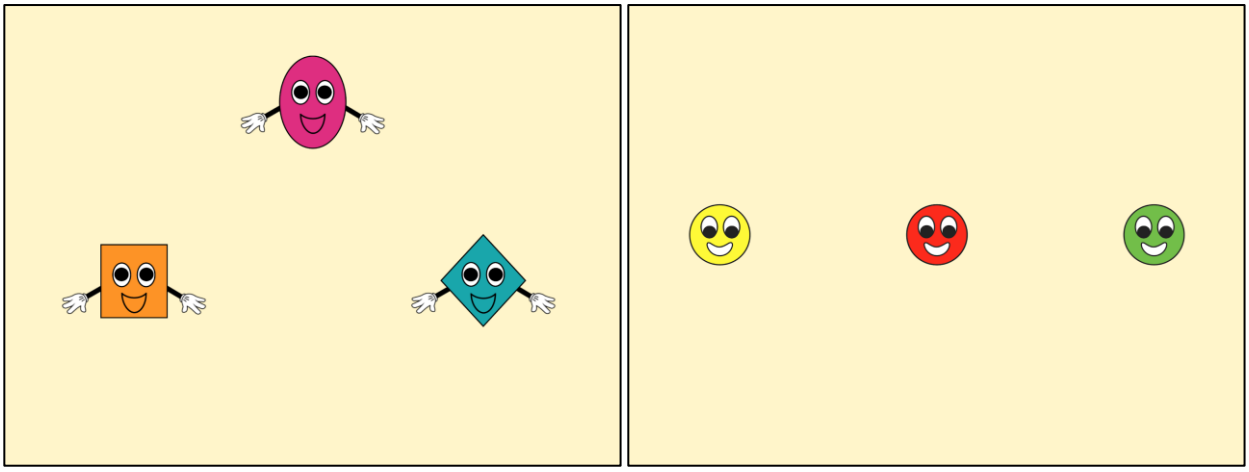


Figure 2.3. The characters involved in the object preference sequence (left) and the vocalization sequence (right) in Experiment 3.

Procedure. The general procedure for each sequence was highly similar to that of the previous experiments, with the main difference in the wording of three-character warm-up

questions. The order of sequences was counterbalanced across participants. The structure of each sequence was as follows:

Introduction. The central character, aka the copier, appeared and was introduced, e.g. "This is Ben. Let's see what he does." He or she then performed all the actions he or she would later get a chance to imitate later on. For the preference and movement sequences, the experimenter narrated the actions as they were performed (e.g. "Ben wore that hat," "Jenny moved her arms.") In the vocalizations sequence, repeating each vocalization was deemed unnecessary, so the experimenter simply said "Emily says different kinds of things" after the character had said a number of utterances.

Side character action trials. The two side characters (i.e. target and non-target) were introduced with their names, and children watched two warm-up trials involving them. The imitator character was not present. After they performed their respective actions, the child was asked who had performed each action. For instance, the experimenter asked "Who wore this mitten? Was it Tommy or was it Jason?" Positive feedback was provided for correct answers. The actions were shown again and the questions were repeated if the participant answered incorrectly.

Propositional scaffolding. Children viewed two trials in which the central character imitated one of the side characters, and were asked about the imitator's action. These trials were analogous to the imitator action trials in the previous experiments. Importantly, however, the wording of the question was different. On the movement trials, for instance, the experimenter said, "Did Jenny bend backwards like Alice, or move her arms like Zoe?" This was in contrast to how the question was worded before, "Did Jenny bend backwards or move her arms?" Correct

answers were rewarded with positive feedback, while the actions and question were repeated in cases of incorrect answers.

Imitation. Participants watched three further trials involving all three characters in which the imitator watched the side characters act, then imitated the target.

Test. Children viewed an incomplete trial in which the copier watched as each of the side characters performed their actions, but did not perform any action when his or her turn came. Participants were then asked three types of test questions:

Prediction question. The experimenter asked the child to predict what the imitator would do, e.g. "What happened? Can you help me figure out what Emily is going to say? Do you think she's going to say shay-shay like Hannah, or do you think she's going to say roo-roo like Katie?" The experimenter responded to the child's answer with neutral feedback.

Affiliation question. The experimenter asked the child to infer who the imitator preferred, e.g. "Who do you think Emily likes best?" Children were once again provided with neutral feedback.

Explanation. After the test questions for the last sequence, whichever it happened to be for a particular participant, the experimenter asked the child to provide their reasoning behind their choice on the affiliation question.

As the task was quite long and repetitive, all children viewed an unrelated animation that depicted a dominance relationship between two cartoon characters and were asked some questions about it after completing either one or two sequences.

Design. The order of sequences, and identity of target of imitation for each sequence was counterbalanced across children. For each subject, the side of the target across the sequences was

counterbalanced. Within each sequence, the order of action by the side characters (e.g. target acting first or non-target acting first) was counterbalanced across trials.

Coding. A participant's responses were excluded from the final dataset if they had failed to provide an answer to the first test question, i.e. prediction question, for at least two out of the three sequences, but kept in the dataset otherwise. Coding for side-character action trials and propositional scaffolding trials was similar to previous experiments, such that responses were counted as correct if children answered correctly the first time the questions were asked. Regarding the test questions, a choice for the target of imitation was counted as correct.

Explanations for affiliation choices were transcribed and coded. If a response referenced copying or doing the same thing, it was categorized as imitation-related (e.g. "Keeps doing the same thing", "Because they were saying the same thing"). All other responses were coded as unrelated (e.g. "I don't know", "Because I like him.")

Results and discussion

Adults. We first consider responses from adult participants.

Analyses of movement sequence. Analyzing the movement sequence separately allows us to better compare this experiment and the previous ones which only involved bodily movements. Responses to side-character action trials and propositional scaffolding trials were both 96.25% accurate, which was above chance (b = 3.24, SD = .59, z = 5.51, p = 3.5e-08). On the test questions, adults responded with 83% accuracy to the prediction question, which was significantly higher than chance (33/40 correct, p = 4.2e-05). Their responses to the affiliation question were 88% accurate, and also significantly better than chance (35/40 correct, p = 1.4e-06). This means that adults remembered the actions the characters performed and were successful at answering both inferences correctly.

Comparison between adult responses in this experiment and Experiment 2 found no differences across the two, meaning that adults succeeded at answering both inferences (affiliation and prediction) equally with or without the imitation being highlighted for them.

Analysis of all action sequences. Adults responded with 93.8% accuracy on side-character action trials, and 94.2% accurate on propositional scaffolding trials. Both sets of performances were significantly better than chance (side-character action trials: $b = 3.31$, $SD = .57$, $z = 5.78$, $p = 7.52e\text{-}09$; propositional scaffolding trials: $b = 3.41$, $SD = .64$, $z = 5.34$, $p = 9.54e\text{-}08$).

As each participant provided three responses for each test question across all sequences, answers were compared against chance by a mixed-effects logistic regression analysis that contained only the intercept and the subject as a random effect (glmer(Response ~ 1 + (1|Subject)). Participants responded to the prediction question with 89.2% accuracy, which was significantly higher than expected by chance ($b = 5.76$, $SD = 2.41$, $z = 2.39$, $p = .017$). Their responses to the affiliation question were 90% correct, which was again better than expected by chance ($b = 3.34$, $SD = 1.17$, $z = 2.86$, $p = .0042$). There was no main effect of action type or sequence order or any interaction in either affiliation or prediction performance.

Children. We next analyze responses from children for bodily movements and all action sequences respectively.

Analyses of bodily movements. Participants answered the side-character action trials 90% accurately, and the propositional action trials 89.5% correctly. Both were significantly better than expected by chance (side-character action trials: $b = 7.36$, $SD = 1.2$, $z = 6.02$, $p = 1.8e\text{-}09$; propositional scaffolding trials: $b = 7.4$, $SD = 1.2$, $z = 6.12$, $p = 9.4e\text{-}10$). A mixed-effects binary logistic regression model including experiment as a fixed effect and subject as a random effect

was run on pooled responses for imitator action trials from Experiment 2 and the analogous propositional scaffolding trials for the movements sequence from the current experiment. This found no difference between the experiments, thus demonstrating that the wording of the question (i.e. "Did Jenny move her arms like Alice…" vs. "Did Jenny move her arms…") did not affect performance. In other words, the fact that the question referenced the other two characters in the current experiment neither helped nor hindered children's rate of remembering the imitator's action correctly.

On the test questions, children responded with 72.8% accuracy on prediction, which was better than expected by chance (83/114 correct, p = 1.2e-06). Responses for affiliation were 75.2% accurate and significantly higher than chance (85/113 correct, p = 7.3e-08). According to a binary logistic regression including age in months, there was no effect of age on either prediction or affiliation responses.

Next, a binary logistic regression model compared responses from Experiment 2 and the movement sequence data in Experiment 3. There was a significant difference between the two experiments for prediction responses (b =.76, SE = .25, z = 3.2, p = .002), and also for affiliation responses (b = .79, SE =.25, z = 3.1, p = .002). That is to say, children's performance on both prediction and affiliation in the current experiment, which included propositional scaffolding, was better than that of their peers in Experiment 2, which did not.

Comparing children and adults. A binary regression model tested for a possible difference between children and adults. As measuring age in months would not be suitable to use for adults, age was treated as a categorical variable (four categories: three-, four-, five-year-olds, and adults). There was no significant difference between children and adults, or between any of the child age ranges, for either prediction or affiliation. In other words, children were just as

successful as adults in inferring affiliation and predicting who the imitator's future target would be.

Analyses of all action sequences. Responses for side-character action trials were high at 90.2% accuracy, and significantly better than expected by chance (b = 2.8, SE = .27, z = 10.34, p < 2e-16). Responses were also high for propositional scaffolding trials, which children answered with 90.3% accuracy, and were once again significantly better than chance (b = 2.5, SE = .21, z = 12.03, p < 2e-16).

Children responded to the prediction question correctly 69% of the time. Overall performance on prediction was significantly better than chance (b = .88, SE = .15, z = 5.94, p = 2.91e-9). A similar analysis was carried out on affiliation responses, which were 71% correct, and also significantly better than chance (b = 1.12, SE = .19, z = 5.9, p = 3.35e-9).

See Figures 2.4 and 2.5 for graphs of the results according to age and action type. Several regression models were fitted to children's responses for the prediction question to test for possible effects of age (a continuous variable in months), action type, and sequence order. The best model (according to a comparison between models using *anova()* in R) was an additive one including age and action type as fixed effects and subject as a random effect, which found a significant main effect of age (b = .03, SE = .01, z = 2.28, p = .02). There was no significant difference between action types, or an interaction between age and action type.

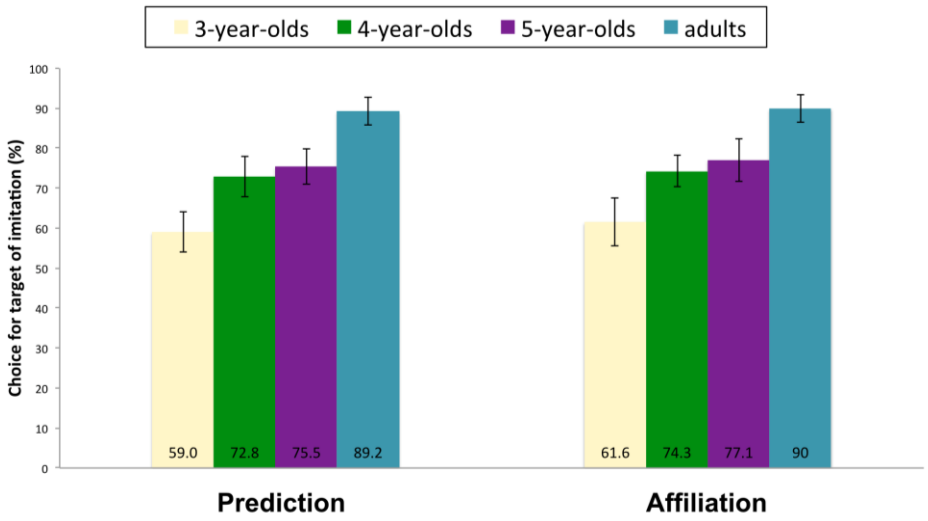


Figure 2.4. Performance for prediction and affiliation questions in Exp 3 according to age

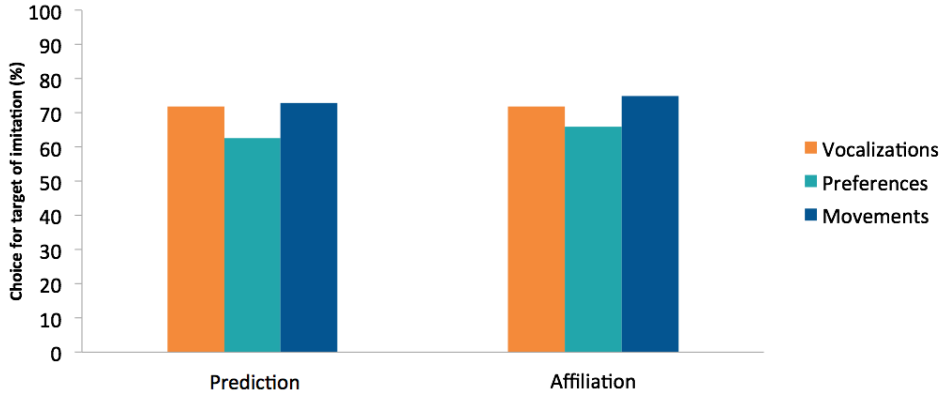


Figure 2.5. Performance for prediction and affiliation questions in Exp 3 according to action type

Comparing children and adults. An additive regression model with action type and age (four categories: three-, four-, five-year-olds, and adults) as fixed effects and subject as a random effect was conducted. This analysis found significant differences between the age groups, but no significant effect of action type. Pairwise comparisons of age groups (adjusted for multiple comparisons according to the Holm-Bonferroni method) showed that adults outperformed all groups of children (adults vs. three-year-olds: $p < .0001$; adults vs. four-year-olds, $p = .018$; adults vs. five-year-olds, $p = .033$). There were no significant differences among the children. A

similar analysis was conducted on affiliation responses. Pairwise comparisons adjusted for multiple comparisons according to the Holm-Bonferroni method found that adults outperformed three-year-olds ($p = .0001$), and four-year-olds ($p = .03$), and responded marginally better than five-year-olds ($p = .059$). Once again, there were no significant differences among the child age groups. In short, children were comparable to each other overall, but outperformed by adults when all three action types were taken into account.

Analyses of explanations. Explanation responses were coded into the two categories of imitation-related and imitation-unrelated. Asking for explanations was added to the procedure midway through data collection. Therefore, only a subset of the children had explanation responses. Explanations were available for 20 three-year-olds, 31 four-year-olds, 35 five-year-olds, and all 40 adult participants. Of those who answered the affiliation question correctly, 17% of three-year-olds, 54% of four-year-olds, 72% of five-year-olds, and 100% of adults gave imitation-related responses. This pattern shows an increasing rate of appealing to imitation with age. Another way to consider the explanations is to examine what proportion of subjects who talked about imitation also responded to the affiliation question correctly. This method shows similar rates across age, since of those participants who referenced imitation, 100% of three-year-olds, 85% of four-year-olds, 95% of five-year-olds, and 100% of adults answered the affiliation question correctly. That is to say, participants who demonstrated explicit awareness of the imitation in the scenes almost always inferred affiliation too. This provides support for the notion that attention to imitation as the basis of the inference is key to success.

This experiment yielded several findings. First, highlighting imitation through propositional scaffolding provided in the form of questions (e.g. “Did Jenny move her arms like Alice or her leg like Zoe?”) led to children at all ages succeeding in responding correctly to the

prediction and affiliation questions. Second, children performed comparably across three types of actions, namely, bodily movements, vocalizations, and preferences for objects. Third, there was an age effect in children's performance for all three sequences overall. However, the age effect did not show up as a significant difference between the age ranges when compared to each other, nor was there an age effect in responses for movements. Thus, while some change is evidently happening within the preschool period, it is not a drastic one. The next experiment highlights imitation differently by priming the concept and asks if children will once again succeed at the inferences.

Experiment 4

This experiment tested whether priming the concept of imitation would lead children to succeed in inferring affiliation and continued imitation of the same target. The hypothesis being tested is that children succeeded in the previous experiment because the propositional scaffolding led them to attend to the imitation happening in the scenes, and answer the test questions while bearing imitation in mind. In an experimental test of this hypothesis, this experiment primes children with the idea of imitation at the start, once again cuing children to attend to imitation. Importantly, this is achieved using a different set of characters and a different action type than those used at test. Further, no propositional scaffolding is given for the action on which the test questions are asked. If highlighting imitation in fact led children to success in the previous experiment, they should also succeed in this case.

This experiment tests for a possible deflationary account of the success in Experiment 3. The propositional scaffolding used in Experiment 3 led to success. However, it is not clear how. One possibility is that asking the scaffolded version of the question helped children choose the correct answer on the test questions simply because the correct answer to one was the correct

answer to the other. Children were asked the scaffolded question on the first two trials that showed imitation, and they were praised for correct answers. If they answered incorrectly, the scene was replayed and they were asked the question once again. In other words, it was made clear to them that the target (or the target's action) was the correct answer to the propositional scaffolding questions. Children then saw three further instances of imitation before answering the test questions. Given this procedure, one strategy for answering the test question correctly could have been to simply choose the same character that had been reinforced in the propositional scaffolding trials. Some indirect evidence against this comes from the explanations offered by children, because almost all children who referred to imitation in their explanations answered the affiliation question correctly, suggesting that those who truly showed explicit knowledge of the imitation were also successful at inferring affiliation. Nonetheless, the current experiment is in a better position to rule out this deflationary account.

Method

Participants. As none of the previous experiments found a difference in performance between four- and five-year-old children, this study includes a mixed sample of both ages. A sample of 40 four- and five-year-old children was tested (M = 60.0 months, 20 female). Half the children were four years old (48-59 months), and half were five years old (60-72 months). An additional two participants were excluded due to not completing the task. According to preset exclusion criteria, yet another participant's responses were not included in the analyses due to answering the critical question at the end of the priming phase incorrectly.

Materials. The stimuli involving the three female characters used in Experiments 1-3 were used. Part of the stimuli used for the vocalization sequence in Experiment 3 was also used.

Procedure. The procedure consisted of two sequences: priming with vocalizations, and test with movements. The first sequence was a shortened version of the vocalizations sequence from Experiment 4, and used to prime the concept of imitation. This included the following:

Priming sequence.

1. *Introduction.* The central character for the vocalizations characters was introduced and performed a number of vocalizations to demonstrate that she was able to utter them all.
2. *Side-character action trials.* Two trials were shown involving the side characters, and children's memory for their utterances was probed.
3. *Propositional scaffolding trials.* Two trials with all three characters present were shown, and the central character imitated the target. Children were asked about the imitator's action, e.g. "Did Emily say vow-vow like Hannah or zoh-zoh like Katie?"
4. *Imitation question.* Children were asked who the central character had imitated, "Who was Emily copying?"

Test sequence. Following priming, participants completed the second sequence involving movements, and answered test questions. The procedure for this sequence up to the test questions was identical to that used in Experiments 2.

1. *Introduction.* The central character was introduced and performance of all actions she would get the opportunity to imitate later on.
2. *Side-character action trials.* Two trials of side-character actions, which children were asked about.

3. *Imitator action trials.* Two trials of the imitator copying the target, which children were asked about (e.g. “Did Jenny move her arms or her leg?”)
4. *Imitation.* Three further trials of imitator copying the target. The wording for the imitator action trials was not scaffolded, e.g. “Did Jenny move her head or her leg?”
5. *Test.* The experimenter asked the affiliation question, “Who does Jenny like best?” and the prediction question, asking them to predict which character’s action the copier would repeat on an incomplete trial, e.g. “Is Jenny going to bend sideways like Alice or move her arms like Zoe?” The order of the questions was counterbalanced across participants.

Design. The side of the targets of imitation was counterbalanced across children. Importantly, the target of the priming sequence was always on the opposite side of the target of imitation for the test sequence.

Results

Participants responded to the side-character action trials with 98.8% accuracy for the vocalization sequence, and with 96.3% accuracy for the movements sequence. Both sets of responses were significantly better than chance performance according to logistic binary regression models including the intercept and subject as a random effect ($p < .0001$ for both). Accuracy was at 81.3% for the propositional scaffolding trials in the vocalization sequence, and at 96.3% for the imitator action trials in the movements sequence. Responses were once again significantly better than chance ($p < .0001$ for both). Thus, children were once again successful at answering the questions that probed their memory of the actions performed by each of the side characters. Further, they also easily remembered the actions performed by the imitator,

regardless of whether the question referenced the side characters in the case of the propositional scaffolding of the vocalization, or whether it did not reference it as in the imitator action trials of the movements sequence.

As mentioned above, the criterion for being included in the final dataset was to answer the imitation question in the priming sequence correctly. Only one child (out of 41 tested) answered incorrectly, and her data were excluded from further analyses. With regard to the test questions, 77.5% of children responded correctly to the affiliation question, which was significantly higher than chance (31/40 correct, $p = .0007$). In response to the prediction question, 70% of responses were correct, which was also significantly better than chance, (28/40 correct, $p = .017$). Binary logistic regression models found no significant effect of age (in months) or question order and no interaction for either prediction or affiliation responses.

Comparing the responses in the current experiment with responses from children of the same ages (i.e. four- and five-year-olds) in Experiment 3 found no differences in prediction question ($p = .61$) or affiliation question ($p = .93$) responses. These analyses provide evidence that propositional scaffolding and imitation priming led children to comparable rates of success in inferring affiliation and prediction. Similarly comparing responses here to those from Experiment 2 found a significant difference between answers to the affiliation question ($b = .87$, $SD = .41$, $z = 2.11$, $p = .035$), but no difference in prediction responses across the two experiments. The difference in performance on affiliation could reflect the efficacy of priming in helping children succeed. The lack of difference on prediction, however, is ambiguous. One possible reason that performance in the current experiment is not significantly higher than that for Experiment 2 could be the vast difference in sample size (40 vs. 160). In other words,

performance may be indeed better in the current task, but the smaller sample size may not provide enough power to capture the difference.

To summarize, children were primed with the notion of imitation ahead of watching the movement actions that were used in the previous experiments, and answered prediction and affiliation questions. They succeeded in answering both questions. The target of imitation in the priming sequence was consistently on the opposite side from the target of imitation in the test sequence, so children's success in answering the questions could not have come from simply persisting in choosing the character who was on the same side as the response for the priming imitation question ("Who was Emily copying?"), which they knew the answer to.

General Discussion

The current set of studies yielded several findings. First, three-year-old children recognized the target of an individual's copying behavior from third-party observations of imitation (Exp 1). Second, despite this ability to recognize imitative behavior, three-, four-, and five-year-olds' failed to spontaneously make inferences about an imitator's proclivity to continue copying the target in the future (Exp 2). Children did infer the imitator's affiliation towards the target, but the effect was small in magnitude. We next found that preschoolers answered both questions more successfully when the imitation was highlighted through propositional scaffolding, drawing attention to the fact that the imitator was copying the target (Exp 3). Additionally, children's performance was comparable across different types of actions, bodily movements as well as imitation of vocalizations and object preferences. In the final experiment, children were primed with the concept of imitation ahead of viewing the same scenes as Experiment 2, and this time answered the test questions correctly. Unlike children's differing

levels of success across the experiments, adults' performance was consistently high on both inferences.

The current set of results shed light on the inferential capacities of preschoolers with regard to observed imitation. The novel finding is that although children recognized the imitation explicitly, their ability to spontaneously infer on the basis of it was limited. In other words, children did not appear to explicitly encode the imitation unless it was scaffolding for them, or unless they were primed with the notion of imitation. Even the youngest group of preschoolers at three years of age possessed the ability to make the affiliation inference that infants made. However, the results also show that this ability only manifests under certain conditions, and it could have been left undiscovered but for the last two experiments.

Although the current experiments do not provide conclusive answers, they provide a foundation on which to speculate about the possible reason for the previously reported failure of four-year-olds (Over & Carpenter, 2014), and children's relative lack of success in Experiment 2, which didn't highlight imitation. The reason could be that children were sidetracked in making clear inferences because the hypothesis space they considered was richer and more varied than that considered by infants. In other words, preschoolers may have been focusing on many aspects of the scene, instead of just considering the most relevant one, i.e. imitation. This could have led them to respond unsystematically as a group. Infants, on the other hand, demonstrated success in looking time experiments because they were much more limited in the features of these events they could encode. That is to say, preschoolers' failure is not a signal of their inability to make the inference, but rather of their lack of ability in prioritizing among several possible bases for it. Importantly, the bases they considered include the correct one, namely imitation. This is why highlighting imitation simply acts as a signal, which guides their inference by limiting the

hypothesis space they take into account. This speculative account is difficult but important to test. Furthermore, it could hold promise for resolving the other cases of discontinuity, such as theory of mind or solidity constraints.

A starting point for testing the possibility laid out above is to further test why highlighting imitation allowed children to make the inferences, and rule out deflationary accounts of their success. The interesting possibility is that highlighting the imitation allows children to realize that making inferences about affiliation and prediction are licensed and meaningful in the context of the imitative roles in the scene, i.e. who is copying whom. However, another possible mechanism is that the act of emphasizing the imitation changes the pragmatics of the situation, such that children think that the experimenter's intention is for them to answer the questions based on the imitation. Yet another possibility is that what made the difference was not stressing imitation per se, but rather emphasizing the social nature of the scenes and showcasing the dependence of one individual's actions on another's. If so, children may succeed at inferences based on imitation even if they are just put in a more social frame of mind, for instance through affiliation or ostracism priming, which have been shown to increase prosocial actions (Over & Carpenter, 2009a) and invoke increased imitation (Over & Carpenter, 2009b) respectively. Resolution of the mechanism is left to future work.

An interesting aspect of the results is that inferences about affiliation were not more difficult to make than predicting an imitator's future action. It is unlikely that children, or even adults for that matter, would expect an individual to continue imitating a target indefinitely. It is, however, reasonable to expect an individual to copy the target again when the question implies that the individual's choice is between the action of the past target and the past non-target. In order to infer that the copier will continue to imitate the same target as before, children need to

notice that imitation is taking place and to encode which of two candidates is the target of imitation. Experiment 1 showed that even three-year-olds were capable of this. Thus, just by appealing to simple consistency, the answer is within reach. In order to infer affiliation, however, an observer needs to invoke a richer social framework which recognizes the link between imitation and liking, which would seem to require extra representational machinery than what is needed for the prediction question. The data were not consistent with this analysis. Experiment 2 revealed that children succeeded in answering affiliation, but not prediction. Experiments 3 and 4 found that children succeeded on both questions. Thus, it seems that children's ability to make the two inferences were similar overall. In sum, there was no added cost to inferring affiliation between an imitator and a target.

To summarize, this set of experiments provides evidence showing that preschoolers as young as three years of age can infer affiliation and continued imitation of the same target from observed imitation, so long as the imitation in the scenes is highlighted for them. This in turn suggests that the reasoning abilities based on imitation that infants demonstrated (Powell & Spelke, 2017) do not simply vanish in early childhood (Over & Carpenter, 2014). Rather, they manifest when children are guided towards noticing what really matters.

Paper 3:

Inferences about yielding in hierarchical relationships:

A cross-cultural comparison

Narges Afshordi, Anni Kajanus, and Felix Warneken

Introduction

Dominance hierarchies organize the social worlds of both humans (Fiske, 1992) and other species (e.g. Drews, 1993). Furthermore, engagement in dominance behavior and recognition of it appears early in ontogeny. Infants between 11 and 16 months observed in a daycare setting showed evidence of dominance relationships, and could be identified as being dominant or subordinate in relation to other individuals based on the outcomes of their agonistic disputes (Russon & Waite, 1990). Interestingly, infants in the group preferred to imitate individuals who were ranked highly overall. In third-party tasks, 10-month-old infants expect a larger individual to win right of way over a smaller one, showing they have access to size as a cue to dominance (Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011), and slightly younger infants expect members of larger groups to win over those of smaller ones (Pun, Birch, & Baron, 2016). Further, representations of dominance roles in pairs of individuals are stable during infancy, meaning that infants expect an agent who has won a zero-sum goal conflict to win again in other similar situations by a year of age, and in completely new situations by 15 months (Mascaro & Csibra, 2012). Additionally, infants succeed at inferring transitivity of dominance relations around 11 months (Gazes, Hampton, & Lourenco, 2015; also see Mascaro & Csibra, 2014), and at 17-month-old expect dominant individuals to receive more resources (Enright, Gweon, & Sommerville, 2017). This collection of findings demonstrates that even preverbal infants represent dominance and harbor expectations about how dominance plays out in different settings.

Like infants, preschool-aged children also maintain and recognize dominance relationships. An observational study of four- and five-year-olds in a preschool reported dominance relationships to be relatively stable over time (LaFreniere & Charlesworth, 1983). It

also found that aggressive physical exchanges decreased over the course of the year, perhaps reflecting the entrenchment of the dominance hierarchy and the lessened need to assert status at each turn. Similar to findings from infancy, dominant children were the objects of attention, affiliation, and popularity amongst their peers (LaFreniere & Charlesworth, 1983), a pattern corroborated by other studies (for a review, see Hawley, 1999).

Prestige as a path to status

Despite the early facility with dominance relationships, dominance is not the only hierarchical relationship at work in human societies. Henrich and Gil-White (2001) proposed prestige as a uniquely human process that provides the individual with similar benefits and access to resources as dominance, but operates on an entirely different premise. Instead of relying on force or threat of force, prestige is status achieved through "merit in the eyes of others" (Henrich & Gil-White, 2001, pg. 170). Prestigious individuals typically possess highly-prized expertise, which others strive to witness in action and learn from. As a result, prestigious parties exert considerable influence over "clients" but are careful to avoid enforcing their views on others. In turn, lower-status individuals treat prestigious persons with respect, seek proximity to them, express admiration for and pay tribute to them. The key similarity between dominance and prestige is that they both bring about deference from others, but critically through different mechanisms. Deference towards a dominant person is driven by fear and an intention to avoid a showdown, which would most likely result in the lower-status person losing. Deference towards a prestigious person, in contrast, is given freely as lower-status individuals are motivated to do so in order to gain more proximity to the prestigious person, which in turn translates to learning opportunities.

The notion of different paths to social power has been tackled through a developmental lens as well. Prior to Henrich and Gil-White's approach to prestige (2001), Hawley (1999) put forth a similar-minded but developmentally-couched theory of strategies that lead to social dominance. This framework differentiated between coercive and prosocial strategies, which can be thought of as analogous to dominant and prestigious processes respectively, and reviewed the developmental literature to suggest that infants' and younger preschoolers' repertoire of behaviors only includes coercive strategies. These coercive strategies buy attention, affiliation, and influence among other subordinate group members. As children develop, however, a bifurcation happens such that prosocial strategies also come into play. Hawley suggested that prosocial and coercive tendencies more or less coexist between the ages of four and seven. Among older elementary schoolers, however, coercive strategies cease to be effective in winning attention, affiliation, and influence, giving way to prosocial strategies instead. There is evidence for children's use of both strategies in the intermediate years. In a quasi-experimental study with four- to seven-year-olds, high-status children were paired with low-status partners and left alone with a novel toy for a short play session (Hawley, 2002). High-status children used both prosocial strategies (e.g. making suggestions, demonstrations, offering objects) and coercive strategies (e.g. directly taking objects, physical aggression, insults) to gain control of the desired toy, and prosocial strategies were used twice as often as coercive ones. Interestingly, the so-called prosocial strategies were highly self-serving and perhaps even manipulative (e.g. taking control of the toy by suggesting the subordinate partner perform the less desirable role). Nonetheless, they still project a sense of soft, rather than harsh, power. In short, there is much similarity between the concepts of prestige and prosocial strategies, although prestige in its

ultimate sense seems to be a more sophisticated process than what younger social players exhibit.

Besides engaging in hierarchical relationships, preschoolers make a number of inferences about them as third-party observers. Three-year-olds expect larger groups to overcome smaller ones (Lourenco, Bonny, & Schwartz, 2016). By age five or younger, preschoolers recognize the "boss" or the "person in charge" based on cues of physical supremacy, decision-making ability, age asymmetry, resource asymmetry (Charafeddine et al., 2015), resource control, goal achievement, being asked permission (Gulgoz & Gelman, 2016), being imitated (Over & Carpenter, 2014), and even non-verbal gestures such as posture and head angle (Brey & Shutts, 2014). Additionally, children make a number of inferences about high-status characters, and expect them to win competitive games (Charafeddine et al., 2015), endorse their testimony and labels for novel objects (Bernard et al., 2016), and even allocate more resources to them (Charaffeddine et al., 2016). Finally, four-year-olds took advantage of bystander attention, a cue theoretically linked to prestige (Henrich & Gil-White, 2001), to imitate the actions of an actor they had seen being observed by others (Chudek, Heller, Birch, & Henrich, 2012). They were not, however, systematic in responding to questions about who they preferred to play with or who was more popular. Although this literature has for the most part framed the findings in terms of dominance understanding (or as prestige understanding in the case of the final study mentioned above), most of the cues provided to children are in fact ambiguous and could be indicative of either dominance or prestige (e.g. age asymmetry, resource asymmetry, asking permission). Thus, while it may seem on the surface that preschoolers have a solid grasp of dominance as a concept, and perhaps a shakier one of prestige, there is in fact no conclusive evidence for or against them distinguishing one from the other. Consequently, the studies

reviewed above provide evidence that children are generally sensitive to hierarchical relationships as observers. Hence, the question arises whether children can in fact distinguish between dominant and prestigious displays of status as third-party observers. Experiment 1 tests this question specifically by asking whether children understand that prestigious individuals are liked and that dominant individuals are feared. Moreover, we test two age groups of children (5-7 years, and 9-12 years) to examine whether this differentiation happens before or after Hawley's (2002) proposed bifurcation of strategies at around eight years of age.

Yielding in Chinese culture

The second question at the heart of this paper is whether children infer that dominant and prestigious characters will have a higher chance at accessing a resource in comparison to a low-status character. Chinese culture presents an intriguing case for studying third-party expectations about yielding and deferring resources. From a young age, children are encouraged to show modesty, give precedence to others, and demonstrate a degree of selflessness that is not typically expected of young children in Euro-American cultures. For instance, toddlers are taught and expected to give up prized toys to children younger than themselves (Kajanus, in prep.). The fable of Kong Rong--a first-century scholar and descendent of Confucius--is a telling tale of this value in Chinese culture (Xu, 2017). This famous story recounts how at four years of age, his father tasked him with dividing up a basket of pears between himself and his brothers. Kong Rong gave away all but the smallest pear. When asked by his father why he did so, he said that he should be respectful of his elder brothers. When pressed on why he gave a bigger pear to his younger brother than the one he kept for himself, he replied that as the older sibling, he should yield and take care of his younger brother.

Beyond this moral tale, there is direct ethnographic evidence that children are expected to defer to those who are lower (or equal) status than themselves, which in many cases translates to older children yielding to younger ones. For instance, Chinese adults often break up a conflict between two children by addressing the older one with criticisms such as "Why can't you just yield to her?" (Kajanus, in prep). Further, prestigious children are especially expected to yield to others. Kajanus asked 9- and 10-year-old children in two schools in China to give the names of others they liked most to play with. The children predominantly mentioned the same four or five individuals, who were socially engaged with others, liked by teachers, and strong academically. Importantly, she observed that these universally popular children were rarely the ones to instigate fights or conflict. Additionally, parents highly valued yielding in conflicts or avoiding them completely, and the parents of the most prestigious boys in both schools explicitly talked about this. These ethnographic observations strongly suggest that the relationship between status and yielding in Chinese culture is different from Euro-American contexts.

Further, there is a current push and pull between between this traditional Chinese value widely taught to children on the one hand, and the concern that strongly pressuring children to behave in line with this ideal may promote dishonesty and hypocrisy. Xu (2017) recounts an interesting case that received much attention on social media involving an angry father who posted a picture of his seven-year-old son's test at school in which he was docked points for responding that he would not turn down the bigger pear if he were in Kong Rong's position. The child clearly understood the moral of the story, as he had answered a separate question correctly by saying that the story demonstrated the values of deference and modesty. What outraged the father, and in which he found a large number of sympathizers online, was that the child was reprimanded for expressing his authentic opinion. In her ethnography of preschool-aged children

in China, Xu (2017) talks about how many adults similarly feel that asking children to show modesty and give up possessions--which is difficult for them to do--puts them in positions that compromise the genuineness of the child. Adults acknowledge that showing modesty is only worthy if done out of sincerity, but they nonetheless sometimes reward children for demonstrating the behavior in order to promote the value. Xu (2017) observed such a situation in a preschool where a child who had won exclusive access to a popular play area was asked by the teacher to give it up to another child who was eager to play there. Following the child's acceptance, the teacher allowed both children to play in the desired spot to reward his behavior, but worried that such instances could turn the modesty into scheming. A final piece complicating this picture arises from modern family dynamics in China over the past two generations. While the primary context for practicing modesty and deference for children would traditionally have been sibling relationships, the fact that the majority of children do not have siblings as a result of China's longstanding one-child policy diminishes the situations in which young children can express it. To sum up, showing modesty and yielding to others, even those lower in status to oneself, is highly valued and encouraged, but within a complex social context which also acknowledges the difficulty that children may have with it.

All of this leads to the hypothesis that the inference that a high-status party, especially a prestigious one, would be the winner of a desired resource is far from straightforward for Chinese individuals, and possibly different from inferences made by individuals living in Euro-American contexts. Alternatively, it is possible that the norm of yielding by prestigious individuals does not manifest in third-party expectations about the outcomes of conflicts. The second experiment tests this hypothesis in adults, and the final experiment tests it with in age groups of children (5-7 years, and 9-12 years).

Experiment 1

The aims of the first experiment were to examine 1) whether children recognize that dominance and prestige both result in higher status, 2) whether they can differentiate between dominant and prestigious roles, and 3) whether children in the UK and China would respond similarly. Three characters were involved in the experiment, a relatively low-status central character, a high-status dominant individual, and a high-status prestigious individual. In order to establish one high-status individual as dominant and the other as prestigious, participants viewed two animated scenes, one involving only the central and dominant characters, the other involving only the central and prestigious characters. Across the two scenes, a number of cues were manipulated in order to showcase the difference between the dominance and prestige roles. First, the interactions were started by different parties. In the dominance scene, the dominant character began by asking the central character what she was doing. In contrast, the central character started the prestige scene by explaining what she was doing to the prestigious character. Second, the dominant character enforced her opinion on the central character, while the prestigious character shared her opinion only after she was consulted for advice. Similarly, while the central character tried to resist the dominant character's choice, she actively sought to see the prestigious character's selection. Thus, while the central character ended up making the same choice as both the dominant and prestigious characters, how this came about was different.

In order to test recognition of the status difference, participants were asked who was older between the low-status character and each of the two high-status characters separately. Previous studies have asked children to identify the person who "is in charge" (Brey & Shutts, 2014), or the "boss" (in German, Over & Carpenter, 2014; in French, Charafeddine et al. 2015; 2016). While it is possible for a prestigious character to be in charge or the boss, the source of

their status lies in influence and popularity, rather than sheer power. Therefore, the question about status was worded more neutrally and enquired who was older in each pair, as an older person can be either dominant and a bully or prestigious and admired. Crucially, using this wording allowed us to ask the question in the same way for both the dominant and prestigious individuals. Besides this, the experiment tested whether participants made judgments about the two high-status characters that aligned with the source of their status.Three additional questions were asked in order to examine whether participants distinguished between the dominant and prestigious characters: who the central character would approach when given the choice, who she liked more, and who she feared more. The correct answers to these questions are to infer that the the central character would approach the prestigious individual and would like that one more, while being more fearful of the dominant character.

Two groups of different ages were recruited in China and the UK, a younger group (5-7 years), and an older group (9-12 years). Given preschoolers' ability to infer status from a host of cues as mentioned above, and the fact that children in our study were five years and older, it is reasonable to expect that they would infer the status differences here as well. More specifically, the prestige scene included the subordinate copying the choice of the prestigious character, and even five-year-olds associate being imitated with higher status (Over & Carpenter, 2014). In the dominance scene, the dominant character enforced her decision on the subordinate, another cue that children as young as three years associate with higher status (Charafeddine et al., 2015). With regard to distinguishing between dominance and prestige, Hawley's framework (2002) suggests that this may be beyond the reach of children younger than eight years as they are still relatively accepting of coercive (or dominance-like) strategies. This hypothesis is backed by evidence about children's third-party preferences between dominant and subordinate individuals.

While two-year-olds liked the dominant party (Thomas, Abramyan, Lukowski, Thomsen, & Sarnecka, 2016), three- to five-year-olds did not systematically prefer either party, although there was a non-significant decrease in preference for the dominant individual. By eight years, however, children showed a significant preference for a subordinate character (Charafeddine et al., 2016), suggesting that dominant behaviors are not favored or easily tolerated at this age. In line with this, our younger group may be unable to infer that the dominant individual would be avoided, disliked, and feared. On the other hand, high-status four- to seven-year-old children engage in more prosocial than coercive behaviors (Hawley, 2002). Given this proclivity to prestige-like behavior, it is alternatively possible that younger children would distinguish between prestige and dominance as well. Regardless of how younger children respond, older children are expected to succeed at this task under both accounts.

The question of whether children in the UK and China respond similarly is interesting for two reasons. First, all of the developmental findings about status inferences come from children in Euro-American cultures, and it is unknown whether children outside of those cultures, Chinese children in this case, interpret the cues the same way. Second, in light of the cultural value of prestigious individuals yielding to others in China, the question arises as to whether other aspects of prestige understanding, including inferences based on cues, might also be affected. The comparison between children in the two countries will help answer this question.

Method

Participants. In the UK, forty children in the younger group (M = 5.7 years, range = 5.0-6.8 years, 19 female) and 40 children in the older group (M = 10.4 years, range = 9.3-11.1 years, 21 female) took part in the study. All children were recruited from and participated at a public school in central London. According to the national rating of schools in the UK, the school has a

Good rating (on a four-point scale of Outstanding, Good, Satisfactory, and Inadequate). The children attending the school came from a mix of working- and middle-class backgrounds, and the school is highly diverse in terms of race and ethnicity. In China, forty younger children (M = 5.8 years, range = 5.2-6.3 years, 19 female) and forty older children (M = 10.9 years, range = 10.15-12.95 years, 22 female) participated. All children were recruited from and participated at their school or preschool in Nanjing, China. Although no formal rating is available for the school and preschool, they are both considered to be at average level, similar to the school in London. Children predominantly came from working-class backgrounds.

Materials. The stimuli were animations involving three cartoon characters: an orange diamond-shaped character in the center, a blue rectangle-shaped character on the left, and a teal oval-shaped character on the right. All characters had similar facial features (eyes and mouth), as well as arms and legs. Each character's eyes moved whenever necessary, and their mouths moved as they spoke. The character in the center (i.e. orange diamond-shaped) was always in the role of the low-status individual, while each of the other two figures portrayed the dominant or prestigious characters. Whether the dominant role was assigned to the rectangle character on the right or the oval character on the left was counterbalanced across participants.

Each participant viewed two main scenes. One scene showed an interaction between the central and the dominant characters (dominance scene), while the other involved the central and prestigious characters (prestige scene). In both scenes, the characters were shown sitting behind desks and engaging in an art project, either coloring a picture of a house, or sticking a picture of an animal onto a frame (see Figure 3.1 for a sample). Whether the coloring situation was used to show dominance or prestige was counterbalanced across participants.

The gender of all three characters was matched to the participant in order to avoid any gender-specific expectations about status. Thus, animations were customized to gender and language (English and Mandarin). For each set, the conversations were recorded by three different adults, one assigned to each role (i.e. central character, prestigious, dominant). The recordings were then passed through a high-pitch filter using a sound editing software in order to make them sound child-like and appropriate for use with cartoon characters. The voice actors for the English version were American-accented native speakers of English, and the actors for the Chinese version were native speakers of Mandarin. The conversation script was translated into Mandarin by a native speaker, and was then discussed and adjusted for accuracy and comparability with the English version by the translator and another researcher highly fluent in Mandarin.

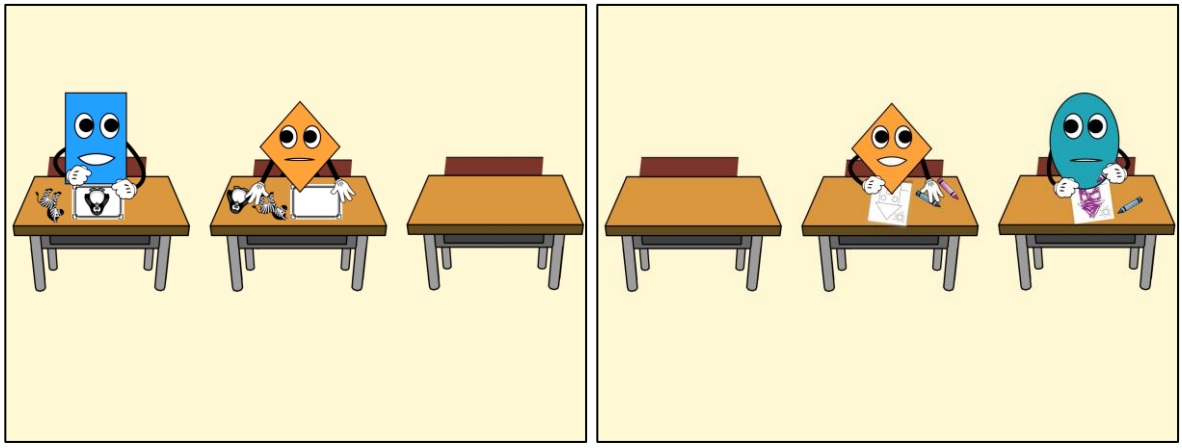


Figure 3.1. Sample of sticker situation with rectangle character (left) and coloring situation with oval character (right).

All animations were made using Keynote, and saved as video recordings. The experiment was presented using Keynote on a laptop computer. The laptop keyboard was covered in order to avoid distraction, and slides were advanced using a mouse.

Procedure. The central character was shown alone on the screen and was introduced to participants, e.g. “This is Dimo” for the English version. The character was given a novel name as opposed to a proper one to avoid the possibility of any participants associating the character with someone they knew with the same name. In Mandarin, the name given to the central character translates approximately to “Little Orange”, once again allowing the experimenter to refer to the character without using a proper name. (It is important to note that using “little” (*xiao* in Mandarin) in the central character’s name did not provide children with a cue to status, as this is a common way of generating nicknames in Mandarin. Further, many children referred to the other characters as “Little Blue” or “Little Green”, demonstrating that they did not view “little” as indicative of status in this context.) After the introduction, the dominance and prestige scenes were played. Participants saw each scene played twice in a row.

Dominance scene. The interaction began with the dominant character asking Dimo what she was doing. Dimo then explained that she had two choices for her art project (choosing blue or purple in the coloring situation, choosing a penguin or a zebra in the sticker situation), at which point the dominant character stated the choice she had made and told Dimo that she should also choose the same option. Dimo resisted the choice by saying that she preferred the other option. However, she subsequently made the same choice at the dominant character’s insistence, and completed the art project. The dominant character’s finished art project was visible from the beginning of the scene. The dominant character’s tone of voice was relatively aggressive throughout, and Dimo sounded relatively sad and defeated when stating that she would make the same choice as the dominant character at the end.

Prestige scene. This interaction started by Dimo informing the other character that she had two choices for her art project, and then enquired the prestigious character’s opinion on

which one to choose. The prestigious character responded by informing Dimo of the choice she had made herself. Dimo then asked to see the prestigious character's project, at which point the prestigious character turned her project around to face Dimo in order for her to see it. After being shown the project, Dimo declared that she would make the same choice as the prestigious character, and completed her project accordingly. The prestigious character's tone of voice was confident and positive throughout.

Detailed transcripts of the conversations for both scenes can be found in Table 3.1 (coloring situation) and Table 3.2 (sticker situation). The order of the dominance and prestige scenes, whether the dominance scene involved coloring or sticking a picture (and vice-versa for the prestige scene), and the identity of the dominant character (left-hand or right-hand character) were counterbalanced across participants.

Table 3.1. Conversation script for coloring situation in English

Dominance scene	**Prestige scene**
Dom: What are you doing?	-
Dimo: I'm going to color my picture. I can color it in blue, or I can color it in purple.	
	Dimo: What do you think I should do?
Dom: I like purple. You should color it in purple!	*Pres:* I like purple. I colored my house in purple.
Dimo: But I want to color it in blue.	*Dimo:* Can I see your picture?
Dom: No! Color it in purple!	*Pres:* Yes! Here's my picture!
Dimo: Okay. I'll color it in purple.	*Dimo:* I'll color mine in purple too.

Table 3.2. Conversation script for sticker situation in English

Dominance scene	Prestige scene
Dom: What are you doing?	-
Dimo: I'm going to pick an animal to put on my picture. I can pick the zebra or I can pick the penguin.	
	Dimo: What do you think I should do?
Dom: I like the penguin. You should pick the penguin!	*Pres:* I like the penguin. I put the penguin on my picture.
Dimo: But I want to pick the zebra.	*Dimo:* Can I see your picture?
Dom: No! Pick the penguin!	*Pres:* Yes! Here's my picture!
Dimo: Okay. I'll pick the penguin.	*Dimo:* Then I'll pick the penguin too.

Test questions. After watching both the dominance and prestige scenes, five test questions were asked. Figure 3.2 shows the pictures that were shown as each question was asked.

- *Older (Dimo-Dom):* Showing a picture of Dimo and the dominant character at the beginning of the dominance scene, the experimenter asked, "One of these is older than the other. Which one is older?"
- *Older (Dimo-Pres)*: Showing a picture of Dimo and the prestigious character at the beginning of the prestige scene, the experimenter asked, "One of these is older than the other. Which one is older?"
- *Approach*: The experimenter asked "Dimo wants to sit down. Who is Dimo going to sit next to?" This question was asked while showing Dimo standing in the middle of the screen. The other two characters (i.e. dominant and prestigious) were sitting on two separate benches on either side of the screen looking at Dimo, and both benches had room for another person to sit on.

- *Affiliation*: Showing a picture of all three characters standing, the experimenter asked “Who does Dimo like more?”
- *Fear*: While still showing the same picture as the previous question, the experimenter asked, “Who is Dimo more scared of?"

The order of the two “Which one is older?” questions for each participant was the same as the order of the dominance and prestige scenes. In other words, if a participant watched the dominance scene first, they were also asked who was older between Dimo and the dominant character first. The remaining three questions were always asked in a fixed order. The order of the two older questions on the one hand, and the three remaining questions (approach, affiliation, fear) on the other, was counterbalanced across participants.

Coding. The dependent variable for the “Which one is older?” questions was the percentage of participants choosing the high-status character (dominant or prestigious depending on the pair of characters it was asked about). The dependent measure for the remaining three questions (approach, affiliation, fear) was the percentage of choices of the prestigious character. Although the correct response to the fear question is to choose the dominant character, responses were coded in this way so that they would be comparable to the approach and affiliation question responses in the analyses.

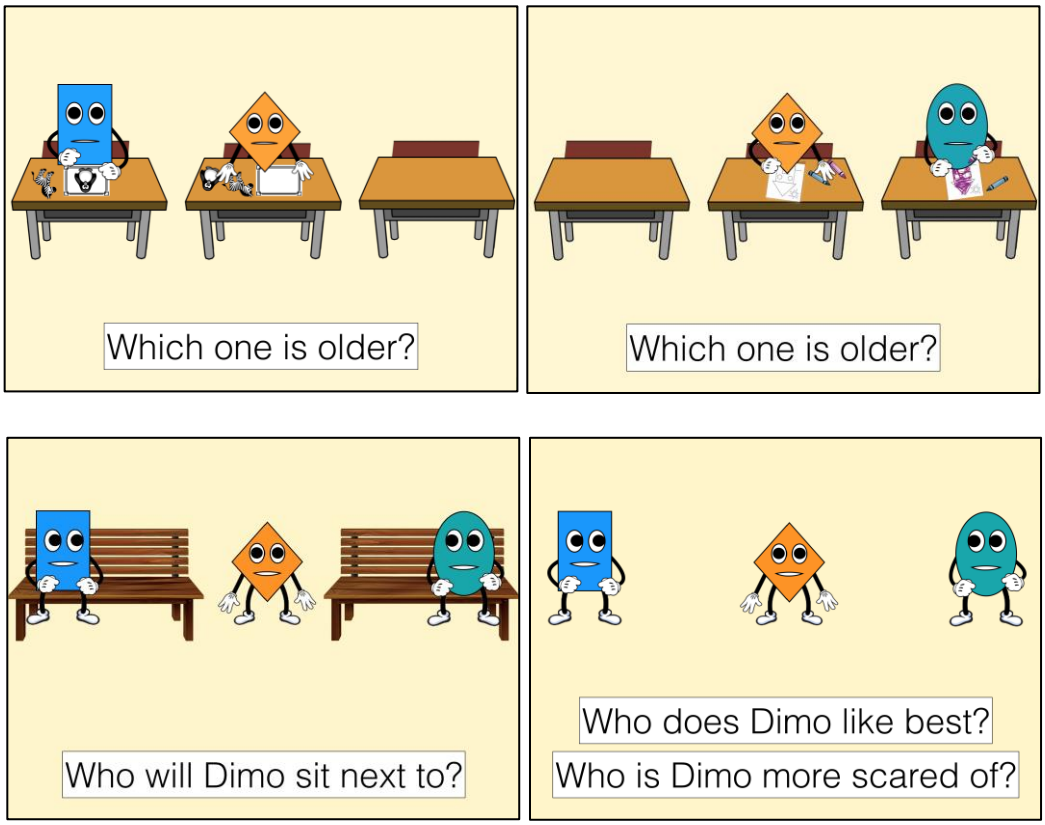


Figure 3.2. Pictures shown when asking test questions in Experiment 1

Results

Figure 3.3 shows the pattern of responses for each of the questions.

Inferring status. Children were significantly better than chance at answering who was older between the subordinate and the dominant character (140/160 correct, $p < 2.2e\text{-}16$). A binary logistic regression including age (younger vs. older) and country (UK vs. China) found no main effects or interactions. Similarly, children were better than chance at identifying who was older between the subordinate and the prestigious characters (138/160 correct, $p < 2.2e\text{-}16$). Once again, a binary logistic regression failed to find a main effect of age or country or any interaction between the two.

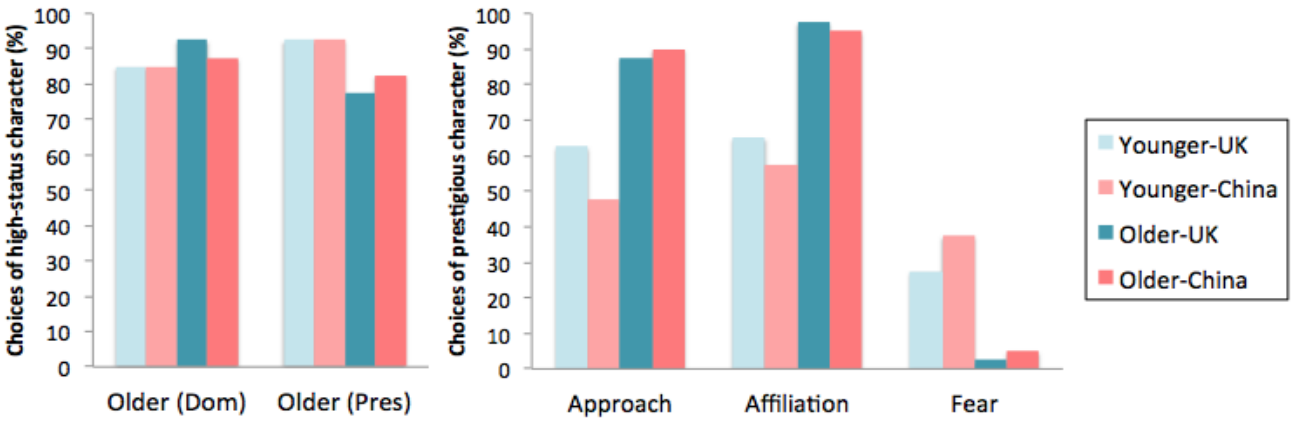


Figure 3.3. Responses to test questions in Exp. 1

Distinguishing dominance and prestige. In order to test whether children at each age distinguished between the dominant and prestigious characters, two mixed-effects models including question type (approach vs. affiliation vs. fear) and country (UK vs. China) as fixed effects and subject as a random effect were run, one on the data from younger children, and one on the data from older children. This analysis found that for both groups, there was a significant effect of question type, such that responses for the fear question were significantly lower than responses for approach or affiliation. In other words, children responded to the fear question differently from the other two questions. Neither model found an effect of country, meaning that children in UK and China responded similarly to each other. What these analyses show is that children's responses were different depending on whether they were asked who Dimo would approach or liked more on the one hand, and who Dimo was more scared of on the other. In short, children were distinguishing between the questions that the prestigious character was the more reasonable choice for and the question on which the dominant character was the more reasonable choice. A number of analyses listed below tested were conducted on the questions separately.

Approach question. Children were successful at answering who the subordinate character preferred, choosing the prestigious character significantly more than expected by chance (115/160 choices of prestigious character, $p < 2.98e\text{-}8$). A binary logistic regression found a significant main effect of age ($b = -1.4$, $SD = .58$, $z = -2.48$, $p = .013$), but no effect of country or an interaction. Separately testing younger and older children against chance revealed younger children to respond randomly (44/80, $p = .434$), and older children to respond significantly higher than chance (71/80, $p = 4.4e\text{-}13$).

Affiliation question. As a group, children said the central character would like the prestigious character significantly more often than chance (126/160 choices of prestigious character, $p = 1.25e\text{-}13$). A regression analysis found a main effect of age ($b = -3.04$, $SD = 1.07$, $z = -2.86$, $p = .0043$), but no main effect of country, and no interaction involving the variables of age and country. Testing each of the age groups against chance found younger children to be marginally (49/80, $p = .057$) and older children to be significantly higher than chance (77/80, $p < 2.2e\text{-}16$).

Fear question. Children's responses were significantly below chance overall, picking the prestigious character less often than chance (29/160 choices of prestigious character, $p < 2.2e\text{-}16$). According to a binary logistic regression, there was a main effect of age ($b = -2.7$, $SD = 1.07$, $z = -2.5$, $p = .012$), but no effect of country or an interaction. Both younger (26/80 correct, $p = .002$) and older children (3/80 correct, $p < 2.2e\text{-}16$) chose the prestigious character significantly less than expected by chance.

Discussion

Children at both ages recognized that the central character was lower in status to both the dominant and the prestigious individuals. Responses were consistently high for older and

younger children, and for children in the UK and China. With regard to questions differentiating between dominance and prestige, however, there were differences between the older and younger children. For all three questions, older children outperformed their younger counterparts. Nonetheless, a comparison of responses showed that both younger and older children's responses were significantly different for the questions on which the prestigious character was the correct choice (i.e. who Dimo would sit next to, and who Dimo liked more) and the question on which the dominant character was the correct choice (i.e. who Dimo was more scared of). Thus, although there is a difference in the degree of differentiation that younger and older children showed, even younger children distinguished between the dominant and prestigious characters. This finding suggests that Hawley's (1999) proposed bifurcation of strategies at around eight years does not translate to the same age being the cutoff for children distinguishing between dominance and prestige as observers.

Despite the contrast between younger and older children, responses on these differentiation questions once again did not show any differences between children in the UK and China. The lack of difference across any of the questions between British and Chinese participants shows that the cues provided to demarcate dominance and prestige, as well as the inferences following from them, were equally accessible for participants in both countries. Further, the culturally different approaches to thinking about prestige across the two countries appears to not have an effect on how cues to prestige and dominance are understood by children. This paved the way for asking about a possible cultural difference in expectations about the outcome of conflicts. The next experiment takes the first step by asking whether adult participants show any evidence of this.

Experiment 2

This experiment asked adults in the UK and China who would come out as the winner in a conflict over a desired resource between the central character and the dominant individual on the one hand, and between the central character and the prestigious individual on the other. As discussed above, the ethnography of Chinese culture (Kajanus, in prep.; Xu, 2017) points towards a strong value of yielding by high-status prestigious individuals towards lower-status ones. Importantly, this particular norm does not appear to have a corresponding counterpart in Euro-American culture. Therefore, the aim of this experiment was to test whether the value placed on yielding to others translates to differences in third-party expectations about yielding between adults in China and the UK. While the ethnographic evidence provides a strong suggestion for this to be the case, it is an open question whether it would be borne out empirically. Thus, it is possible that directly testing expectations about who should yield would find a difference between British and Chinese adults, particularly when the question concerns a conflict between prestigious and subordinate individuals. Further, if the difference does manifest empirically, the results would provide us with a sense of the magnitude of the effect. Alternatively, it is possible that adults in the two countries would respond similarly to each other, and that the norm, while culturally important, does not lead Chinese adults to have different expectations about third-party yielding.

In order to avoid using language that would make assumptions about how the end result of the conflict came about (e.g. winning or taking as opposed to yielding or giving up), the choices were presented to participants non-verbally. Participants saw two possible outcomes for a conflict scene, such that each picture showed one of the individuals happily in possession of the desired resource, and they were asked which one would happen.

Method

Participants. Forty adults in the UK were recruited through Amazon Mechanical Turk (M = 32.9 years, 12 female). All had lived more than 10 years in the UK, and the overwhelming majority had spent their whole lives in the UK. All participants identified as being white (European) or black[1]. Additionally, forty adults in China (M = 28.1 years, 15 female) were recruited through an ad on the crowdsourcing website Witmart. (Amazon mTurk is not a viable option for recruiting in China.) All of the Chinese participants had lived their entire lives in China and identified as being of Han ethnicity[2]. Participants were paid 1 USD for their time.

Materials. The animations used in Experiment 1 to establish the roles of the three characters were used again. Two new animations were prepared depicting conflict scenes. In one scene, Dimo and the oval character started out on opposite sides of a walkway, at the center of which an ice-cream cone stood on a table. They simultaneously expressed excitement about the ice-cream (e.g. in English saying, "Wow!"). They then moved towards the center while smiling. Upon arriving at the center, they stopped, their smiles changed into neutral expressions, and they both raised their arms, reaching for the ice-cream. The animation ended at this point. Two pictures were made to present two possible resolutions to the scene (see Figure 3.4). One showed the central character, Dimo, holding the ice-cream and smiling while the oval character looked on with a neutral expression, and the other showed the opposite, i.e. the oval character smiling and holding the ice-cream while Dimo looked on with a neutral expression. The second scene involved Dimo and the rectangle character. The structure of the scenes and the resolution

[1] According to preset exclusion criteria, participants who indicated any racial group other than white (European) or black, or those who had lived in the UK less than 10 years were excluded, as it was not clear what cultural expectations they may have about yielding.

[2] According to preset exclusion criteria, participants who identified as any ethnicity other than Han or those who had lived outside of China for an extended period of time were excluded because the ethnographic evidence came from Han individuals living in China.

pictures were similar to the first scene, but the desirable object in this scene was a chocolate popsicle. Across both conflict scenes, the characters' relative sides to each other were preserved (i.e. Dimo on the left and the oval character on the right in one scene, and Dimo on right and the rectangle character on the left in the other scene). Once again, the correspondence between the dominant and prestigious characters and the oval- and rectangle-shaped figures was counterbalanced across participants.

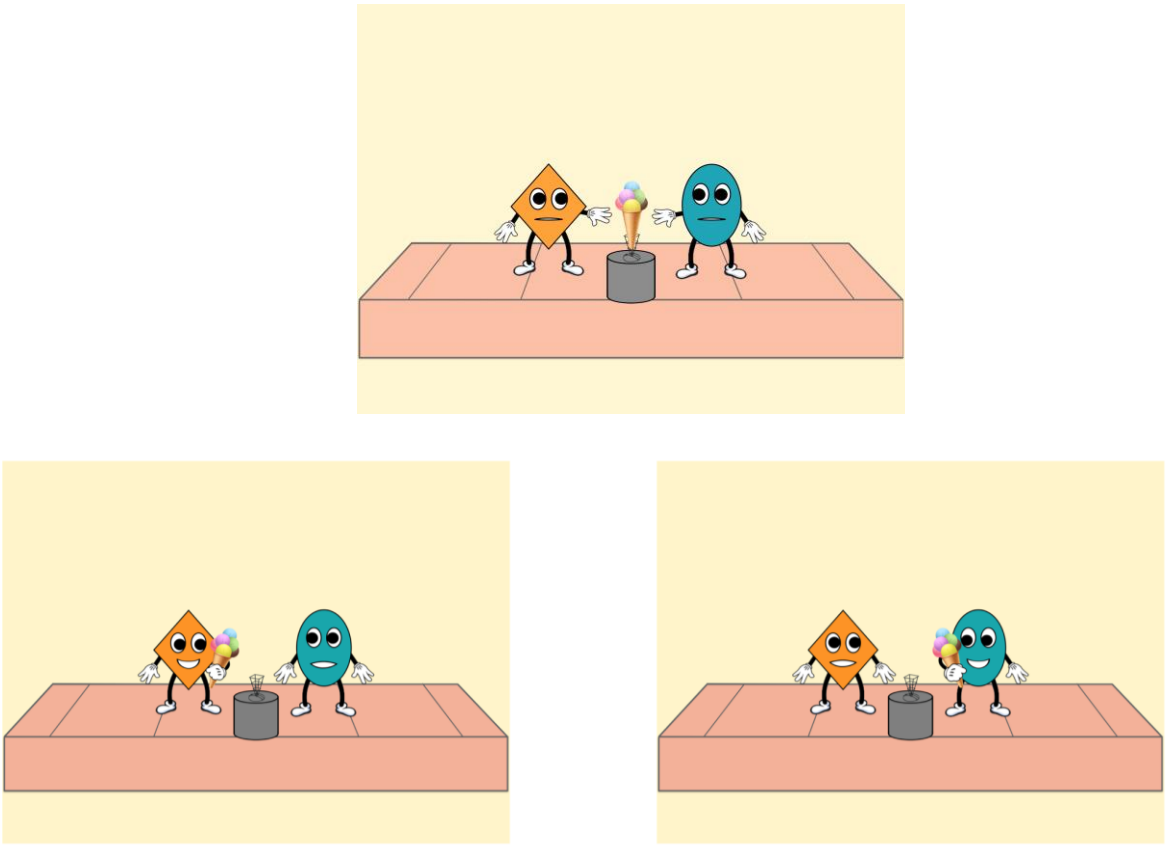


Figure 3.4. The ending of the conflict scene between Dimo and oval character (top), and the two possible resolutions presented to participants (bottom).

Procedure. Participants were informed at the start that the experiment was intended for children and that researchers were interested in their intuitions. The start of the procedure was the same as Experiment 1. Dimo was introduced, and then participants watched the dominance

and prestige scenes once each. They then watched the conflict scene for the first character they had seen in the experiment, i.e. if a participant saw the prestige scene first, they also watched the conflict scene between Dimo and the prestigious character first. Above the video showing the conflict scene, participants were told, "After this video, we will ask you what's going to happen next. Please watch the video completely." Afterwards, participants saw the resolution pictures side by side and were asked, "What do you think will happen next?" Thus, each participant provided their response for who would win the desired resource between Dimo and the dominant character, and Dimo and the prestigious character. After responding for both cases, they were asked for explanations of each choice, "You picked this picture when we asked you what would happen. Why do you think this one will get it?"

Coding. The dependent variable on the forced-choice questions was the percentage of choices for the dominant or prestigious characters. Explanations were coded into the following categories:

- *Status.* Referring to the status difference in some way, by focusing on the subordinate or high-status characters, e.g. that one was forcing him to choose the zebra, the blue one likes to bully others, Dimo always does what others tell him to do.
- *Norm.* Appealing to a moral or social norm, typically turn-taking, sharing, yielding, e.g. Dimo already got one, you should share, that would be fair.
- *Physical.* Relating the outcome to a physical property of the character in some way, e.g. that one is faster, was standing closer, this one has the same shape as the ice-cream.

- *Other.* Miscellaneous explanations that do not fit into other categories, e.g. I don't know, this one likes it more, this one wants to have it.

Results

Figure 3.5 shows the rate of responses for both groups of adults. Overall, participants were significantly more likely than chance to choose the high-status character in answering the yielding questions (131/160 correct, $p < 2.2e\text{-}16$). A mixed-effects binary logistic regression including country (China vs. UK), case type (Dimo against dominant vs. Dimo against prestigious) and their interaction term as fixed effects, and subject as a random effect (glmer(Response ~ Country*CaseType + (1|Subject)) found a significant interaction between country and case type ($b = -12.06$, $SD = 3.84$, $z = -3.14$, $p = .0017$). Separate comparison of Chinese and UK participants for each case revealed a significant difference for the prestige case ($b = -1.54$, $SD = .7$, $z = -2.2$, $p = .027$), but no difference in the dominance case ($b = .16$, $SD = .57$, $z = .29$, $p = .77$). Both groups were significantly higher than chance on both questions (Chinese adults-Dominance, 33/40 correct, $p = 4.2e\text{-}5$; Chinese adults-Prestige: 29/40 correct, $p = .006$; UK adults-Dominance: 32/40 correct, $p = .0002$; UK adults-Prestige: 37/40 correct, $p = 1.9e\text{-}8$). Explanations among both groups predominantly appealed to status (84% among UK sample, 88% among China sample), showing that participants were reasoning about the conflict in light of the status differences between the characters.

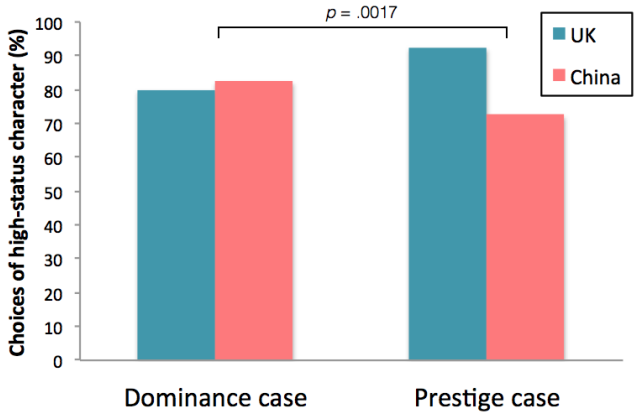


Figure 3.5. Adults' responses to yielding questions in Exp. 2

Discussion

This experiment tested expectations about who would end up with a resource desired by both parties when the subordinate character was pitted against a dominant character or (separately) against a prestigious one. Both British and Chinese adults inferred that the higher-status character would win the resource, regardless of whether the status was rooted in dominance or prestige. However, Chinese adults were significantly less likely to choose the prestigious person as the winner compared to British adults. What this finding suggests is that for Chinese participants, the answer to who would claim the resource is more complicated than simply considering the difference in status. Thus, we found support for the hypothesis that the value on yielding in Chinese culture translates into differences in third-party expectations. Nevertheless, this effect is subtle, and does not mean that Chinese adults expect the subordinate to win, or that they have no prediction at all. Given its subtlety, it is important to explore whether any difference can be observed among children in the two countries on the same task. The next experiment addresses this.

Experiment 3

The aims of this experiment were twofold. First, to test whether children of different ages would differ in their inferences about the winner of a conflict between the subordinate character and each of the high-status characters. Second, to examine whether children in the UK and China would respond differently from each other. Given the difference found between adults in the two countries in Experiment 2, this experiment examines the developmental trajectory of the inference. Possibly, children in both age groups may respond like adults in their countries, thus also demonstrating the difference observed between the adults in the prestige case. On the other hand, perhaps learning the norm of yielding to others when in positions of prestige is difficult for Chinese children, and takes longer to develop. If so, it is possible for children to start off the same across the two countries, and only for older children to show signs of a difference akin to what adults showed.

Method

Participants. Forty one children in the younger group (M = 5.8 years, range = 5-7.3 years, 21 female) and forty children in the older group (M = 10.33 years, range = 9.2-11.2 years, 20 female) participated in London, UK. Similarly, 40 younger children (19 female, M = 5.8 years, range = 5.1-6.3 years) and 40 older children (20 female, M = 10.98 years, range = 10.3-12 years) took part in the study in Nanjing, China. Children were recruited from and participated at the same schools as Experiment 1.

Materials. The materials were identical to those in Experiment 2.

Procedure. The general procedure was the same as Experiment 2. Children were introduced to the central character, then watched the dominance and prestige scenes before watching the conflict scenes and answering the yielding questions. In contrast to adults who

watched the dominance and prestige scenes once, children saw each twice in a row. Before watching the conflict scene, the experimenter said, “After this video, I’m going to ask you what’s going to happen next. Let’s watch!” Afterwards, participants saw the resolution pictures side by side and were asked, “What do you think will happen next? Is this going to happen, or is this going to happen?” Finally, the experimenter asked children for explanations regarding their choices, saying, “You picked this picture when I asked what would happen. Why do you think this one will get it?

Coding. Once again, the dependent variable was the percentage of choices of the dominant or prestigious characters. Explanations were coded as in the previous experiment into the following categories: status, norm, physical, and other.

Results

Responses are presented in Figure 3.6. A binary logistic regression analysis including age group (younger vs. older) as a fixed effect and subject as a random effect found a main effect of age (b = .36, SD = .11, z = 3.2, p = .001). Given the difference between the younger and older children, the two age groups were analysed separately.

Younger children. This age group performed at chance level overall (82/162 correct, p = .94). A mixed-effects binary logistic regression analysis including country (China vs. UK) and case type (dominance vs. prestige) and their interaction as fixed effects and subject as a random effect found no main effects or interaction.

Older children. A comparison against chance showed that older children responded with better than chance accuracy overall (108/160 correct, p = 1.1e-05). A regression analysis including country (China vs. UK) and case type (dominance vs. prestige) and their interaction as fixed effects and subject as a random effect found a significant main effect of country (b = -1.04,

SD = .53, z = -1.97, *p* =.049) and no other main effect or interaction. Given the difference between Chinese and British children, each group was compared to chance separately, revealing the UK sample to be above chance (61/80 correct, *p* = 2.7e-06), and the Chinese sample to be at chance (47/80 correct, *p* = .15) in their responses.

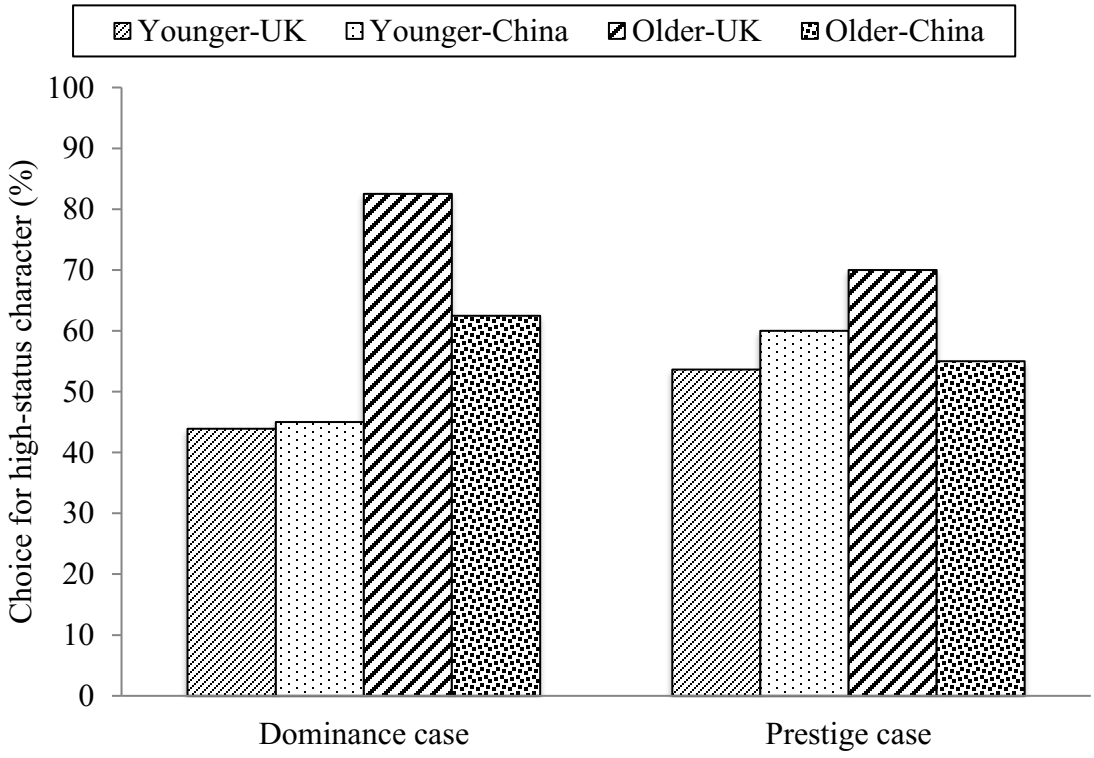


Figure 3.6. Choices (%) for high-status character in the dominance and prestige cases for each age group

Explanations. Responses were categorized into the four categories of status, norm, physical and other (see Figure 3.7). The categorization of responses provided two main insights. First, explanations appealing to status were almost non-existent among the younger group (~5%), increased with age, such that most of the older children mentioned status (~65%), and culminated in almost unanimous mention of status by adults (~90-95%). The pattern of increase in status-relevant explanations was true both when focusing on participants who chose the high-

status character (i.e. dominant or prestigious) and when considering responses from all participants, regardless of how they responded on the yielding questions (see Table 3.3). Thus, the younger group appeared to be construing the scenes in terms of shallow, surface features instead of status, but older children and even more so the adults were cognizant of the role of status in determining the outcome of the conflicts. Second, the proportion of status explanations offered by each age group was similar across British and Chinese participants, as is clear in both Table 3.3 and Figure 3.7. In other words, participants in the UK and China were just as likely as each other to give status-relevant explanations.

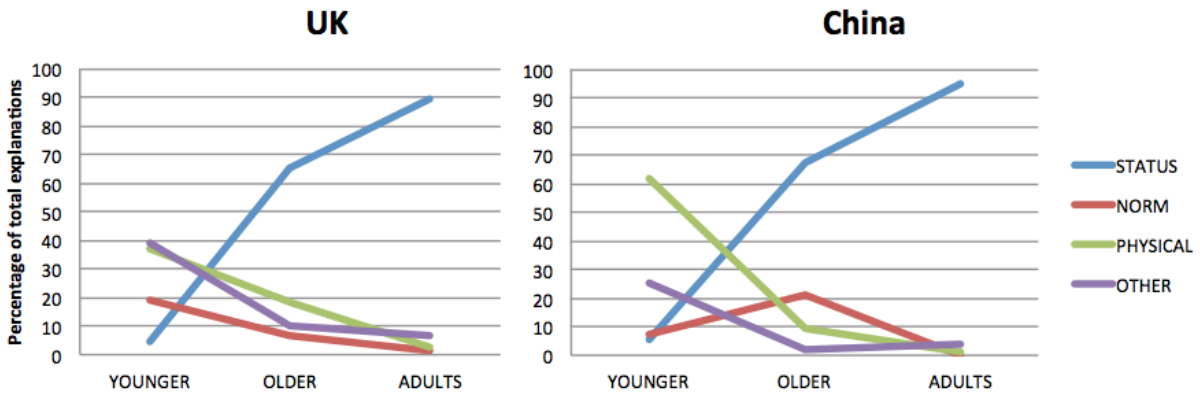


Figure 3.7. Breakdown of explanations from all participants by category across ages and countries. Proportion of status-relevant responses (out of all explanations) increased with age, but did not differ by country, being similar across UK and China.

Table 3.3. Percentage of status-related explanations out of all explanations

Age group	Country	Subjects who chose high-status character	All subjects
Younger	UK	5%	3%
	China	6%	6%
Older	UK	65%	55%

	China	68%	60%
Adults	UK	89%	84%
	China	95%	88%

Additionally, we counted the number of times Chinese participants mentioned one party yielding the resource to the other, using terms that are typically used to talk about the concept (e.g. *rang*). First, this tally further corroborated that younger children did not view the scenes as relevant to status since they mentioned yielding only three times (4% of all explanations), while older children did so 26 times (33% of all explanations). Second, among the older children who chose the high-status character as the winner, children were just as likely to mention yielding in the prestige case as in the dominance case (40% of prestige case responses vs. 23% of dominance case responses, $\chi^2(1) = 1.5$, $p = .21$). Thus, children were equally likely to think that the low-status character would yield to either the prestigious or the dominant character. Third, among those who chose the subordinate character as the winner, children were more likely to think that this would come about as a result of the prestigious character yielding to Dimo than the dominant character yielding (50% of prestige case responses vs. 13% of dominance case responses, $\chi^2(1) = 4.95$, $p = .026$)[3]. This suggests that older children differentiated in how they thought dominance and prestige affected the conflict outcomes, as a larger number of them harbored the expectation that the prestigious character would yield to the subordinate.

Discussion

[3] For comparison, we also tallied mentions of yielding (e.g. letting the other one have it) among the older British children. Unlike Chinese children, rates of yielding explanations were similar across prestige and dominance cases among those who chose the subordinate character (25% vs. 28% respectively, $\chi^2(1) = .03$, $p = .86$).

Younger children (5-7 years) in neither country made consistent inferences about the outcome of the conflicts. This is a surprising finding, given their clear success in identifying the high-status parties in Experiment 1 (i.e. answering "Which one is older?" correctly). The explanations show that younger children almost exclusively focused on reasons having nothing to do with status in both countries, mainly providing irrelevant and superficial responses (e.g. "He likes it." , "He was faster."). Although 5- to 7-year-old children were able to identify the higher-status parties and even distinguished between prestige and dominance in Experiment 1 (i.e. inferring that the dominant character was less likely to be liked and approached and more likely to be feared), they did not connect the status dynamics to how the conflict would be resolved.

Among the older group (9-12 years), children in the UK reacted similarly to adults in the UK such that they thought that both high-status characters would get the resource when matched against the low-status character. In contrast, older children in China did not respond systematically on either case, and made no inference about high-status characters winning. Despite their inconsistent responses, their explanations were just as focused on status as those of older British children. Further, a larger number of children who picked the low-status character as the winner thought that this would be the result of the high-status character yielding in the prestige case than in the dominance case. Thus, despite the lack of consistent inferences, there was some evidence to suggest that older children in China were aware that prestigious individuals may be more likely to yield to a subordinate. The developmental trajectory in inferences about the winner of the conflict scenes seems to be relatively protracted, as neither younger nor older children in China performed similarly to Chinese adults. This may be due to

the situational complexities that govern when a high-status individual is likely to yield in Chinese culture.

General Discussion

The present experiments yielded several important findings. First, 5- to 7-year-olds (younger group) and 9- to 12-year-olds (older group) in both the UK and China easily identified both dominant and prestigious characters as high-status. Further, both age groups inferred that a prestigious character is more likely to be approached and liked, but less likely to be feared. Thus, all children distinguished between prestigious and dominant characters, although older children were significantly more successful. Second, adults in UK and China inferred that a high-status character would win in a conflict over a resource with a subordinate, regardless of whether the high status was achieved through dominance or prestige. Interestingly, Chinese adults were significantly less likely than British adults to think that the prestigious character would win. Third, older children in the UK inferred that the high-status characters would win, while older Chinese children made no systematic inference. Children in both countries, however, provided status-relevant explanations for their choices. Younger children in both countries failed to make consistent predictions about the who would win and provided shallow, status-irrelevant justifications for responses.

One of the current paper's main contributions is to the developmental literature on children's third-party understanding of hierarchical relationships. While much previous work has dealt with inferences about dominance (and possibly prestige), it remained unknown whether young children differentiate the two. The first experiment took a novel approach to this question by asking children to judge dominant and prestigious characters in a within-subjects design. The first important finding was that even for the younger group of children (5-7 years), cues to

prestige (asking for advice, following said advice, copying behavior) were enough to merit an inferences about status. In fact, children were just as successful at recognizing that the prestigious character was higher-status as they were for the dominant character. The next finding was that even younger children distinguished between prestigious and dominant characters in a third-party situation, as they chose the prestigious character significantly more when asked who the subordinate approached or liked than when asked who she feared. This finding is particularly interesting in light of the fact that children do not shun coercive (or dominance-like) strategies in their peer groups until eight years (Hawley, 1999). As our younger group were younger than eight years, the finding rules out the possibility that shunning coercive strategies is the result of realizing their divergence from prosocial strategies. Rather, children understand the difference earlier, but may fail to reject coercive practices for other reasons, which require further investigation. Moreover, the ability to distinguish dominance and prestige improved significantly with age, such that older children were far more successful at answering the questions. Finally, there were no differences between children in the UK and China in their responses to the questions about any of the questions, providing the first empirical evidence of children from non-Euro-American cultures understanding cues to status similarly to children in Europe and the US.

The paper's other contribution is to provide evidence for the role of culture in shaping inferences about hierarchical relationships. Adults in the UK and China were similar in that they both inferred that high-status characters would win against a subordinate. They did differ, however, in the rate that they made this inference with regard to the prestige case. Chinese adults were less likely to think that the prestigious person would win the resource. This difference, while subtle, is a sign of the cultural difference reflecting the value specifically placed on yielding to others when in a position of prestige (Kajanus, in prep.). The cultural difference

manifested in the older child participants as well, but in a different way. Older children in the UK (9-12 years) inferred that the high-status party would win the conflict, regardless of whether the character was prestigious or dominant. In contrast, older Chinese children responded similarly to younger children in both countries, demonstrating no systematic prediction about who would win in either of the conflict cases. Although they seem similar on the surface, children's explanations provide critical evidence that distinguishes the older Chinese children from younger children in both countries. Unlike younger children's explanations, which were shallow and unrelated to status, older children in China provided just as many status-relevant explanations for their choices as their British counterparts. This leads to the conclusion that the lack of systematic inference was due to age in younger children, but a result of culture in older Chinese children. Moreover, older children in China who thought the subordinate would win the conflict were more likely to mention the prestigious character yielding than the dominant character yielding. This piece of evidence aligns with the value of yielding by prestigious individuals. Together, these findings suggest that the difference between older children in China and the UK came is rooted in the cultural difference.

On the topic of Chinese participants' choices, two points merit further discussion. With regard to the adults, the fact that Chinese participants thought that both dominant and prestigious parties would win the conflict gives room for pause. While the choice of the dominant character as the winner is not surprising, choosing the prestigious character as the winner was not the only possible pattern of results. Given the value of prestigious individuals showing restraint and giving up resources to lower-status parties, Chinese adults might have chosen the subordinate as the winner, or given unsystematic responses as a group. A possible reason behind the actual pattern of data rather than these two alternatives could be that the situation shown in the

animations is quite simplistic and lacks finesse. Yielding to others in lower positions is a highly sophisticated social skill, one that is honed through practice, and one that is dependent on many aspects and nuances of the situation. It is possible that given the cartoonish appearance of the characters, the simplicity of the exchanges, and the prestigious character's unabashed (i.e. juvenile) desire for the resource, most adults viewed the situation as one in which the high-status character lacked the sophistication to yield. Therefore, the high-status character's winning was the more straightforward choice. A number of them did, nonetheless, possibly consider the situation in terms of the prestigious character yielding, thereby bringing about the difference with the British adults.

Another puzzling piece of the findings has to do with the older Chinese children's responses in Experiment 3. In a reverse situation from the adults, the children held no expectation of either party winning in either conflict case. The lack of a clear prediction may be understandable for the prestige case, but it would have been less surprising had children inferred that the dominant character would win. After all, unlike prestigious persons, dominant individuals are not highly regarded nor expected to yield in Chinese culture (Kajanus, in prep.). As the data do not clarify the reason for this result, we can only speculate. One likely account is that children at this age have learned that high-status parties sometimes yield to lower-status individuals, but have yet to master the rules and nuances of when this happens. They may not have yet figured out how other factors such as behavioral characteristics (e.g. prestige vs. dominance) or different hierarchies (e.g. age, gender, position) factor into this. Thus, it is possible that in both the dominance and prestige cases, some children thought the high-status individual should yield because they opted for that character. Another point to keep in mind is that dominance and prestige processes are just that, processes, rather than defining characteristics

of an individual (Henrich & Gil-White, 2001). In fact, any person may be dominant or prestigious towards different people and in different situations. While our stimuli simplified this complexity by presenting the characters as unidimensionally dominant or prestigious, it is possible that children may have thought that a dominant character could be prestigious in a different situation, and thus susceptible to the need to yield to subordinates.

Experiment 3 also unearthed a surprising finding from a developmental perspective. Although younger children recognized the status difference between the characters just as easily as older children, they did not infer that the higher-status party would come out as the winner of the conflict in either case. This failure to infer cannot be attributed to a cultural effect, as British and Chinese children performed similarly. When asked to justify their choice of who would be the winner in the conflict, children's explanations centered on superficial physical features of the characters or other shallow aspects. Overall, almost no children referenced status in their explanations, confirming that their failure to infer is a real consequence of how they construed the scene. Thus, although younger children extracted status rankings from watching the interactions between the characters in Experiment 1, they were unable to automatically incorporate this into their inference when asked about their future behaviors in Experiment 3. The reason behind this failure is unknown, but it is possible that children were considering several different aspects of the scenes instead of focusing on the key one, i.e. status. Future work could explore this possibility further by highlighting the hierarchical nature of the scenes and examining its effect on children's inferences. In fact, manipulations such as propositional scaffolding and concept priming that were found to be effective in Paper 2 may prove beneficial in this context as well.

Relying on large amounts of ethnographic evidence, Fiske (1992) laid out a theory of four elementary forms of social relationships in human societies, one of which is authority ranking, which corresponds to hierarchical relationships marked by status differences. One of Fiske's key claims about the relational models is that they are universal across different cultures and history. Being universal, however, does not mean that they are identical in their implementations. In fact, Fiske stresses that each model can be implemented in myriad ways. Thus, even though children may be endowed with an innate knowledge of each model (Thomsen & Carey, 2013), they must learn the details of what each one looks like and exactly how it operates in their particular social milieu. These experiments provide evidence for one such culturally-influenced aspect of hierarchical relationships across two cultural backdrops. The range of culturally-customized features and cues to relationships is underexplored, as is how children come to learn them and how long it takes them to do so. We offer an initial glimpse of one particular aspect and hope that future work will go on to uncover many more.

Conclusion

Infants and children, like adults, are immersed in a rich world of social relationships. They are interested in them from the start, and their understanding of them develops and diversifies with age in fascinating ways. This dissertation built on decades of prior research by tackling a small set of questions relevant to this development, provided a number of answers, and brought forth even more questions. A brief summary of the papers and a number of future directions are discussed here.

Paper 1 asked four-year-old children to infer who was friends with whom based on reports linking pairs of individuals together. This investigation was undertaken in order to gauge children's concepts of friendship, to determine whether methodological shortcomings of previous work had underestimated what they understood, and to broaden our understanding by testing new dimensions of the concept. We found that children thought that an individual was friends with another child if they had engaged in a joint activity, but not if they had an arbitrary links (Exp. 1), if they were similar on a meaningful dimension (e.g. a skill) but not if they had an arbitrary link (Exp. 2), and that they privileged joint activities over similarities (Exp. 3). The final experiment found that children thought individuals linked through a joint activity, but not individuals who had an arbitrary link, were more likely to be playmates, and more likely to like the same game, but not any more likely than chance to like the same food. Overall, the results provide a schema of children's early friendship concepts and expectations. Their concept recognizes that friendship is centered on joint activities, playing together, meaningful similarities, and a shared liking for the same games.

This set of findings in part confirms (e.g. playing is central to early friendships), extends (e.g. similarity cues can be the basis of friendship inference), and contrasts with previous work (e.g. children disregard arbitrary information when inferring friendship). Future research on this

topic should examine whether three-year-olds, younger than the children tested here, make the same inferences as the four-year-olds. According to one study, while four-year-olds think that a child would prefer to be friends with another child of the same gender, three-year-olds do not (Shutts, Roben, & Spelke, 2013). So it is possible that three-year-olds' third-party inferences, and by extension their conception of friendship, would be different when tested in a task such as the one used in this paper. Nonetheless, the above finding is focused on gender and does not speak to the broader concept of friendship, hence making this an empirical question worth investigating.

Another interesting avenue for future work is to match social category cues against friendship cues like the ones provided in this experiment, and test whether children distinguish one from the other. For instance, if participants were told that child A and child B belong to the same group and wear the same kinds of clothes, but that child C helped child A clean up her room, would children make any differentiated inferences about child B and child C? For instance, would they infer that child A prefers child B or child C? Or that child C knows more personal information about child A than child B, given their relationship? There has been much theoretical interest and empirical research focused on children's notions of social categories and social kinds, for instance centered on race, gender, or even arbitrary features (e.g. Hirschfeld, 2001). While categories are arguably a strong driving force of social dynamics, social relationships are another strong--if not stronger--current of structure and influence. The question of whether, and if so in what ways, they are differentiated in the conceptual repertoire of children would go a long way towards clarifying how social knowledge is structured in early childhood.

Finally, the paper found that four-year-olds did not harbor expectations of shared food preferences amongst individuals linked by friendship-relevant cues (e.g. joint activities). Given

the finding that infants expect individuals that have affiliated to share food preferences (Liberman, Woodward, Sullivan, & Kinzler, 2016), it is possible that infants did so because they viewed the affiliation as a signal to a communal relationship, or perhaps belonging to the same social category. Thus, the question arises whether preschoolers would infer shared food preferences between family members or between group members, even though they did not for friends.

Paper 2 asked whether preschool-aged children infer affiliation from observed imitation, and whether they expect an imitator to continue imitating the same target as before in the future. This project was partly aimed at clarifying why four-year-olds failed to infer affiliation between an imitator and target (Over & Carpenter, 2013), while infants demonstrated this expectation (Powell & Spelke, 2017). We found that even three-year-olds have an explicit representation of "copying" (Exp. 1). Despite this, three- to five-year-olds were not particularly successful at readily inferring affiliation or continued imitation (Exp. 2), suggesting that they might not spontaneously encode the imitation while watching it, or realize that it is relevant to the inferences they were asked to make. In line with this conjecture, preschoolers were significantly better at making the same inferences if the imitation in the events was emphasized for them early on in the task through propositional scaffolding (e.g. "Did Jenny move her head like Alice, or her leg like Zoe?"), or if the concept of imitation was primed ahead of the task (Exp. 4). Overall, the paper does not ultimately answer why infants have succeeded in the past and four-year-olds have not. It does, however, show that the discontinuity may be less severe than suggested by previous findings because with propositional scaffolding, even three-year-olds can make the inference that Over and Carpenter's (2014) four-year-olds failed to make. Thus, one account of the apparent discontinuity could be that preschoolers may consider a number of possible hypotheses--

including but not limited to imitation--when asked who the imitator likes, but that infants' succeed because their ability to consider any possibility besides imitation is much more limited. As a result, when the correct hypothesis, i.e. imitation, is highlighted for preschoolers, they easily come to the conclusion that the imitator prefers her target and that she will continue to imitate the same person going forward. This proposal is clearly difficult to test, but attempts at gaining empirical evidence for or against it would be valuable.

Another point that was not resolved by this paper is whether it is necessary to highlight imitation in order for preschoolers to make the ensuing inferences, or whether any social priming could have had the same effect. Over and Carpenter (2014) told children at the start of their task that they would be asked about who the central character likes more. This highlighted affiliation as the inference of interest. And while it did not help four-year-olds, who failed nonetheless, it possibly helped five-year-olds, who succeeded in contrast to the failure of five-year-olds who received no support in our study (Exp. 2). Thus, it would be worth testing whether a stronger social manipulation, such as priming the concept of friendship, may achieve the same findings as the propositional scaffolding or priming of imitation in our study. A further question concerns the nature of the inference that children make when they answer the affiliation question correctly. In these experiments, we asked children who the imitator "liked best", which is not indicative of a single relationship type. For instance, children may think of the imitator and target as friends that have mutual affection for one another. Over and Carpenter (2014) reported that five-year-olds inferred that a target of imitation was the "boss", so there seems to be an element of hierarchical relationships embedded in imitation. In line with this, children might view the two individuals as engaged in a hierarchical prestige relationship, where the imitator (i.e. the low-status party) harbors more affection for the target (i.e. high-status party) than the other way

around. It is alternatively possible that children view the affiliation in a non-committal manner so far as it relates to a specific relationship, and would require more information before making a judgment.

In Paper 3, we explored the role of age and culture (China vs. UK) on expectations about the cues and consequences of dominance and prestige. This study explored not only whether children differentiate dominance from prestige, but also whether culturally-specific expectations about the behaviors of prestigious and dominant people are learned at the same time as they make this distinction. In Chinese culture, prestigious individuals are skilled at and expected to yield to others, even those in lower status to themselves, whereas there is no such expectation in the UK. Experiment 1 found that children at around 6 and 10 years in both countries used relevant cues (e.g. being asked for advice vs. forcing opinion on others) to recognize prestigious and dominant individuals as high-status, and that they differentiated between the two relationships. The next two experiments tested adults' and children's expectations about the winner of a conflict between high- and low-status parties. We found a cultural difference such that adults in China were less likely than those in the UK to infer that a prestigious person would win against a subordinate. Older children in the UK inferred that both dominant and prestigious individuals would win against the subordinate, but Chinese children of the same age had no consistent expectations, even though they viewed the scenes in terms of status. Younger children in both countries also failed to make systematic inferences. This paper provides evidence of a culturally-influenced expectation about hierarchical relationships, both in adulthood and in childhood.

A puzzling finding from this paper was that five- to seven-year-old children failed to infer that a high-status individual would win access to a desired resource, despite recognizing the

status difference. This is surprising because preschoolers infer higher status from a host of cues (e.g. decision-making ability). More specifically, they expect higher-status individuals to possess more resources (Charafeddine et. al, 2015). In fact, even infants expect a character shown to be dominant by controlling a territory to succeed at winning a desired resource (Mascaro & Csibra, 2012). In light of all this, understanding why even relatively old children failed to make a similar inference in the current experiment would be valuable. One possibility is that this null finding may signal another case of discontinuity between infancy and childhood, such as that reported for theory of mind understanding (success in infancy, e.g. Onishi & Baillergon, 2005, vs. failure until four years, e.g. Wellman, Cross, & Watson, 2001) or for solidity constraints on objects passing barriers (success in infancy, e.g. Spelke, Breinlinger, Macomber, & Jacobson, 1992, vs. failure until three years, e.g. Berthier, DeBlois, Poirier, Novak, & Clifton, 2000). Another possibility is that the discontinuity is less robust, and that given the right conditions, children would succeed at the inference. One way to tease these two apart would be to harness the methods from Paper 2 in such a task. In other words, manipulations like propositional scaffolding (i.e. drawing attention to the fact that one party is dominant over the other), or priming the concept of dominance ahead of the task might change children's failure into success.

Finally, this paper showcases the utility of cross-cultural experimental research that is inspired by anthropological findings. Importantly, such work can explore universal aspects of social relationship concepts from a separate angle than that of research with infants. Clearly, one of the key merits of infancy research on the topic of social relationships is its potential to discover universal features of understanding. Infants' concepts are more likely than those of children to be innate and universal due to their limited social experiences, and this is even more true the younger infants get. Even so, investigations with children from different cultures provide

a complementary wedge into the problem of what is universal, with the added benefit of allowing us to test more nuanced and complex concepts than with infants. In the case of this paper, Chinese and British children differentiated between dominance and prestige at the same age (around 6 years). This was the case despite presumably widely differing types and amounts of talk and exposure to the two concepts in the two countries given their cultural differences. As such, it provides initial support for the hypothesis that this distinction between dominance and prestige, and the age at which children achieve it, may be universal. More broadly, cross-cultural work that compares contexts with considerable distinctions can shed light on similarities in addition to differences, and both are theoretically significant. That is to say, similarities across populations with documented differences are indirect, but still invaluable evidence in support of universality. Taking seriously the promise of ethnographic evidence in locating possible points of interest and charting out the similarities and differences is an important strategy for understanding social relationship concepts.

References

Aboud, F., Mendelson, M., & Purdy, K. (2003). Cross-race peer relations and friendship quality. *International Journal of Behavioral Development, 27*(2)*,* 165-173.

Beier, J. S., & Spelke, E. S. (2012). Infants' developing understanding of social gaze. *Child Development, 83*(2), 486-496.

Bernard, S., Castelain, T., Mercier, H., Kaufmann, L., Van der Henst, J. B., & Clément, F. (2016). The boss is always right: Preschoolers endorse the testimony of a dominant over that of a subordinate. *Journal of Experimental Child Psychology*, *152*, 307-317.

Berndt, T. J. (1981). Relations between social cognition, nonsocial cognition, and social behavior: The case of friendship. In J.H. Flavell & L. Ross (Eds.), *Social cognitive development: Frontiers and possible futures*, (pp. 176-199). Cambridge, UK: Cambridge University Press.

Berthier, N. E., DeBlois, S., Poirier, C. R., Novak, M. A., & Clifton, R. K. (2000). Where's the ball? Two-and three-year-olds reason about unseen events. *Developmental Psychology*, *36*(3), 394-401.

Bigelow, B. J. (1977). Children's friendship expectations: A cognitive-developmental study. *Child Development*, 246-253.

Bigelow, B. J., & La Gaipa, J. J. (1975). Children's written descriptions of friendship: A multidimensional analysis. *Developmental Psychology*, *11*(6), 857-858.

Brey, E., & Shutts, K. (2015). Children use nonverbal cues to make inferences about social power. *Child Development*, *86*(1), 276-286.

Bugental, D. B. (2000). Acquisition of the algorithms of social life: A domain-based approach. *Psychological Bulletin*, *126*(2), 187-219.

Carey, S. (2009). *The origin of concepts*. Oxford, UK: Oxford University Press.

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.

Carpenter, M., Uebel, J., & Tomasello, M. (2013). Being mimicked increases prosocial behavior in 18-month-old infants. *Child Development, 84*(5)*,* 1511-1518.

Cashdan, E. (1998). Adaptiveness of food learning and food aversions in children. *Social Science Information, 37*, 613-632.

Chalik, L., & Rhodes, M. (2014). Preschoolers use social allegiances to predict behavior. *Journal of Cognition and Development, 15*(1), 136-160.

Chalik, L., Rivera, C., & Rhodes, M. (2014). Children's use of categories and mental states to predict social behavior. *Developmental Psychology, 50*(10), 2360-2367.

Charafeddine, R., Mercier, H., Clément, F., Kaufmann, L., Berchtold, A., Reboul, A., & Van der Henst, J. B. (2015). How preschoolers use cues of dominance to make sense of their social environment. *Journal of Cognition and Development*, *16*(4), 587-607.

Charafeddine, R., Mercier, H., Clément, F., Kaufmann, L., Reboul, A., & Van der Henst, J. B. (2016). Children's allocation of resources in social dominance situations. *Developmental Psychology*, *52*(11), 1843-1857.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910.

Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, *33*(1), 46-56.

Clark, M. S., & Mills, J. (1994). Communal and exchange relationships: Controversies and research. In R. Erber & R. Gilmour (Eds.), *Theoretical frameworks for personal relationships*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Corsaro, W.A. (1981). Friendship in the nursery school: social organization in a peer environment. In S. R. Asher, & J. M. Gottman, (Eds.), *The development of children's friendships* (pp. 207-241). Cambridge, U.K.: Cambridge University

Press.

Damon, W. (1979). *The social world of the child*. San Francisco, CA: Jossey-Bass.

Danziger, K. (1957). The child's understanding of kinship terms: A study in the development of relational concepts. *The Journal of Genetic Psychology, 91,* 213–232.

Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category based induction. *Child Development, 77*, 539-553.

Diesendruck, G., Salzer, S., Kushnir, T., & Xu, F. (2013). When choices are not personal: The effect of statistical and social cues on children's inferences about the scope of preferences. *Journal of Cognition and Development, 16*(2), 370-380.

Drews, C. (1993). The concept and definition of dominance in animal behaviour. *Behaviour*, *125*(3), 283-313.

Dunbar, R. I. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*(04), 681-694.

Dunn, J. (1993). *Young children's close relationships: Beyond attachment.* London: Sage.

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

Dunn, J., Cutting, A. L., & Fisher, N. (2002). Old friends, new friends: Predictors of children's perspective on their friends at school. *Child Development*, *73*(2), 621-635.

Enright, E. A., Gweon, H., & Sommerville, J. A. (2017). 'To the victor go the spoils': Infants expect resources to align with dominance structures. *Cognition*, *164*, 8-21.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14),

9602-9605.

Fawcett, C. A., & Markson, L. (2010a). Similarity predicts liking in 3-year-old children. *Journal of Experimental Child Psychology, 105*(4), 345-358.

Fawcett, C. A., & Markson, L. (2010b). Children reason about shared preferences. *Developmental Psychology, 46*(2), 299-309.

Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological Review*, *99*(4), 689-723.

Frazier, B. N., Gelman, S. A., Kaciroti, N., Russell, J. W., & Lumeng, J. C. (2012). I'll have what she's having: The impact of model characteristics on children's food choices. *Developmental Science*, *15*(1), 87-98.

Furman, W., & Bierman, K. L. (1983). Developmental changes in young children's conceptions of friendship. *Child Development*, 549-556.

Gazes, R. P., Hampton, R. R., & Lourenco, S. F. (2015). Transitive inference of social dominance by human infants. *Developmental Science*, *20(2).*

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature, 415*(6873), 755-755.

Gülgöz, S., & Gelman, S. A. (2017). Who's the boss? Concepts of social power across development. *Child Development*, *88*(3), 946-963.

Guralnick, M. J., & Groom, J. M. (1988). Friendships of preschool children in mainstreamed playgroups. *Developmental Psychology, 24*(4), 595-604.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-559.

Hawley, P. H. (1999). The ontogenesis of social dominance: A strategy-based evolutionary perspective. *Developmental Review, 19*(1), 97-132.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission.

Evolution and Human Behavior, *22*(3), 165-196.

Heyman, G., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology, 77*(1), 1-19.

Hirschfeld, L. A. (1989). Rethinking the acquisition of kinship terms. *International Journal of Behavioral Development, 12(4),* 541–568.

Hirschfeld, L. A. (2001). On a folk theory of society: Children, evolution, and mental representations of social groups. *Personality and Social Psychology Review*, *5*(2), 107-117.

Horner, V. & Whiten, A. 2005 Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition, 8*,164–181.

Howes, C. (1983). Patterns of friendship. *Child Development*, 1041-1053.

Howes, C. (1987). Social competence with peers in young children: Developmental sequences. *Developmental Review, 7*(3), 252-272.

Howes, C. (2009). Friendship in early childhood. In K.H. Rubin, W.M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 180-194). New York, NY: Guilford Press.

Howes, C., & Farver, J. (1987). Toddlers' responses to the distress of their peers. *Journal of Applied Developmental Psychology, 8*(4), 441-452.

Hruschka, D. J. (2010). *Friendship: Development, ecology, and evolution of a relationship* (Vol. 5). Berkeley, CA: University of California Press.

Johnson, S. C., Dweck, C. S., & Chen, F. S. (2007). Evidence for infants' internal working models of attachment. *Psychological Science*, *18*(6), 501-502.

Jones, S. S. (2007). Imitation in infancy the development of mimicry. *Psychological Science, 18*(7), 593-599.

Jones, S. S. (2009). The development of imitation in infancy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1528),

2325-2335.

Kajanus, A. (in prep). *Cooperation, Competition and Violence in Child Development – A Cross-Cultural Perspective.* Monograph proposal to be submitted to Stanford University Press.

Kalish, C. W., & Shiverick, S. M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development, 19*(3), 401-416.

Kavanagh, L. C., Suhler, C. L., Churchland, P. S., & Winkielman, P. (2011). When it's an error to mirror: The surprising reputational costs of mimicry. *Psychological Science, 22*(10), 1274-1276.

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 221-236.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577-12580.

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27*(4), 623-634.

Király, I., Csibra, G., & Gergely, G. (2004, May). *The role of communicative-referential cues in observational learning during the second year.* Poster presented at the 14th Biennial International Conference on Infant Studies, Chicago, IL.

Ladd, G. W., & Kochenderfer, B. (1996). Linkages between friendship and adjustment during early school transitions. In W. M. Bukowski, A. F. Newcomb, & W. W. Hartup (Eds.), *The company they keep* (pp. 322–345). Cambridge, U.K.: Cambridge University Press

LaFraniere, P., & Charlesworth, W. R. (1983). Dominance, attention, and affiliation in a preschool group: A nine-month longitudinal study. *Ethology and Sociobiology*, *4,* 55–67.

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon

effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior, 27*(3), 145-162.

Liberman, Z., Kinzler, K. D., & Woodward, A. L. (2014). Friends or foes: Infants use shared evaluations to infer others' social relationships. *Journal of Experimental Psychology: General*, *143*(3), 966-971.

Liberman, Z., Woodward, A. L., Sullivan, K. R., & Kinzler, K. D. (2016). Early emerging system for reasoning about the social nature of food. *Proceedings of the National Academy of Sciences*, *113*(34), 9480-9485.

Lourenco, S. F., Bonny, J. W., & Schwartz, B. L. (2015). Children and Adults Use Physical Size and Numerical Alliances in Third-Party Judgments of Dominance. *Frontiers in Psychology*, *6*.

Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1567), 1158-1167.

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist, 45*(4), 513-520.

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition, 124*(2), 227-233.

Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12 months. *Cognition*, *123*(1), 50-60.

Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences*, *109*(18), 6862-6867.

Mascaro, O., & Csibra, G. (2014). Human infants' learning of social structures the case of dominance hierarchy. *Psychological Science*, *25*(1), 250-255.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838–

850.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75-78.

Meltzoff, A. N., & Moore, M. K. (1999). Persons and representation: Why infant imitation is important for theories of human development. In J. B. G. Nadel (Ed.), *Imitation in infancy: Cambridge studies in cognitive perceptual development* (pp. 9–35). New York, NY: Cambridge University Press.

Molina, M., Van de Walle, G. A., Condry, K., & Spelke, E. S. (2004). The animate-inanimate distinction in infancy: Developing sensitivity to constraints on human actions. *Journal of Cognition and Development*, *5*(4), 399-426.

Neal, J. W., Neal, Z. P., & Cappella, E. (2014). I know who my friends are, but do you? Predictors of self-reported and peer-inferred relationships. *Child Development*, *85*(4), 1366-1372.

Newcomb, T. M. (1956). The prediction of interpersonal attraction. *American Psychologist, 11*(11), 575-586.

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science, 11*(5), 722-731.

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*(1), 222-231.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? S*cience*, *308*(5719), 255-258.

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., Clark, S. & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology, 26*(10), 1334-1338.

Over, H., & Carpenter, M. (2009a). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science, 12*(3).

Over, H., & Carpenter, M. (2009b). Eighteen-month-old infants show increased helping following priming with affiliation. *Psychological Science, 20*(10), 1189-1193.

Over, H., & Carpenter, M. (2013). The social side of imitation. *Child Development Perspectives, 7*(1), 6-11.

Over, H., & Carpenter, M. (2014). Children infer affiliative and status relations from watching others imitate. *Developmental Science*, *18*(6), 917-925.

Piaget, J. (1928). *Judgement and reasoning in the child.* London: Routledge and Keagan Paul.

Powell, L., & Spelke, E. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*(41), E3965-E3972.

Powell, L. & Spelke, E. (2017). Human infants' understanding of social imitation: Inferences of affiliation from third party observations. *Cognition, 170*, 31-48.

Pun, A., Birch, S. A., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences*, *113*(9), 2376-2381.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Roberts, S. B. G., Arrow, H., Lehmann, J., & Dunbar, R. I. M. (2014). Close social relationships: an evolutionary perspective. In R.I.M Dunbar, C. Gamble, & J.A.J. Gowlett (Eds.), *Lucy to Language: The Benchmark Papers* (pp. 151-180). Oxford, UK: Oxford University Press.

Russon, A. E., & Waite, B. E. (1991). Patterns of dominance and imitation in an infant peer group. *Ethology and Sociobiology*, *12*(1), 55-73.

Saxe, R. (2013). The new puzzle of theory of mind development. In M. Banaji & S. Gelman (Eds.), *Navigating the Social World: What infants, children, and other species teach us* (pp. 107-112). New York: Oxford University Press.

Schwarz, J. C. (1972). Effects of peer familiarity on the behavior of preschoolers in a novel situation. *Journal of Personality and Social Psychology, 24*(2), 276-284.

Selman, R. L. (1981). The child as a friendship philosopher. In S. R. Asher, & J. M. Gottman, (Eds.), *The development of children's friendships* (pp. 242-272). Cambridge, U.K.: Cambridge University Press.

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental Science*, 13(4), 599-610.

Shutts, K., Kinzler, K. D., & DeJesus, J. M. (2013). Understanding infants' and children's social learning about foods: Previous research and new prospects. *Developmental Psychology, 49*(3), 419-425.

Shutts, K., Kinzler, K. D., McKee, C. B., & Spelke, E. S. (2009). Social information guides infants' selection of foods. *Journal of Cognition and Development, 10*(1-2), 1-17.

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*, 14(1), 35-62.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587-592.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*(4), 605-632.

Spokes, A. C., & Spelke, E. S. (2016). Children's expectations and understanding of kinship as a social category. *Frontiers in Psychology*, *7*, 440.

Spokes, A. C., & Spelke, E. S. (2017). The cradle of social knowledge: Infants' reasoning about caregiving and affiliation. *Cognition*, *159*, 102-116.

Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. *British Journal of Psychology, 101*(2), 311-323.

Thomas, A. J., Abramyan, M., Lukowski, A., Thomsen, L., & Sarnecka, B. W. (2016). Preferring the Mighty to the Meek: Toddlers Prefer Novel Dominant Agents. In A. Papafragou, D. Grodner, D. Mirman, & J.C. Trueswell (Eds.) *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Thomsen, L., & Carey, S. (2013). Core cognition of relational models. In M. Banaji & S. Gelman (Eds.), *Navigating the Social World: What infants, children, and other species teach us* (pp. 17-22). New York: Oxford University Press.

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *Science*, *331*(6016), 477-480.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78*(3), 705-722.

van Baaren, R. B., Holland, R. W., Kawakami, K., & Van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science, 15*(1), 71-74.

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development*, *18*(1), 61-71.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301-1303.

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

Xu, J. (2017). *The Good Child: Moral Development in a Chinese Preschool*. Stanford, CA: Stanford University Press.

Youniss, J., & Volpe, J. (1978). A relational analysis of children's friendship. *New Directions for Child and Adolescent Development*, 1978(1), 1-22.

Appendix

Appendix A – Paper 1

Facts from Experiments 1-3 in Paper 1

N.B.: Each of the following facts includes a specific character's name. In actuality, the character each fact was used with differed between participants in line with counterbalancing.

Joint activity facts from Experiments 1 and 3

Helping

1. This boy and Jayden went camping last summer, and this boy helped Jayden put up a tent. They made it together.
2. Sophie was making a card for her Mom and this girl helped her. They made the card together.
3. Matthew's room was messy and this boy helped him clean it up. They worked together and got the room tidy.
4. Abigail's cat was missing and this girl helped her find it. They looked for it together and found it sleeping in the backyard.

Invitation

1. Maya invited this girl to the zoo. Last week, they went to the zoo together and they say all sorts of different animals.
2. Jake invited this boy to a sleepover at his house last month. This boy went to Jake's house and he took his sleeping bag with him.
3. Last week, Zoe invited this girl to the cinema. They went to the cinema and watched a movie together.

4. Logan invited this boy to his birthday party. This boy went to the birthday part at Logan's house and he took a present with him.

Arbitrary link facts from Experiments 1 and 2

Coincidental proximity

1. On Saturday, Jayden was sitting outside this café eating lunch. At the same time, this boy was in a taxi that went by.
2. On Sunday, Sophie was at the stadium watching a baseball game and this girl was in the plane that flew over the stadium.
3. A couple of weeks ago, this boy was riding in this train. Right as the train was passing, Matthew was playing with his kite nearby.
4. Yesterday, this girl was sitting in class at school and Abigail rode past the school's building on her bike.

Coincidental commonality

1. On this girl's birthday last year, it was snowing. It was also snowing on Maya's birthday this year.
2. This boy's birthday is in July and he went to the cinema on her birthday. Jake's birthday is in July too and he went swimming on his birthday.
3. Two years ago, this girl moved to a new house. Last year, Zoe moved to a different school.
4. It was cold when this boy and his class went to the zoo last month. It was also cold when Logan went shopping with his Mom yesterday.

Similarity facts from Experiments 2 and 3

Similarity in ability

1. Look at this drawing. Jayden is really good at drawing colorful pictures. This boy is also good at drawing pictures.
2. Sophie is really good at singing. She sang in her school play last year. This girl is good at singing too.
3. Matthew is very good at counting numbers. He can count all the way up to 100. This boy is good at counting too.
4. Abigail is really good at running. She can run pretty fast. This girl is also really good at running as well.

Similarity in experience

1. Last week, Maya's class went to the park and she saw the trees. This girl went to the park yesterday and saw the trees too.
2. Jake went to the aquarium with his Mom and saw the fish. A month ago, this boy went to the aquarium with his class and he saw the fish too.
3. Zoe went to the beach for the holidays last summer. This girl went to the beach for the holidays too.
4. Logan has a dog and he plays with him every day. This boy has a dog too.

Appendix B – Paper 1

Facts from Experiment 4 in Paper 1

N.B.: Each of the following facts includes a specific character's name. In actuality, which character the facts were used with differed between participants in line with counterbalancing.

Joint activity facts

Helping

1. This boy helped Jayden plant a flower. After planting it, they watered it together.
2. This girl helped Abigail make a card for her mom. They made it together.
3. This girl helped Zoe clean up the toys at their school. They cleaned it up together.
4. This girl helped Maya look for her cat. They found the cat in the yard together.

Invitation

1. This boy invited Logan to his birthday party. He came and brought a present.
2. This boy invited Jake to go to the playground. They went there together.
3. This boy invited Matthew to the zoo. They went there and saw the animals.
4. This girl invited Sophie to a sleepover at her house. They had fun.

Arbitrary link facts

Coincidental proximity

1. This girl was in this bus here. Maya was in the car that passed the bus.
2. This boy was riding on the train. His train went through Jayden's town.

3. This boy was in this shop. Jayden walked past the shop on his way home.
4. This boy went on holiday on a plane. His plane went over Matthew's house.

Coincidental commonality

1. This boy has a river in his town. There's a river in Logan's town too.
2. This girl got snow where she lives. It was snowing in Abigail's town too.
3. This girl's birthday is in July. Sophie's birthday is in July as well.
4. This girl's house is far from her school. Zoe's house is far from her school too.

Appendix C – Paper 1

Motivation for including replication of Experiment 1 in Experiment 4 in Paper 1

In order to ask about preferences, the task needed to involve a preferred item, which made for more complex trials, both visually and verbally. Additionally, given the repetitiveness of the procedure, the number of trials that each participant could reasonably complete was limited. Therefore, part of the motivation for replicating the finding from Experiment 1 was to examine whether children would make meaningful inferences about friendship when responding to fewer trials, each of which involved more visual and verbal information. If these more complicated trials precluded children from inferring friendship because they are unable to process the information, their responses to the preference trials may also be susceptible to the same low-level, attentional concerns. Thus, any failure to infer shared preferences may reflect issues with the task rather than children's true performance. If, however, subjects responded similarly to Experiment 1 when asked about friendship here, it would suggest that the current paradigm is robust enough to also explore preference inferences. This would mean that the pattern of responses for preferences would by extension be a more reliable signal of children's true abilities. Thus, if children failed to infer any shared preferences between characters linked by positive interactions, the failure would be interpretable.

Friendship Session Procedure. Four sets of characters and facts were used with eight items (four food and four game pictures). Characters and facts were presented twice, once with food items and once with game items. Each set of character pictures and facts was separated from its own repetition by three trials. Repetition of the characters and facts within the session held everything but the type of item and test question constant. The test question in the first four-trial block was who the central character's friend was, regardless of whether the block involved

food or game items. In the second four-trial block, subjects were asked to identify the central character's playmate, e.g. "I think Jayden plays with one of these boys a lot. Can you help me figure out which one he plays with?" Further, the order of games or foods in the session was also counterbalanced.

Preference session Procedure. Four sets of characters and facts were used together with eight items. No pictures or facts were the same as the friendship session. Both blocks in the preference session asked who would share the central character's preference for an item. Whether the items in the first block were games or foods was counterbalanced across subjects. As in the friendship session, the characters and facts were presented once again in the second block.

Design. Each participant viewed 16 trials, divided into two 8-trial sessions: a friendship session and a preference session. Order of sessions was counterbalanced across participants. Each session included a four-trial block of food items and a four-trial block of game items. Whether games appeared in the first or second block of each session was counterbalanced across participants. Side of target on screen, order of target fact presentation, and trial gender were counterbalanced within and across sessions and participants. Helping and invitation facts (matched against proximity and commonality facts the same number of times) were interspersed across trials with the constraint that two of each fact type appeared in each block (and therefore session). Finally, order of each triad of child pictures, as well as the order of fact pairs was counterbalanced across children.