Psychological and Sociomoral Reasoning in Infancy

Authors:
Renée Baillargeon[1][*], Rose M. Scott[2], Zijing He[3], Stephanie Sloane[1], Peipei Setoh[1], Kyong-sun Jin[1], Di Wu[4], and Lin Bian[1]

Affiliations:

[1]Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA.

[2]School of Social Sciences, Humanities, and Arts, University of California Merced, 5200 North Lake Road, Merced, CA 95343, USA.

[3]Department of Psychology, Sun Yat-sen University, 135 Xin Gang Xi Road, Guangzhou, Guangdong 510275, China.

[4]Department of Psychology, Cedarville University, 251 N. Main St., Cedarville, OH 45314, USA.

*Correspondence to: rbaillar@illinois.edu.

Introduction

Consider the following scene: A man wearing a backpack is pacing leisurely back and forth in a large airport room. As he strolls, he occasionally crosses his arms, twirls the dangling straps of his backpack, or stuffs his hands in his pant pockets. At one point, he sits down, takes off his backpack, and removes from it a bag filled with assorted gummy bears; as he peers intently inside the bag, he selects and eats, one at a time, five red gummy bears.

As adults, we would have no difficulty interpreting the man's actions. We might view his initial actions—pacing, crossing his arms, and so on—as intentional (as opposed to accidental) though not as directed toward any particular goal. In contrast, we might perceive his subsequent actions—removing the bag of gummy bears from his backpack and eating five red ones in succession—as both intentional and goal-directed. In analyzing these actions, we might build an explanation that attributes to the man a causally coherent set of motivational and epistemic mental states: He wants to eat gummy bears, he is particularly fond of red ones, and so when he spies one in the bag, he removes it and eats it.

Next, consider a new scene: As the man is happily chewing on red gummy bears, he notices a second man approaching who is carrying two heavy suitcases. At this point, our scene might unfold according to different scenarios. In one, the first man greets the second man, offers to carry one of the suitcases, and holds out his bag of gummy bears. In another scenario, the first man continues to watch the second man but makes no move to approach him. In yet another scenario, the first man sticks out a leg to trip the second man, causing him to fall heavily.

As adults, we would interpret and evaluate the first man's actions in the three scenarios very differently. In the first scenario, we might infer that the two men have a social group in common: They might be friends, coworkers, or relatives, for example. In the second scenario, we

would conclude that the two men are strangers. In both the first and second scenarios, we would view the first man's behavior as acceptable: Offering help and sharing food are expected prosocial behaviors in interactions with ingroup, but not outgroup, individuals. In contrast, the first man's behavior in the third scenario would seem to us beyond the pale: Unprovoked harmful actions, even against outgroup individuals, are generally viewed as unacceptable. We would categorize the first man as an antisocial lout, and we might file away distinctive characteristics as possible markers of a social group to be avoided in the future.

Our discussion of the two scenes above illustrates the rich analyses that adults spontaneously engage in when watching others act. What are the developmental origins of these interpretations? Over the past 25 years, there has been a great deal of research on social cognition in infancy. This research can be roughly organized into two sets of questions that map neatly onto the two scenes above. First, when watching an agent act in a scene, how do infants represent and interpret the agent's actions? Second, in scenes with two or more agents, what expectations do infants possess about how the agents should interact? How do infants represent, interpret, and evaluate the successive actions that take place in the course of an unfolding interaction? In this chapter, we refer to the first set of questions as *psychological reasoning*[1] and to the second set as *sociomoral reasoning*.

Not surprisingly, developmental researchers began by focusing their attention on early psychological reasoning and, as it became apparent that even young infants possess sophisticated psychological expectations, gradually turned their attention to early sociomoral reasoning. The organization of this chapter reflects this historical perspective: The first and longer section reviews some of the key findings on infant psychological reasoning, whereas the second section summarizes new findings on infant sociomoral reasoning.

PART 1: EARLY PSYCHOLOGICAL REASONING

1-I. CAN YOUNG INFANTS INTERPRET OTHERS' ACTIONS?

A. A Methodological Revolution

At the start of the 1980s, a revolution took place in infancy research as investigators began to adapt available looking-time methods to explore young infants' *physical world* (e.g., Baillargeon, Spelke, & Wasserman, 1985; Kellman & Spelke, 1983; Leslie, 1984; for a review, see Baillargeon, Li, Gertner, & Wu, 2011). Until then, early physical reasoning had typically been assessed through *action tasks* (i.e., tasks that required infants to produce overt manual actions). Unfortunately, negative results with young infants were often difficult to interpret: Did young infants truly lack the physical expectations under investigation, as was generally assumed (e.g., Piaget, 1954), or were they merely unable—due to ancillary difficulties—to produce the actions necessary for success (e.g., Bower, 1977)? By using looking-time methods, researchers could begin to disentangle cognitive from response limitations in young infants.

One looking-time method that proved particularly helpful in exploring early physical reasoning was the *violation-of-expectation* (VOE) method. This method relies on infants' robust tendency to look longer at events that violate, as opposed to confirm, their expectations. In daily life, as infants encounter unexpected events (i.e., events inconsistent with their current state of knowledge), they naturally scrutinize the events to revise their faulty expectations and bring them in line with their observations; the VOE method takes advantage of this natural tendency in order to uncover what expectations, if any, infants possess about specific events (e.g., Luo & Baillargeon, 2005b; Wang, Baillargeon, & Brueckner, 2004).

During the 1980s and early 1990s, many new VOE tasks were designed to explore various facets of infants' physical world; studies of object permanence, object support, and

object tracking, among others, revealed that, under optimal conditions, even young infants demonstrated sophisticated expectations about physical events (e.g., Baillargeon, & DeVos, 1991; Needham & Baillargeon, 1993; Spelke, Breinlinger, Macomber, & Jacobson, 1992; Wynn, 1992). These new findings (though initially quite controversial) ushered in new perspectives on the cognitive competences of young infants and led to new proposals on the mechanisms responsible for their attainment.

One such proposal, consistent with the Rationalist view that humans possess domain-specific cognitive systems (e.g., Chomsky, 1965), is that infants are born with a *physical-reasoning* system, an abstract computational system that provides them with a skeletal causal framework for reasoning and learning about physical events (e.g., Baillargeon, 2008; Gelman, 1990; Leslie, 1994b; Spelke, 1994). According to this proposal, the physical-reasoning system's initial causal framework includes several core concepts and principles. For example, the principle of *persistence* states that, all other things being equal, objects persist, as they are, in time and space (e.g., Baillargeon, 2008). Beginning with the seminal work of Spelke and her colleagues (e.g., Spelke, Phillips, & Woodward, 1995), a great deal of research has examined infants' sensitivity to the principle's corollaries: An object cannot spontaneously appear or disappear (*continuity*), occupy the same space as another object (*solidity*), break apart (*cohesion*), fuse with another object (*boundedness*), or change its appearance (*unchangeableness*) (e.g., Baillargeon, Li, Ng, & Yuan, 2009a; Cheries, Wynn, & Scholl, 2006; Needham, 1998; Newcombe, Huttenlocher, & Learmonth, 1999; Saxe, Tzelnic, & Carey, 2006; Wilcox, 1999).

In the mid-1990s, encouraged by the positive findings from VOE investigations of infants' physical world, researchers began to use VOE tasks to explore infants' *psychological world* as well. Until then, investigations of early psychological reasoning had relied primarily on

infants' overt actions in social contexts—for example, in gaze-following, point-following, social-referencing, imitative-learning, or language-comprehension tasks. Although many positive results were obtained with infants in the second year of life (for reviews, see Baldwin & Moses, 2001; Carpenter, Nagell, & Tomasello, 1998; Meltzoff, 2002), negative results with younger infants were, here again, difficult to interpret. Did young infants truly lack any psychological understanding of others as intentional agents, as some researchers claimed (e.g., Barresi & Moore, 1996; Butterworth & Jarrett, 1991; Tomasello, 1999)? Or did exclusive reliance on overt social responses tend to underestimate young infants' psychological understanding, as other investigators suggested (e.g., Leslie, 1994b; Premack, 1990; Trevarthen, 1998)? VOE tasks provided a way to address this issue and revealed rich psychological-reasoning competences in young infants, as explained in the next sections.

B. The Principle of Rationality

As adults, our expectations about others' actions are guided by a principle of *rationality*: All other things being equal, we expect agents to act rationally—indeed, this is what makes it possible for us to interpret and predict their actions (e.g., Dennett, 1987; Fodor, 1987). Like the principle of persistence discussed in the last section, the principle of rationality has multiple corollaries. One such corollary is *consistency*: Agents' actions should be consistent with their mental states (e.g., their goals, dispositions, knowledge, and so on). Thus, we would expect a woman who was preparing to go jogging to be putting on sneakers rather than ski boots, and we would be surprised if a friend who hated slasher films arranged an eveningfest of "Halloween" movies. Another corollary is *efficiency*: We expect agents to expend as little effort as possible to achieve their goals (e.g., Gergely, Nádasdy, Csibra, & Bíró, 1995; Gergely & Csibra, 1997). Thus, we would be puzzled if a man walked around the edges of a large empty parking lot to get

to his car across the lot, or if a customer dug through a pile of identical packages of printer paper to get the one at the bottom, instead of simply taking the one at the top. Yet another corollary is *normality*: We expect agents to act in ways we take to be normal. Thus, we would be alarmed if a friend announced that he planned to serve blood pudding made with his own blood at a dinner party, or if a neighbor told us she was considering having all her teeth removed to help her stay on her diet. In general, when observing an agent act in a scene, we attempt to build a *psychological explanation* that specifies the mental states that underlie these actions. A psychological explanation is well-formed if it posits mental states (1) that are possible, given what is known about the agent and the scene (e.g., it does not attribute knowledge that the agent could not possibly have), (2) that are causally coherent, and (3) that portray the agent's ongoing actions as rational—consistent, efficient, and normal.

Initial VOE investigations of the young infant's psychological world all tended to focus, implicitly or explicitly, on the rationality principle. Two sets of experiments, in particular, had a profound impact on theory and research on early psychological reasoning. One set, which explored infants' sensitivity to *consistency* constraints, was conducted by Amanda Woodward and her colleagues (e.g., Woodward, 1998, 1999; Woodward & Sommerville, 2000). The other set, which focused mainly on infants' sensitivity to *efficiency* constraints, was conducted by Gergely Csibra, György Gergely, and their colleagues (e.g., Csibra, Gergely, Bíró, Koós, & Brockbank, 1999; Gergely, Bekkering, & Király, 2002; Gergely et al., 1995). In the next sections, we discuss each set of experiments in turn. (fMRI findings with adults also support the distinction between consistency and efficiency constraints: For example, there is evidence that the right temporo-parietal junction is preferentially recruited when participants detect consistency violations (e.g., Cloutier, Gabrieli, O'Young, & Ambady, 2011; Saxe & Wexler,

2005), whereas the superior temporal sulcus is preferentially recruited when participants detect efficiency violations (e.g., Brass, Schmitt, Spengler, & Gergely, 2007; Grèzes, Frith, & Passingham, 2004)).

C. Are Infants Sensitive to Consistency Constraints?

1. Do infants attribute preferences to agents and expect them to act in a manner consistent with these preferences?

In a seminal series of experiments, Woodward (1998, 1999) devised a novel *two-object* VOE task to examine whether young infants (1) could attribute to a human agent the goal of obtaining a specific object and, if yes, (2) would expect the agent to maintain this goal and to adjust her actions, when the object's position changed, in order to again obtain the object, in accordance with the consistency principle. In the experiments, 5- to 9-month-olds first received habituation trials in which they saw the following event: Two toys, a ball and a teddy bear, rested on an apparatus floor; a woman's bare hand reached through a curtained opening in the right wall of the apparatus, grasped one of the toys (e.g., the teddy bear; henceforth the target toy), and paused until the infant looked away and the trial ended. In each habituation trial, the hand grasped the target toy, and trials were repeated until looking time decreased to a preset criterion (infants received 6 to 14 habituation trials). Following habituation, the toys' positions were switched, and the infants received a static display trial showing the toys in their new positions. Finally, the infants received test trials in which the hand grasped either the target toy (*old-object* event) or the other toy (*new-object* event). At all ages, infants looked reliably longer at the new- than at the old-object event, suggesting that (1) they recognized the hand as a part of a human agent; (2) during the habituation trials, they interpreted the hand's actions as directed at the goal of obtaining the target toy; (3) during the test trials, they expected the agent to maintain this goal

and to reach for the target toy in its new position, in accordance with the consistency principle; and hence (4) they detected a violation when she reached for the other toy instead.

Woodward's (1998, 1999) results were subsequently confirmed in multiple laboratories (e.g., Luo & Baillargeon, 2005a; Martin, Onishi, & Vouloumanos, 2012; Shimizu & Johnson, 2004; Spaepen & Spelke, 2007; Schlottmann, Ray, & Surian, 2012). As time went on, however, researchers came to the realization that a richer interpretation of the results was called for. Specifically, it appeared that during the habituation trials, (1) the infants understood that the agent's actions were directed at the goal of obtaining the target toy; (2) they noticed that the agent always chose the target toy over the other toy (i.e., she never attempted to obtain the other toy); and (3) on the basis of this unvarying choice information, they attributed to the agent a positive disposition toward, or *preference* for, the target toy.[2] From this perspective, the infants thus looked reliably longer at the new-object test event because the agent, after demonstrating a preference for the target toy, now sought the other toy, in violation of the consistency principle. (Adults, of course, also use choice information to infer preferences: We would attribute to Uncle Bernie a fondness for cheesecake if he regularly selected cheesecake at dessert buffets).

Part of the evidence for this preference-based interpretation came from experiments that contrasted Woodward's two-object task with a novel *one-object* task in which only the target toy was present in the habituation trials; the other toy was not added until the display and test trials; e.g., Bíró, Verschoor, & Coenen, 2011; Luo & Baillargeon, 2005a; Robson & Kuhlmeier, 2013; Song, Baillargeon, & Fisher, in press). Infants in these one-object tasks looked about equally at the new- and old-object test events: Although they still understood in the habituation trials that the agent's actions were directed at the goal of obtaining the target toy, they no longer had choice information they could use to infer her disposition toward the toy. Did she have a *neutral*

disposition toward it and repeatedly reach for it simply because it was the only toy available? Or did she have a *positive* disposition toward it and repeatedly reach for it because she liked it? Because infants could not determine the agent's disposition toward the target toy, they could not predict whether she would reach for it or for the other toy when both were available, and so they tended to look equally when she reached for either toy. These negative results made clear that infants in two-object tasks do not simply attribute to the agent the goal of obtaining the target toy in the habituation trials and expect her to maintain this goal in the test trials (had this been the case, responses in one- and two-object tasks would have been the same). Rather, infants in two-object tasks attribute to the agent a preference for the target toy and expect her to continue acting on this preference in the test trials, in accordance with the consistency principle.[3]

Woodward's (1998, 1999) two-object task has proven enormously productive: As will become clear, it has been used with a myriad of variations to explore different facets of infants' psychological reasoning. In the next sections, we summarize some of the key results that have been obtained concerning infants' ability to represent and to reason about agents' preferences.

2. What preferences do infants attribute to agents?

When infants see an agent repeatedly choose a teddy bear over a ball in a two-object task, what sort of preference do they attribute to the agent: a preference for that teddy bear in particular or a preference for teddy bears in general? Experiments by Spaepen and Spelke (2007) with 12-month-olds suggest that the second alternative is correct, at least by the end of the first year. If in the habituation trials the agent continually chose a toy from one taxonomic category (e.g., a red tow truck) over a toy from the *same* taxonomic category (e.g., an orange dump truck), then in the test trials infants tended to look equally at the new- and old-object events. In contrast, if in the habituation trials the agent continually chose a toy from one taxonomic category (e.g.,

the red tow truck) over a toy from a *different* taxonomic category (e.g., a white male doll), then infants looked reliably longer at the new- than at the old-object event; moreover, infants did so whether the *same* exemplars (e.g., the red tow truck and the white male doll) or *new* exemplars (e.g., the orange dump truck and a black female doll) were used in the test trials. Together, these results suggest that when the agent reached repeatedly for a truck as opposed to a doll, infants attributed to her a general preference for trucks and expected her to act on this preference even when new exemplars were introduced in the test trials. (When the agent repeatedly selected one truck as opposed to a different truck, infants did not attribute to the agent a specific preference for the selected truck. Did they attribute to the agent the goal of obtaining a truck, but no preference for trucks in general? Additional research is needed to determine whether contrastive categories are necessary for infants to attribute preferences—or, to put it another way, whether one-category tasks are equivalent for infants to one-object tasks, in that they provide no information about underlying dispositions).[4]

The preceding results should not be taken to suggest that infants can attribute only taxonomic preferences (e.g., for trucks) to agents: When provided with appropriate evidence, infants are clearly capable of attributing other types of preferences. For example, after watching an agent choose a red toy pepper over a black cup, and a red pyramid over a yellow toy house, 16-month-olds attributed to the agent a preference for a *property-based* category, red objects (Luo & Beck, 2010). Infants are also sensitive to *functional* differences between objects: When an agent faced two toy penguins that were identical except that one could be disassembled and one could not, 18-month-olds attributed to the agent a specific preference for the penguin that could be disassembled if she used it to hide a key inside it, but not otherwise (Scott & Baillargeon, 2009). Together, these results suggest that, although in the second year of life the

default encoding of object preferences may be taxonomic, infants can easily be led to alternative encodings by exposure to appropriate observations.

Although infants do not attribute a specific preference to an agent who repeatedly selects one of two different objects from the same taxonomic category (e.g., a red tow truck as opposed to an orange dump truck), they are still sensitive to the agent's attentional and emotional responses and can use these to predict which of the two objects the agent is likely to act on next. In a novel task, Phillips, Wellman, and Spelke (2002) habituated 12-month-olds to the following event: A female agent sat centered behind two toy cats that differed only in color (one was grey and the other orange). To start, the agent looked at one of the cats (cat-A) with a positive facial expression while saying "Oo-oo, look at the kitty" in a pleasant voice. Next, a curtain was closed to hide the scene, and when it reopened the agent held cat-A. The test events followed the same structure; in the matching event, the agent looked at cat-B and then held cat-B, and in the non-matching event she looked at cat-A, but then held cat-B. The infants looked reliably longer at the non-matching than at the matching event, suggesting that they attended to the agent's gaze and emotional expressions and used these to predict which cat she would pick up next. According to the preceding analysis, the design of Phillips et al. was effective only because the two toys used belonged to the *same* taxonomic category: Because the infants did not attribute to the agent a specific preference for cat-A, they did not detect a consistency violation when she switched her attention from cat-A to cat-B. And indeed, when Sodian and Thoermer (2004) attempted to replicate the experiment of Phillips et al. using toys from different taxonomic categories (a ball and a tumbling figure), they obtained negative results. After seeing the agent repeatedly hold object-A as opposed to object-B in the habituation trials, infants were puzzled when she held object-B in the matching (look at object-B, hold object-B) and non-matching (look at object-A,

hold object-B) events. Both events suggested that the agent had changed her preference, for no apparent reason, from object-A to object-B, thus violating the consistency principle.

3. Are preferences agent-specific?

When infants attribute a preference to an agent, do they view this preference as specific to the agent and recognize that a different agent might not have the same preference? Or do they expect all agents to share the same preference? At issue is whether infants recognize that preferences are often attributes of individual agents: Mommy prefers wine, but daddy prefers beer; big sister Jane is fond of reading, but big brother Karl likes sports. To examine this question, Buresh and Woodward (2007) conducted two-object experiments with 9- and 13-month-olds using toys from different taxonomic categories (a toy rocket and a toy animal). The *same-agent* condition was similar to previous two-object experiments, with agent-A acting in both the habituation and the test events. In the *switch-agent* condition, agent-A (e.g., a man) acted in the habituation event, but agent-B (e.g., a woman) acted in the test events (the infants were briefly familiarized with both agents prior to the testing session). In line with prior findings, the infants in the same-agent condition looked reliably longer at the new- than at the old-object test event. In contrast, the infants in the switch-agent condition looked about equally at the two events, suggesting that they had no expectation about which toy agent-B would select: The fact that agent-A had a preference for object-A did not mean that agent-B would share this preference (see also Henderson & Woodward, 2012, for similar results with 9-month-olds).

Although infants tend to view preferences as agent-specific, there are at least two situations where they deviate from this pattern and expect preferences to be shared more generally. First, when a novel animal prefers one food over another, 15-month-olds expect similar animals (as indexed by shape) to share this dietary preference (Setoh & Baillargeon,

2010); in this case, infants' inductive inferences appear to be guided by their conceptual knowledge about biological categories (e.g., Gelman, 2003; Gelman & Markman, 1986). Second, 18-month-olds extend a preference differently depending on its communicative context (Egyed, Király, & Gergely, in press). When agent-A used ostensive-communicative cues (e.g., looked at and addressed the infant) before emoting positively toward object-A and emoting negatively toward object-B, 18-month-olds inferred that object-A was pleasing whereas object-B was not, and they expected agent-B to share the same knowledge and preference. If agent-A did not use ostensive-communicative cues, however, the infants took her actions to reflect only her own dispositions toward the objects, and they had no expectation about which object agent-B would prefer. Infants thus seem to interpret object-directed emotional expressions that are accompanied by ostensive signals as *pedagogical* encounters aimed at "teaching" them the properties of the objects, and they expect other agents to possess the same knowledge about the objects (e.g., Csibra & Gergely, 2009; Gergely, Egyed, & Király, 2007).

Finally, what happens when an agent expresses a preference that differs from the infant's own preference? In a well-known experiment, Repacholi and Gopnik (1997) found that, when an agent expressed a preference for broccoli over goldfish crackers and then requested food, 18-month-olds correctly gave her broccoli, even though they themselves preferred the crackers. Younger, 14-month-old infants tended to give the agent crackers, but it is unclear whether they were acting egocentrically (i.e., had difficulty attributing a preference distinct from their own) or whether they were interpreting her minimal request ("Can you give me some?") more loosely, due to linguistic or pragmatic limitations. Additional research is needed to resolve this issue.

4. What evidence leads infants to attribute preferences? The case of choice information

Older infants. By their first birthday, infants attribute to an agent a preference for object-

A over object-B not only when she repeatedly *grasps* object-A (e.g., Guajardo & Woodward, 2004; Spaepen & Spelke, 2007), as we saw before, but also when she repeatedly *points* to object-A or *looks* at it intently (e.g., Meltzoff & Brooks, 2001; Woodward, 2003; Woodward & Guajardo, 2002). In each case, infants (1) view the agent's actions on object-A as goal-directed, (2) notice that she continually chooses to act on object-A as opposed to object-B, and (3) infer from this unvarying choice that she has a preference for object-A.

Infants in the second year of life also interpret goal-directed actions on identical objects, when statistically improbable, as signaling preferences (just as we inferred, in the first scene presented in the Introduction, that the man had a particular fondness for red gummy bears; e.g., Gweon, Tennenbaum, & Schulz, 2010; Kushnir, Xu, & Wellman, 2010). In one experiment, for example, 19- to 24-month-olds were shown a box that contained different ratios (7:31 or 31:7) of toy ducks to toy frogs (Kushnir et al., 2010). The infants first watched a female agent remove five identical toys (e.g., all ducks) from the box, and then they were asked to select a toy for her from two bowls, one containing ducks and one containing frogs. The infants were reliably more likely to select the same toy as the agent when the box contained few of the toys she chose (e.g., when she removed five ducks from a box of mostly frogs), suggesting that they took her deviation from random sampling to reveal a preference (e.g., a fondness for ducks). Gweon et al. (2011) obtained related results with younger, 15-month-old infants.[5]

Younger infants. During the first year of life, young infants fail at many of the two-object tasks at which older infants succeed. The problem usually stems from the fact that young infants cannot make sense of the action the agent is directing at object-A (i.e., they cannot build a satisfactory explanation for this action), and so they cannot proceed further in their reasoning. At least two factors seem to contribute to young infants' difficulties.

First, young infants sometimes lack sufficient experience to identify the goal of the agent's action on object-A. We can easily imagine, for example, that young infants may be nonplussed when they first observe a parent listen to a cell phone, point a remote key at a car, or lick a fingertip before turning a page; in such cases, infants may appreciate that the parent is acting intentionally, but be uncertain as to *why* the parent is performing these actions. In line with this first factor, researchers have found that 7- to 8-month-olds, who typically have not yet learned the communicative function of pointing gestures (e.g., Carpenter, Nagell, et al., 1998; Morissette, Ricard, & Decarie, 1995), fail at two-object tasks involving pointing (e.g., Kim & Song, 2008; Woodward & Guajardo, 2002). At 7 months, infants fail at two-object tasks involving grasping if the agent wears a glove (even if they are introduced to the gloved agent prior to the testing session), most likely because they are unsure what function the glove is serving (Guajardo & Woodward, 2004). At 3 months, infants fail at two-object tasks involving grasping even if the agent's hand is bare, presumably because the scene presented—a hand grasping a toy and then pausing—is too minimalist to be meaningful at this young age; infants do succeed, however, if (1) during an initial play session, they wear Velcro mittens which allow them to pick up small toys by swiping at them, and (2) during the testing session, the agent's hand wears a similar mitten (Sommerville, Woodward, & Needham, 2005). Infants' play experience with the Velcro mittens no doubt provides them with a supportive context for interpreting the hand's scant actions. Supportive contexts can also help young infants make sense of odd, unfamiliar actions: For example, although 8- to 10-month-olds fail at two-object tasks in which the agent, instead of grasping the target toy, "flops" the back of her hand against it (Woodward, 1999), they succeed if she then uses the back of her hand to push the toy, suggesting a possible goal for this otherwise baffling action (Király, Jovanovic, Prinz, Aschersleben, &

Gergely, 2003).

Second, young infants' limited experience may lead them to view as non-rational some actions that older infants—from their more experienced and more flexible perspective—may deem explicable and rational. For example, consider the result that 7- and 9-month-olds fail at two-object tasks involving looking (Woodward, 2003). One possible interpretation of this negative result is that young infants take the agent's actions to violate the consistency principle: If the agent wants object-A, as her attentional actions suggest, why does she not take it, since it is well within her reach and there is nothing to prevent her from taking it? In line with this interpretation, Luo (2010) found that 8-month-olds succeeded at a two-object task involving looking when the scene provided an explanation for the agent's failure to reach for object-A: Either the agent sat behind a small window that only allowed her to look at object-A, or she sat at a large window but her hands were occupied holding the two handles of a sippy cup (this manipulation was adapted from Gergely et al., 2002, described later). These results suggest that, whereas 8-month-olds are in the habit of reaching for interesting objects within easy reach and interpret others' actions accordingly, 12-month-olds have learned (perhaps via parental admonitions) that one may sometimes look at, but not touch, interesting objects.

Amount of choice information. How much choice information do infants require to attribute preferences? Although many investigators using two-object tasks uphold the tradition of administering 6-14 habituation trials, other investigators have begun to use a smaller, fixed number of trials, referred to as *familiarization* trials. Positive results have been obtained with young infants with as few as four familiarization trials (e.g., Luo & Baillargeon, 2005a; Luo & Johnson, 2009; Luo, Markson, Hennefield, Mou, & vanMarle, 2013; Song et al., in press). For example, 8-month-olds attributed a preference to an agent who grasped object-A as opposed to

object-B in four familiarization trials; infants did not do so, however, if in either the first or the last familiarization trial the agent grasped object-B instead (Luo et al., 2013). Thus, although consistent choices across a few trials may be sufficient for infants to posit a preference, slight inconsistencies under these conditions will prevent such attributions.

5. *What evidence leads infants to attribute preferences? The case of effort information*

Until now, we have focused on tasks where choice information could be used to infer preferences: The agent repeatedly chose a teddy bear over a ball, chose only ducks from a box containing mainly frogs, and so on. Recent research suggests that, besides choice information, infants also use *effort* information to infer preferences. (Adults, of course, do the same: We would infer that Aunt Bertha was very fond of cherries if she dragged a ladder to her cherry tree to eat the last few cherries in the top branches).

To date, the evidence that infants can use effort information to attribute preferences comes primarily from modified one-object tasks (e.g., Bíró et al., 2011; Hernik & Southgate, 2012). For example, Bíró et al. (2011) tested 12-month-olds in two different conditions using videotaped events. In the *less-effort* condition, only object-A (e.g., a teddy bear) was present in the familiarization trials; in each trial, a woman's bare hand reached for object-A, grasped it, and paused. In the test trials, object-A was moved to a new position and object-B (e.g., a block) was placed in the position formerly occupied by object-A; the agent reached either for object-A (old-object event) or object-B (new-object event). The *more-effort* condition was similar except that in the familiarization trials object-A stood inside a closed transparent box; the hand first removed the lid of the box and then grasped object-A. The infants in the less-effort condition tended to look equally at the old- and new-object test events, replicating previous results with similar one-object tasks (e.g., Kim & Song, 2008; Luo & Baillargeon, 2005a). In contrast, the infants in the

more-effort condition looked reliably longer at the new- than at the old-object event. Together, these results suggested that when minimal effort is needed to obtain the only object present in a scene, infants have no basis for determining whether the agent has a neutral or a positive disposition toward the object, and so they detect no violation when she reaches for a different object in the test trials. However, if the agent repeatedly goes to some effort (e.g., produces a means-end sequence of actions) to obtain the object, infants infer that she has a preference for it, and they expect her to continue reaching for it even when a novel object is introduced.

Could the infants in the more-effort condition have used subtle choice information, rather than effort information, to attribute to the agent a preference for object-A? After all, each familiarization trial ended with the agent grasping object-A as opposed to the box (from which object-A was removed). Prior experiments by Woodward and her colleagues suggest that this explanation is unlikely: When an agent performs a means-end sequence of actions to obtain an object, 12-month-olds view the separate actions in the sequence as directed at the overarching goal of obtaining the object (e.g., Sommerville & Woodward, 2005; Woodward & Sommerville, 2000). In one experiment, for example, object-A and object-B (a toy bear and a toy tiger) were placed inside two tinted clear boxes, one blue and one purple (Woodward & Sommerville, 2000). In the habituation event, the agent opened the lid of the box containing object-A, grasped object-A, and paused. In the test events, the toys' (but not the boxes') positions were switched; the agent grasped either the lid of the box containing object-A (old-object event) or the lid of the box containing object-B (new-object event) and then paused, without opening the box. The infants looked reliably longer at the new- that at the old-object event, suggesting that (1) they understood the agent's actions during the habituation trials as a means-end action sequence designed to obtain object-A, (2) they attributed to the agent a preference for object-A (because

she consistently chose it over object-B and/or because she went to some effort to retrieve it), and (3) they expected the agent to act on the other box in the test trials so as to again obtain object-A.

In line with the rationality principle, infants do not view just *any* action sequence that ends with grasping object-A as evidence of a preference for object-A. Thus, Bíró et al. (2011) found that infants looked about equally at the new- and old-object test events in the more-effort condition if in the familiarization trials object-A stood *in front* of the box, so that opening the box was no longer necessary to obtain object-A. Similarly, Woodward and Sommerville (2000) found that infants looked about equally at the old- and new-object events if in the habituation trials object-A and object-B stood *in front* of their boxes. In both experiments, infants were unable to generate a satisfactory explanation for the action sequence repeatedly performed by the agent because it violated the efficiency principle (i.e., why did she always open the box first, since this action did not make it possible or easier to grasp object-A?). Lacking an explanation for the agent's inefficient actions, infants held no expectation about her actions in the test trials and hence they looked about equally at the two test events (we return later to infants' sensitivity to the efficiency principle).

6. What evidence leads infants to attribute preferences? The case of equifinality information

In addition to choice and effort information, infants can use *equifinality* information to infer preferences: If in a continually changing scene an agent keeps adjusting her actions so as to always achieve the same end, infants may attribute to the agent a positive disposition involving that end. Prior research suggests that infants can use equifinality information to attribute preferences for activities as well as for objects.

Preferences for activities. If we saw an infant gleefully throw whatever toy was placed within his reach, using one or two hands as needed, we might attribute to the young scamp a

preference or predilection for throwing objects. In a similar vein, Song and her colleagues examined whether infants could attribute to an agent a predilection for sliding objects (Song & Baillargeon, 2007; Song, Baillargeon, & Fisher, 2005). In one experiment, 9-month-olds first received six familiarization trials in which they saw a female agent slide an object forward and backward on an apparatus floor; six different objects (e.g., a toy fish, a box, a baby shoe, and so on) were used across trials (Song & Baillargeon, 2007). Next, the infants received a static display trial in which they saw two identical trucks placed side by side: One stood inside a short frame that left little room for sliding, and the other stood inside a longer frame that left ample room for sliding. Finally, the infants received test trials in which the agent grasped the truck inside either the short (short-frame event) or the long (long-frame event) frame and then paused.

The infants looked reliably longer at the short- than at the long-frame event, suggesting that (1) they noticed that the agent performed the same activity—sliding—regardless of which object was placed before her; (2) based on this equifinality information, they attributed to the agent a predilection for sliding objects; and hence (3) they expected the agent to grasp the "slidable" truck. Support for this interpretation came from a control condition in which the agent lifted (instead of slid) each object in the familiarization trials; as expected, the infants in this condition looked about equally at the short- and long-frame events, as the agent could lift either truck. Additional results indicated that when the agent slid three (instead of six) different objects in the familiarization trials, positive results were obtained at 13 months (Song et al., 2005), but not at 9 months (Song & Baillargeon, 2007). Moreover, 9-month-olds failed even when the three familiarization trials (with the three different objects) were repeated a second time for a total of six familiarization trials. Infants thus required more equifinality evidence at 9 months than at 13 months to attribute to the agent a preference for sliding objects.

Preferences for objects. Results by Luo (2011b) suggest that infants as young as 3 months are sensitive to equifinality. The infants were assigned to one of two one-object conditions: In the *no-variation* condition, object-A stood in the same location in all of the familiarization trials; in the *variation* condition, in contrast, the location of object-A varied randomly (on the left or right side of the apparatus) across trials. The infants in the variation condition looked reliably longer at the new- than at the old-object test event, whereas those in the no-variation condition looked about equally at the events (as in previous one-object tasks). Together, these results suggested that the infants in the variation condition used the equifinality information in the familiarization trials to attribute to the agent a preference for object-A.

Hernik and Southgate (2012) found that 9-month-olds succeeded at a one-object task that combined effort and equifinality information. In each familiarization trial, the agent had to detour around an obstacle to contact object-A (effort information), and the obstacle varied in width across trials so that the agent had to constantly change paths (equifinality information). Here again, only rational actions led to the attribution of a preference: Infants looked equally at the new- and old-object test events if in the familiarization trials the agent's detour around the obstacle was inefficient.

Mental effort? Although we have interpreted the findings presented in this section as evidence that infants are sensitive to equifinality information, an alternative interpretation might be that infants consider not only information about *physical* effort, as we saw earlier (e.g., Bíró et al., 2011), but also information about *mental* effort. In an equifinal situation, the agent must constantly monitor the scene, detect relevant changes, and adjust her actions so as to attain the same end: obtain the target toy, perform the target activity, and so on. Infants might perceive all of this mental effort as evidence that the agent is really invested in—or has a positive disposition

toward—the target toy or activity. We return later on to evidence that infants consider both physical and mental effort in interpreting and predicting agents' actions.

7. To what agents can infants attribute preferences?

To examine whether infants could attribute preferences to non-human agents as well as to human agents, Woodward (1998) tested 5- to 9-month-olds in a two-object task identical to her original task except that the human hand was replaced with a non-human entity: a rod tipped with a sponge, an arm-shaped screen, or a mechanical claw. In each habituation trial, the non-human entity entered the apparatus through a window in the right wall (the right end of the object never came into view) and contacted the target toy. Results were negative, suggesting that infants could not attribute goals and preferences to the rod, screen, or claw. These results, together with other results described in the next section, led Woodward (2005) to conclude that young infants can initially reason only about human agents and gradually extend their action understanding to non-human agents. As we will see in section 1-II, these findings and conclusions played a central role in the next wave of research in early psychological reasoning.

D. Are Infants Sensitive to Efficiency Constraints?

In addition to consistency, another major corollary of the rationality principle is efficiency: An agent who is pursuing a goal should do so efficiently, expending as little effort as possible. Gergely, Csibra, and their colleagues were the first to explore infants' sensitivity to efficiency constraints (e.g., Csibra et al., 1999; Gergely et al. 1995).

1. Do infants attribute goals to agents and expect them to pursue these goals efficiently?

In their first experiment, Gergely et al. (1995) devised a novel *detour* task that addressed three questions: (1) could infants view a non-human entity as an agent?; (2) could infants attribute to an agent the goal of reaching a specific target?; and (3) would infants expect an agent

to use the shortest possible path to reach a target, in accordance with the efficiency principle? In the experiment, 12-month-olds were first habituated to the following computer-animated event. First, a large red circle and a small yellow circle appeared on either side of the computer monitor (which circle was on which side was varied randomly across trials); between the circles, at the center of the monitor, was a tall black block that functioned as an obstacle. The large circle expanded and contracted, the small circle did the same, and this sequence was repeated a second time. Next, the small circle advanced in a straight line toward the block, retreated to its original position, and then sped toward the block, jumped over it, and landed on the other side. Finally, the small circle moved in a straight line until it contacted the large circle, at which point both circles expanded and contracted as before. Following habituation, the block was removed, and the infants saw two test events. In the *long-path* event, the small circle performed exactly the same actions as in the habituation event, even though the block was now absent. In the *short-path* event, after the two circles expanded and contracted at the start of the event, the small circle simply approached the large circle in a straight line.

The infants looked reliably longer at the long- than at the short-path test event, suggesting three conclusions. First, based on the information available in the habituation trials, the infants determined that the circles were agents (we return later to the question of how infants identify non-human entities as agents). Second, during the habituation trials, the infants attributed to the small circle the goal of contacting the large circle. Finally, during the test trials, the infants expected the small circle to maintain its goal and to pursue it efficiently: With the block removed, a shorter, more efficient path to the large circle became possible, and the infants detected a violation when the small circle ignored this path and followed the longer, less efficient path it had previously used instead. These conclusions were supported by the results of a control

condition identical to the first, experimental condition except that in the habituation trials the block stood *behind* the small circle, near the edge of the computer monitor, and thus no longer functioned as an obstacle. The infants looked about equally at the two test events, suggesting that they could not generate a satisfactory explanation for the small circle's actions in the habituation trials (e.g., why did it advance, retreat, and jump in every trial, instead of simply approaching the large circle in a straight line?) and hence had no expectation about its behavior in the test trials. These results were confirmed in other laboratories in experiments conducted with 12- to 14-month-olds, using videotaped or live events with humans or puppets as agents (e.g., Phillips & Wellman, 2005; Sodian, Schoeppner, & Metz, 2004). Phillips and Wellman (2005) also showed that infants could attribute to an agent the goal of contacting an inanimate target (in the habituation event, a woman reached over an obstacle to grasp a ball), but not the goal of reaching to an empty location (results were negative when the ball was absent).

To provide converging evidence for their findings, Csibra, Bíró, Koós, and Gergely (2003) also tested 12-month-olds in a different detour task. In the habituation event, a large red circle chased a small yellow circle until it passed between two bars separated by a gap. Because the gap was too narrow for the large circle to pass through, it detoured around the bars and then continued its pursuit until both circles disappeared off the computer monitor. In the test events, the gap between the bars was wider so that the large circle could now pass between them. In the long-path event, the large circle still detoured around the bars; in the short-path event, the large circle now followed the small circle through the gap until they both disappeared off the monitor. The infants looked reliably longer at the long- than at the short-path event, suggesting that they attributed to the large circle the goal of chasing the small circle and expected the large circle to do so efficiently. In a control condition, the red circle was the same size as the yellow circle but

still detoured around the bars in the habituation event; the infants looked about equally at the test events, presumably because they could not make sense of the red circle's inefficient actions during the habituation trials and hence held no expectation about its subsequent behavior.[6]

2. Do young infants also expect agents to pursue goals efficiently?

The experiments summarized in the last section indicated that, by their first birthday, infants expect an agent who is pursuing a goal to do so efficiently. In additional experiments, Csibra, Gergely, and their colleagues asked whether younger infants would also show sensitivity to efficiency constraints. Results were mixed: 9- but not 6-month-olds succeeded at the detour task of Csibra et al. (1999), and 9-month-olds failed at the detour task of Csibra et al. (2003).

Inspired by these results and those of Woodward (1998) (recall that young infants attributed a disposition to a human agent, but not to a rod, arm-shaped screen, or mechanical claw protruding from the side of the apparatus), Kamewari, Kato, Kanda, Ishiguro, and Hiraki (2005) asked whether young infants might be more likely to succeed at a detour task involving a human agent as opposed to a non-human entity. In a series of experiments, 6.5-month-olds were habituated to a man, a human-like robot, or a box moving around an obstacle to reach an inanimate target; in test, the obstacle was removed and infants saw, as before, a short- or a long-path event. Results were positive with the man and the robot, but negative with the box, suggesting that young infants could succeed at a detour task only if it involved a human or human-like agent. Additional detour experiments with 7- and 8-month-olds also obtained positive results with a human agent (e.g., Brandone & Wellman, 2009; Southgate, Johnson, & Csibra, 2008). We return in a later section to the negative results with non-human entities.

3. Can infants generate an explanation for an agent's apparently inefficient action?

In the control conditions of the detour tasks discussed in the previous sections, infants

watched an agent perform inefficient actions, such as retreating and jumping for no apparent reason while approaching a target. The negative results obtained in these control conditions suggest that infants could not make sense of the actions they observed and hence held no expectation about subsequent actions. In novel *inefficient-action* tasks conducted with different methods, Csibra, Gergely, and their colleagues examined whether infants might, under some conditions, generate explanations for apparently inefficient actions (e.g., Csibra et al., 2003; Gergely et al., 2002).

One inefficient-action VOE task (Csibra et al., 2003) built on prior evidence that even young infants posit hidden objects to make sense of physical events that would otherwise violate their expectations; for example, when an object moves back and forth behind a screen without appearing in a large opening at the bottom of the screen, infants as young as 3.5 months of age assume that two identical objects are used to produce the event, one traveling to the left and one to the right of the opening (e.g., Aguiar & Baillargeon, 2002). In a related vein, Csibra et al. asked in a VOE task whether 12- and 9-month-olds would posit a hidden obstacle to make sense of an agent's otherwise inefficient action. In the habituation event, a small red circle approached a large screen, jumped over the area behind the screen, landed on the other side, and then moved forward until it contacted a large yellow circle. In the test events, the small circle performed the same actions as before, but the screen was removed at the start of each event to reveal either an obstacle (*obstacle* event) or empty space (*no-obstacle* event). The 9-month-olds looked equally at the two test events, but the 12-month-olds looked reliably longer at the no-obstacle than at the obstacle event, suggesting that (1) they made sense of the agent's inefficient jumping action in the habituation event by positing an obstacle behind the screen and therefore (2) they detected a violation when the screen was removed to reveal a clear path (this interpretation was supported

by the results of a baseline condition: 12-month-olds who saw only the test events showed no preference for either event).

Another inefficient-action task (Gergely et al., 2002) was suggested by a puzzling contrast in findings from imitative-learning tasks. On the one hand, a great deal of research suggested that, by the second year of life, infants readily infer the goals underlying models' actions and selectively reproduce goal-relevant components in their own actions. For example, infants are more likely to imitate actions that are marked as intentional ("There!") than actions that are marked as accidental ("Woops!) (e.g., Carpenter, Akhtar, & Tomasello, 1998; Olineck & Poulin-Dubois, 2005); they are equally likely to reproduce intended outcomes after watching successful or incomplete demonstrations (e.g., Bellagamba & Tomasello, 1999; Johnson, Booth, & O'Hearn, 2001; Meltzoff, 1995; Olineck & Poulin-Dubois, 2009); and they are more likely to reproduce goal-relevant than goal-irrelevant action components (e.g., Carpenter, Call, & Tomasello, 2005). In contrast to this evidence of skilled and selective imitation, other findings suggested that infants in the second year willingly imitate inefficient actions, even when more efficient options are readily available. In a well-known task, Meltzoff (1988) found that, after watching a male model activate a light-box by touching it with his forehead, 14-month-olds who were presented with the light-box one week later were reliably more likely to reproduce this inefficient head action (67%) than were infants in control conditions (0%). But if infants in the second year of life are skilled imitators who selectively focus on goal-relevant components, why did the infants tested by Meltzoff not use the more efficient approach of touching the light box with their hands?

Gergely et al. (2002) speculated that the infants attempted to make sense of the model's head action and assumed that it "must offer some advantage in turning on the light" (p. 755); in

other words, a rational model would not use the inefficient head action to activate the light-box unless there was some reason for doing so. This speculation suggested that infants would be less likely to reproduce the model's head action if they were provided with a different explanation for it: Specifically, if they were shown that the model's hands were otherwise occupied, making the head action no more than an expedient, alternative means of activating the light-box. As predicted, 14-month-olds were reliably less likely to reproduce a female model's head action one week later if her hands were occupied during the demonstration (21%; the model wrapped herself in a blanket which she held with both hands) than if her hands were free (69%; the model wore the blanket loosely and laid her hands on either side of the light-box). Similar results were obtained in other laboratories with infants ages 12 to 18 months using a variety of inefficient actions (e.g., Chen & Waxman, 2013; Paulus, Hunnius, Vissers, & Bekkering, 2011a, 2011b; Pinkham & Jaswal, 2011; Poulin-Dubois, Brooker, & Polonia, 2011; Schwier, van Maanen, Carpenter, & Tomasello, 2006). These additional results indicate that infants' attempts at making sense of inefficient actions are highly sensitive to contextual cues. For example, Paulus et al. (2011b) found that 14-month-olds imitated a model's head action if her hands held balls on either side of the light-box; infants presumably reasoned that, had she wished to, the model could easily have touched the light-box with either a hand or a ball. In contrast, infants did not imitate the model's head action if she performed it while holding her arms wide open above her head; even though her hands were free, her markedly inefficient, extravagant actions no doubt defied explanation, and the infants simply activated the light-box with their hands.

4. A new test of the efficiency principle

One potential limitation of the various detour and inefficient-action tasks described in the previous sections is that infants *are always shown inefficient actions*, leaving the findings open

to deflationary interpretations. One such interpretation stems from recent proposals that early expectations about agents' actions are primarily statistical in nature (e.g., Paulus et al., 2011c; Perner, 2010; Ruffman, Taumoepeau, & Perkins, 2012). In this view, infants gather a wealth of statistical information about the actions agents produce in daily life. Because the inefficient actions shown in detour tasks are *infrequent*, they elicit novelty responses; thus, infants look longer when an agent performs a detour for no apparent reason, not because this action is inefficient, but because it deviates from learned statistical regularities.

One way to address this alternative interpretation is to create a novel test of the efficiency principle that does not involve infrequent actions. To this end, Scott and Baillargeon (2013) devised a novel two-object task in which an agent faced two identical objects, one of which was physically more accessible than the other; the rationale was that if infants were sensitive to efficiency constraints, then they would expect the agent to reach for the object that could be retrieved with less effort (as an analogy, think of an apple tree with two apples, one of which is within easier reach than the other).

In the *identical-objects* condition, 16-month-olds received four familiarization trials and one test trial. In each familiarization trial, a female agent sat centered behind two identical toy pigs; each pig stood in front of a long support (first two trials) or in front of a short platform resting on a long support (last two trials). As the agent watched, an experimenter's gloved hand entered the apparatus, placed each pig on its support or platform, and left. The agent then grasped the handle of the right or left support (counterbalanced across trials), pulled it, grasped the pig, and paused. The familiarization trials thus served to establish that the agent wanted a pig and did not care which one she obtained. In the test trial, a transparent cover and a transparent container stood centered on the right and left supports, respectively. The gloved hand placed the

right pig on its support and covered it with the transparent cover; next, the hand moved the container to the front end of the left support, placed the left pig in the container, and left. The agent then grasped the handle of the support with the pig under the cover (*more-effortful* event) or the handle of the support with the pig in the container (*less-effortful* event) and paused, without pulling the support.

The infants who saw the more-effortful event looked reliably longer than those who saw the less-effortful event, suggesting that they (1) attributed to the agent the goal of obtaining a pig, (2) realized that retrieving the pig in the container would require fewer actions (pull support, grasp pig) than retrieving the pig under the cover (pull support, lift cover, grasp pig), (3) expected the agent to choose the pig that could be obtained with less effort, in accordance with the efficiency principle, and hence (4) detected a violation when the agent grasped the support with the cover. Infants thus demonstrate sensitivity to efficiency constraints even in the context of typical, everyday scenes that do not involve inefficient actions. Support for this interpretation came from a *different-objects* condition in which one of the pigs was replaced with a toy apple. In each familiarization trial, the agent demonstrated that she preferred the apple over the pig. In the test trial, the apple was placed under the cover and the pig was placed in the container. The infants now looked reliably longer when the agent grasped the support under the container. Even though retrieving the apple required a longer action sequence than retrieving the pig, the infants expected the agent to perform this sequence, in accordance with the consistency constraint, because it was necessary to obtain her preferred toy.

5. Does the efficiency principle extend to mental effort?

The efficiency experiments discussed until now have all focused on agents' *physical* effort: We have seen that infants expect agents to choose shorter paths, or shorter action

sequences, to achieve their goals. But do infants also consider *mental* effort when reasoning about agents' actions? If an agent faced two identical objects that were physically equally accessible, but one was visible and one was hidden, would infants view the visible object as mentally more accessible (e.g., as easier to attend to or keep in mind) and hence expect the agent to reach for it rather than for the hidden object? To address this question, Scott and Baillargeon (2013) tested 16-month-olds with a two-object task adapted from that in the previous section.

In the familiarization trials shown in the *identical-objects* condition, an experimenter's gloved hand placed two identical toy pigs on placemats (first two trials) or in shallow containers (last two trials); the agent then grasped the right or left pig (counterbalanced across trials) and paused, indicating that she wanted a pig and did not care which pig she obtained. In the test trial, an opaque and a transparent cover stood behind the right and left pigs, respectively; a small screen lay flat on the apparatus floor, centered in front of the left pig. After a few seconds, the screen was rotated upwards to hide the pig. The gloved hand then covered each pig with its cover and left; the top of the transparent cover protruded above the screen. Next, the agent grasped the opaque cover (*more-effortful* event) or the transparent cover (*less-effortful* event) and paused.

The infants looked reliably longer if shown the more- as opposed to the less-effortful event, suggesting that they (1) attributed to the agent the goal of obtaining a pig, (2) determined that, for the agent, the pig visible under the transparent cover was mentally more accessible than the pig hidden under the opaque cover, (3) expected the agent to choose the pig that could be obtained with less effort, in accordance with the efficiency principle, and hence (4) detected a violation when the agent grasped the opaque cover. Infants thus consider efficiency in mental effort when determining which of two identical objects an agent is likely to select. Results from a *different-objects* condition supported this interpretation. In this condition, one of the pigs was

replaced with a toy apple, the agent preferred the apple over the pig in the familiarization trials, and the apple was placed under the opaque cover in the test trial. The infants now looked reliably longer when shown the less-effortful event: Even though keeping in mind a hidden object may require more mental effort than keeping in mind a visible object, infants expect an agent to expend this effort when it is necessary to obtain a preferred toy.

Together, these results and those reported in the previous section suggest two broad conclusions. First, the efficiency principle is quite abstract: Infants expect agents to expend as little physical or mental effort as possible when pursuing their goals. Second, infants rank consistency above efficiency, at least in situations where the effort required for obtaining a more preferred object is not much greater than that required for obtaining a less preferred object. Adults, of course, respond similarly: We would not be surprised if a friend traveled extra miles to dine at his preferred restaurant, ignoring nearer but less preferred restaurants. Efficiency will typically be assessed relative to the specific goal an agent has elected to pursue—unless this goal requires such an outlay of resources as to become irrational.

1-II. IS EARLY PSYCHOLOGICAL REASONING EXPERIENCE-DRIVEN OR SYSTEM-BASED?

As we saw in section 1-I, beginning with the publication of the seminal results of Csibra, Gergely, Woodward, and their colleagues (e.g., Csibra et al., 1999; Gergely et al., 1995; Woodward, 1998, 1999), a wealth of findings appeared that dramatically increased our understanding of early psychological reasoning. These findings supported the view that, when observing an agent act in a scene, infants attempt to make sense of the agent's actions; if they infer that these actions are driven by a specific goal or disposition, then they expect the agent's subsequent actions to follow the consistency and efficiency principles. Although this

accumulating evidence made clear that infants' action knowledge was far more sophisticated than had hitherto been believed, one important issue that remained unresolved concerned the developmental origins of this knowledge. By the early 2000s, there were at least two main views on this issue; we refer to them as the experience-driven and the system-based view.

Proponents of the *experience-driven* view generally assumed that infants' own experiences, as intentional agents performing goal-directed actions and as participants in social interactions, play a critical role in the development of their ability to understand others' actions (for reviews, see Meltzoff, 2005; Tomasello, Carpenter, Call, Behne, & Moll, 2005; Woodward, 2005). For example, Woodward (2005) proposed that, as infants learn to produce various goal-directed actions, they become able to understand similar actions by similar agents, in part due to innate neurocognitive mechanisms (such as mirror neurons) for establishing equivalences between their own actions and those of other agents. Over time, through the application of cognitive comparison and conceptual abstraction processes to the representations of their own actions and those of other agents, infants gradually construct a more abstract and more general understanding of intentional action, which can then be applied, broadly and flexibly, to novel actions and novel agents. According to Woodward (2005), "infants' action knowledge is experience-driven, constrained by developmental progressions in the motor domain, and dependent on innate pathways for establishing mirroring systems...real-world action knowledge provides the developmental basis for more abstract conceptions of intentional action" (p. 254).

On the other hand, proponents of the *system-based* view generally held that infants are born with a psychological-reasoning system that provides them with a skeletal causal framework for making sense of the actions of agents (e.g., Baron-Cohen, 1995; Gergely & Csibra, 2003; Johnson, 2003; Leslie, 1994b; Premack, 1990). This initial framework includes concepts such as

agent, *goal*, and *disposition*, as well as principles such as *consistency* and *efficiency*. According to the system-based view, these concepts and principles are not gradually constructed from local actions and mirrored perceptions, but rather guide from the start infants' psychological reasoning. Advocates of the system-based view did not, of course, deny that experience plays an important role in infants' expectations about specific agents and situations. Although infants' psychological-reasoning system provides them with an abstract interpretive framework (e.g., they do not need to learn what a *goal* is), they must still learn, for example, what goals different agents may pursue at different times, what objects, foods, or activities different agents may prefer, and so on.

Was early action knowledge primarily experience-driven or system-based? Part of the evidence that initially seemed to support the experience-driven view was that young infants typically failed at VOE tasks involving non-human agents. Recall that Woodward (1998) found that 5- to 9-month-olds failed at two-object tasks in which the human agent's arm was replaced with a rod, an arm-shaped screen, or a mechanical claw; that Csibra et al. (1999) found that 9- but not 6-month-olds succeeded at a detour task in which a small circle approached a large circle; that Csibra et al. (2003) found that 12- but not 9-month-olds succeeded at a detour task in which a large circle chased a small circle; and that Kamewari et al. (2005) found that 6.5-month-olds succeeded at a detour task when the agent was a human or a human-like robot, but not a box. These results were consistent with the experience-driven view that "infants begin with local understandings of goal-directed action that become gradually broader over the course of the first year of life" (Woodward, 2005, p. 247). This was often referred to as the *humans-first* hypothesis: Infants begin by reasoning about agents similar to themselves—human agents—but by their first birthday can extend this understanding to non-human agents.

1. New evidence with non-human agents

Was the humans-first hypothesis correct? In time, researchers began to suggest alternative interpretations of the negative results reported in the last section. In particular, one limitation of Woodward's (1998) findings with the rod, arm-shaped screen, and mechanical claw was that all of these novel objects protruded from a window in the right wall of the apparatus—the right end of each object never came into view—so that infants might have been uncertain whether the objects were agents or not. This alternative interpretation suggested that young infants might succeed at a two-object task involving a non-human agent if given clearer evidence that they were facing an agent.

To examine this possibility, Luo and Baillargeon (2005a) tested 5-month-olds in a two-object task that was modeled after that of Woodward (1998) but used a box as the agent. The infants first received two orientation trials in which the box moved back and forth across the center of the apparatus floor. In the familiarization trials, a cylinder and a cone stood on the left and right sides of the apparatus, respectively; in each trial, the box approached and rested against the cone. In the display trial, the locations of the cylinder and cone were switched. Finally, in the test trials, the box approached either the cone (old-object event) or the cylinder (new-object event). The infants looked reliably longer at the new- than at the old-object event, suggesting that they were able to reason about the actions of a non-human agent: They attributed to the box a preference for the cone, and they detected a consistency violation when the box approached the cylinder instead. Negative results were obtained in a one-object task in which only the cone was present in the familiarization trials: Although infants could attribute to the box the goal of approaching the cone in the familiarization trials, they could not determine whether the box had a positive or only a neutral disposition toward the cone, and hence they could not predict which

object—the cone or the cylinder—the box would approach in the test trials. Luo (2011b) subsequently extended these results to 3-month-old infants.

Finally, to test more directly their interpretation of Woodward's (1998) results with the rod, arm-shaped screen, and mechanical claw, Luo and Baillargeon (2005a) conducted another experiment in which the box was equipped with a long handle that protruded through a window in the right wall of the apparatus. As predicted, infants looked equally at the two test events, presumably because they were uncertain whether the box was an agent. Infants did identify the box as an agent, however, when the handle was cut short so that its unfettered right end was always in view, making it appear as though the box was controlling its own actions.

Together, the results of these experiments cast serious doubt on the humans-first hypothesis and more generally on the experience-driven view: If young infants can already reason about non-human agents very dissimilar from themselves, it is unlikely that their action knowledge is derived from local experiences and understandings that are gradually generalized over the course of the first year.

2. How do infants identify non-human agents?

The positive results obtained with younger as well as older infants in experiments with non-human agents naturally gave rise to a key question: How did infants *identify* a novel entity as an agent? Since a circle or a box could be viewed as an agent, it was clear that animal-like properties (e.g., having a face or engaging in biological motion) were not necessary for agency. What sort of evidence, then, led infants to view novel entities as agentive?

Evidence of control. Ground-breaking experiments by Johnson and her colleagues (e.g., Johnson, Shimizu, & Ok, 2007; Shimizu & Johnson, 2004) helped shed light on this question. In these experiments, 12-month-olds received a two-object task in which an oval-shaped "blob"

covered with green fiberfill stood near the front of the apparatus; at the back of the apparatus were two toys, object-A on one side and object-B on the other. During each habituation trial, the blob approached and stopped against object-A. During the test trials, the toys' positions were switched, and the blob approached either object-A (old-object event) or object-B (new-object event). At the start of each habituation and test trial, the blob's front-to-back axis was aligned with the object it approached during the trial. The infants looked about equally at the new- and old-object events, suggesting that they did not view the blob as an agent and held no expectation about its displacements in the test trials.

However, infants did view the blob as an agent—as indicated by the fact that they looked reliably longer at the new- than at the old-object test event—in two key conditions. In one, instead of being aligned with object-A at the start of each habituation trial, the blob faced a position midway between the two toys and "turned" toward the object—as though making a choice—before approaching it (Schlottmann and Ray (2010) obtained similar results with 6-month-olds). In the other condition, prior to approaching object-A in the first habituation trial, the blob participated in a scripted "conversation" with an experimenter; the experimenter spoke English and the blob responded with a varying series of beeps. Interestingly, negative results were obtained (1) if the experimenter spoke but the blob remained silent (suggesting that it was not merely seeing the experimenter talk to the blob that led the infants to view it as an agent), or (2) if the blob beeped but the experimenter remained silent (suggesting that it was not merely observing that the box could produce varying beeps that led the infants to view it as an agent; apparently, variable self-generated behavior, if it appears random, does not constitute evidence of agency). In converging experiments using a gaze-following measure, Johnson et al. (2008) found that, after observing the blob turn toward one of two targets, 14- to 15-month-olds turned

in the same direction if the blob first "conversed" with an experimenter (agent condition), but not if it beeped and the experimenter remained silent (non-agent condition). Deligianni, Senju, Gergely, and Csibra (2011) extended these results to 8-month-olds, who turned in the same direction as a novel non-human object only if it first interacted with them contingently (i.e., produced a specific response when they fixated it).

These results, together with those from previous sections, indicate that for infants, evidence that a novel entity is self-propelled, is spoken to by a human agent, is able to produce an action that could be described as goal-directed (e.g., approaching and contacting object-A), or is capable of generating several different actions (e.g., producing varying beeps and then approaching and contacting object-A), does *not* constitute conclusive proof of agency. Rather, infants categorize a novel entity as an agent if it gives sufficient evidence that it has *control* over its actions. This evidence can take several different forms, including *orienting toward and approaching an object* in a scene (think of the blob used by Johnson, Shimizu, et al. (2007) first turning toward object-A and then approaching it), and *responding to a change* in a scene (think, for example, of the blob used by Johnson, Shimizu, et al. (2007) beeping in response to the experimenter, of the large and small circles used by Gergely et al. (1995) expanding and contracting in turn, or of the box used by Luo and Baillargeon (2005a) detecting and approaching the cone).

Work by Csibra (2008) uncovered yet another form of evidence indicative of control for infants: *choosing different means to achieve the same goal* in a scene.[7] Csibra speculated that Kamewari et al. (2005) obtained negative results in their detour task with 6.5-month-olds when they used a box (as opposed to a human or a human-like robot) as the agent, not because these young infants could not reason about non-human agents, but because the box did not give

sufficient evidence that it was agentive: It reached its target by moving around the right side of the obstacle in each and every habituation trial, making it unclear whether it had control over its behavior. To test this speculation, Csibra repeated the same experiment with 6.5-month-olds except that the box no longer followed a fixed path: It moved randomly around the right *or* the left side of the obstacle across habituation trials. The infants now looked reliably longer at the long- than at the short-path event, suggesting that this slight choice-dependent variation was sufficient to lead them to identify the box as an agent. Csibra's results thus provided converging evidence that, when given appropriate information, even young infants can view a non-human entity as an agent (e.g., Luo, 2011b; Luo & Baillargeon, 2005a; Schlottmann & Ray, 2010).[8]

Evidence of autonomy. In the experiment of Luo and Baillargeon (2005a), infants viewed the box as an agent when it had no handle or a short handle, but not when it had a long handle that protruded through the right wall of the apparatus, even though in each case the box performed exactly the same actions. These results suggest that, in determining whether a novel entity is an agent, infants consider not only whether it is producing controlled actions, but also whether the control for these actions resides within the entity itself—in other words, whether the entity has *autonomous* control over its actions. For the young infants tested by Luo and Baillargeon (2005a) and Woodward (1998), it was presumably unclear whether the long-handle box, rod, arm-shaped screen, and mechanical claw were acting on their own or not.[9]

3. Can an agent be inert?

The results of Kamewari et al. (2005), Johnson, Shimizu, et al. (2007), and Csibra (2008) all indicated that self-propulsion or *autonomous motion* alone is not sufficient to lead infants to categorize a novel entity as an agent: The entity must provide evidence that it has *autonomous control* over its actions. These findings left open the possibility that self-propulsion, though not

sufficient, is still *necessary* for infants to view an entity as an agent. Is that really the case? Adults, of course, can view inert objects as agents, and the world of fiction offers many such examples: Think, for example, of the magic mirror in the fairy-tale "Snow White", the sorting hat in the "Harry Potter" books, the evil ring in the "Lord of the Ring" books, or the computer Hal in the movie "2001: Space Odyssey".

Can infants also view an inert object as an agent? To address this question, Wu and Baillargeon (in Baillargeon, Wu, Yuan, Li, & Luo, 2009b) tested 14-month-olds in a novel two-object task. In the initial trials, a box responded with beeps in a conversation with an experimenter, but otherwise remained stationary. Next, the infants received familiarization trials in which the box stood centered behind two small covers; hidden under the covers were object-A and object-B (a ball and a block). In each trial, a gloved hand lifted and lowered each cover in turn; the box beeped when the left cover was lifted to reveal object-A, but not when the right cover was lifted to reveal object-B. Next, the infants received a display trial in which the box was absent and the hand lifted and lowered each cover to show that the toys' positions had been switched. Finally, in the test trials, the box stood in its original position, the hand lifted and lowered the left and right covers in turn, and the box either beeped when object-B but not object-A was revealed (new-object event) or beeped when object-A but not object-B was revealed (old-object event). The infants looked reliably longer at the new- than at the old-object event, suggesting that they (1) noticed during the familiarization trials that the box beeped only when object-A was revealed, (2) used this choice information to attribute to the box a preference for object-A, and (3) and expected the box to continue to act on this preference in the test trials. This interpretation was supported by the results of a one-object task: When only object-A was present in the familiarization trials, infants looked about equally at the two test events, because they had

no information about the box's disposition toward object-A.

Could infants have concluded that the stationary box was self-propelled, even though they never saw it move on its own, either because it could beep (on the assumption that beeping implied some capacity for self-generated motion) or because it was an agent (on the default assumption that all agents must be self-propelled)? Additional results indicated that this was not the case, and that infants continued to view the stationary box as inert even after it beeped in the familiarization trials to demonstrate its preference for object-A over object-B (for details, see Baillargeon et al., 2009b).

These results, together with those in the previous section, suggest that for infants self-propulsion is neither sufficient nor necessary for agency: Like adults, infants can reason about inert agents. More generally, these various results suggest that *self-propelled object* and *agent* are two distinct concepts for infants, with the former being a physical-reasoning concept (an object capable of autonomous motion has *internal energy*; e.g., Leslie, 1994b; Luo, Kaufman, & Baillargeon, 2009) and the latter a psychological-reasoning concept (as we will see in the next section, an object capable of autonomous control has *mental states*).[10]

1-III. IS EARLY PSYCHOLOGICAL REASONING TELEOLOGICAL OR MENTALISTIC?

The evidence reviewed in section 1-II indicated that young infants can interpret the actions of *any entity* they identify as an agent, whether human or non-human. This evidence cast doubt on the humans-first hypothesis and more generally supported the system-based as opposed to the experience-driven view. Recall that according to the system-based view, infants are born with a psychological-reasoning system that provides them with a skeletal causal framework for interpreting the actions of agents. What remained unclear, however, was the precise nature of

this initial framework. There were at least two broad views on this issue, and we refer to them as the mentalistic and the teleological view.

According to the *mentalistic* view, even young infants are capable of attributing (at least simple) mental states to agents (e.g., Baron-Cohen, 1994; Johnson, 2005; Leslie, 1994b; Luo & Baillargeon, 2005a; Premack & Premack, 1995). According to the *teleological* view, in contrast, early psychological reasoning is non-mentalistic and deals exclusively with physical variables: Infants are initially "mindblind" and cannot yet represent mental states (e.g., Csibra et al., 1999, 2003; Gergely et al., 1995; Gergely & Csibra, 2003). Thus, when watching an agent act in a scene, the psychological-reasoning system generates a teleological explanation that specifies the physical layout of the scene (e.g., the presence and location of obstacles), the agent's actions, and the end-state the agent achieves. These teleological explanations, together with the rationality principle, are sufficient to allow infants to predict future actions. Thus, if agent-A repeatedly jumps over an obstacle and approaches agent-B, infants will expect agent-A to arrive at the same end-state again in the future, in accordance with the consistency principle. Furthermore, if the obstacle is removed, infants will now expect agent-A to travel to agent-B in a straight line, in accordance with the efficiency principle. Proponents of the teleological view argued that, with development, physical variables gradually become incorporated into a mentalistic system that makes sense of intentional actions in terms of goals and other mental states.

Are infants initially capable of only teleological reasoning? If yes, at what age do they transition to mentalistic reasoning? To address these questions empirically, researchers focused on one key assumption about teleological reasoning: Because it deals exclusively with physical variables, teleological reasoning must be "reality-based" (Gergely & Csibra, 2003); *only one* representation of the physical layout is possible, the one construed by the infant. Thus, when

watching an agent in a scene, infants should egocentrically expect the agent to possess the same knowledge about the scene that they themselves possess. For example, if they see an object that the agent (from her specific viewpoint) cannot see, they should still expect her to act as though she knows about the hidden object. In order to overcome this egocentric outlook, infants would need to be able to attribute to the agent a separate (and in this case less complete) perspective on the scene, and such *perspective-taking* cannot be accomplished within the confines of mindblind, reality-based teleological reasoning—it can only be achieved via mentalistic reasoning. (We are referring here to what is sometimes called level-1 perspective taking, the ability to determine *what* others can see or have seen; level-2 perspective taking refers to the ability to determine *how* the same object or scene will appear to observers in different locations; Masangkay et al., 1974).

The preceding analysis provided a straightforward way of testing whether early psychological reasoning was mentalistic or teleological in nature: Positive evidence that infants were capable of simple perspective-taking would argue for mentalistic reasoning, whereas negative evidence would argue for teleological reasoning.

1. Perspective-taking in older infants

To determine whether someone can or cannot see an object, one must possess some knowledge about visual perception. A great deal of research suggested that, by the second year of life, infants already understand under what conditions an agent can see an object: They realize that the agent's eyes must be open and oriented toward the object, and that her line of sight must not be obstructed by an opaque barrier (e.g., Brooks & Meltzoff, 2002, 2005; Butler, Caron, & Brook, 2000; Caron, Kiel, Dayton, & Butler, 2002; Dunphy-Lelii & Wellman, 2004; Lempers, Flavell, & Flavell, 1977). Meltzoff and Brooks (2007) also found that, although 12-month-olds expected an agent to see a target object even when her view was obstructed by a blindfold, brief

experience with the blindfold (held to their eyes) was sufficient to help them understand that the blindfold blocked the agent's view (see also D'Entremont & Morgan, 2006).

Building on these results, subsequent experiments established that infants aged 12 months and older are capable of simple perspective-taking: They keep track of what objects and events an agent can or cannot see, and has or has not seen, and they use that information both to interpret the agent's actions and to guide their own responses to the agent (e.g., Koenig & Echols, 2003; Luo & Baillargeon, 2007; Luo & Beck, 2010; Sodian, Thoermer, & Metz, 2007; Tomasello & Haberl, 2003). For example, Tomasello and Haberl (2003) found that, when an agent requested one of three objects excitedly, 12- and 18-month-olds gave her the one she had not seen previously, suggesting that they kept track of which objects the agent had or had not experienced during the testing session. Repacholi, Meltzoff, and Olsen (2008) reported that, after watching an agent angrily scold an experimenter for playing with an "irritating" toy, 18-month-olds were more likely to play with the toy if the agent did not look at them (read a magazine or closed her eyes) than if she looked at them directly. In an extensive series of experiments, Liszkowski and his colleagues showed that infants aged 12 months and older spontaneously pointed to inform ignorant but not knowledgeable agents (e.g., Liszkowski, Carpenter, & Tomasello, 2007, 2008; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, Striano, & Tomasello, 2006). Thus, 12-month-olds pointed to inform an agent about the current location of an object she was looking for if she was absent when it was moved to a new location (e.g., Liszkowski et al., 2006), or if she was looking away when it fell to the floor (e.g., Liszkowski et al., 2008).

Finally, additional experiments indicated that older infants realize that, just as they may see an object that an agent cannot see, an agent may see an object that they themselves cannot

see (e.g., Chow, Poulin-Dubois, & Lewis, 2008; Moll & Tomasello, 2004; Repacholi, 1998). For example, Moll and Tomasello (2004) found that, when an agent looked behind an opaque barrier with expressions of excitement, 12- and 18-month-olds crawled or walked forward a short distance in order to peek behind the barrier and see what the agent could see.

Together, these results suggested that infants 12 months and older are capable of simple perspective-taking: They realize that an agent's representation of a scene may differ from their own, and in such cases they use the agent's perspective to interpret and respond to her actions. These results provided robust evidence that, by 12 months of age, infants' psychological reasoning is mentalistic rather than teleological in nature. However, these results left unclear the nature of psychological reasoning in the first year of life, and as such left open the possibility that psychological reasoning was initially teleological.

2. Perspective-taking in younger infants

Are young infants capable of simple perspective-taking? Do they recognize that an agent's representation of a scene may be less complete than their own? To address this question, Luo and Johnson (2009) tested 6-month-olds with novel two-object tasks (adapted from Luo & Baillargeon, 2007). In one experiment, infants received four familiarization trials in which an agent sat between two objects, object-A and object-B, and consistently reached for object-A. In one condition, both objects were visible to the infant, but only object-A was visible to the agent: A tall screen hid object-B from her (hidden-object condition). In two other conditions, both object-A and object-B were visible to the agent: Either the screen was shorter so that object-B protruded above it (short-screen condition), or object-B itself was taller so that it protruded above the tall screen (tall-object condition). Following the familiarization trials, the objects' positions were reversed, the screen was removed, and the agent reached for either object-A (old-

object event) or object-B (new-object event). In the short-screen and tall-object conditions, the infants looked reliably longer at the new- than at the old-object event; in the hidden-object condition, in contrast, the infants looked about equally at the two events. The infants thus realized that the agent's repeated actions on object-A during the familiarization trials could not be interpreted as revealing a preference for object-A over object-B if the agent could not see object-B (these results are of course consistent with those of the one-object tasks discussed earlier: For the agent, it was as though object-A was the only object present in the scene). In a second experiment, either the agent sat facing object-A with her back to object-B (back condition), or she sat facing both objects (front condition). As expected, the infants in the front condition looked reliably longer at the new- than at the old-object event, whereas those in the back condition looked about equally at the two events. Here again, the infants concluded that the agent preferred object-A over object-B only when she could see both objects during the familiarization trials.

These results provided evidence that, by 6 months of age, psychological reasoning is mentalistic rather than teleological in nature: Infants did not appear to be mindblind, egocentric observers who believed that the agent must share their reality-based representation of the scene. To the contrary, infants (1) understood that the agent would not see an object they themselves could see if it either stood behind her or was hidden from her by an opaque barrier, and (2) used the agent's incomplete representation of the scene to interpret her actions. Infants concluded that the agent preferred object-A over object-B only when both objects were visible to *her* during the familiarization trials—even though both objects were always visible to *them*.

3. Can young infants attribute inference-based knowledge to others?

The evidence reviewed in the two previous sections suggested that infants aged 6 months

and older are capable of mentalistic reasoning: When they watch an agent act in a scene, their psychological-reasoning system attempts to build an explanation that specifies the agent's likely mental states. In particular, *motivational* states capture what motivates the agent in the scene and include goals and dispositions, whereas *epistemic* states capture how much the agent knows about the scene and include knowledge and ignorance states. We have suggested that a *masking* mechanism makes it possible for an infant to represent what an agent does not know about a scene; this mechanism masks or blocks the information that is unavailable to the agent, enabling the infant to predict and interpret the agent's actions in terms of the remaining, shared information (e.g., Scott & Baillargeon, 2009).

The next question that had to be addressed about epistemic states concerned the range of knowledge states that infants could attribute to others. Specifically, could infants attribute only knowledge states based on *perception*, either present or past, or could infants also attribute knowledge states derived though *inference*? As adults, we recognize that agents may infer a great deal of information about a scene, either based on their previous experiences in the scene (e.g., object-A must be hidden in the red container, because that is where it has always been hidden in the past) or based on their general world knowledge (e.g., object-A must be hidden in the wide container, because it is too wide to fit into the narrow container). If infants could attribute only perception-based knowledge to others, it would suggest that early mentalistic reasoning consists of a relatively simple form of *perception-action* reasoning: Agents act on what they immediately see or have recently seen. In contrast, if infants could attribute inference-based knowledge to others, then it would indicate that early mentalistic reasoning goes beyond simple perception-action reasoning: Knowledge is understood in fairly abstract terms, as information that can be gathered directly through perception or indirectly through inference.

He and Baillargeon (2012) devised a novel two-object task to examine whether 5-month-olds could attribute inference-based knowledge to an agent. The infants in the *inference* condition received four familiarization trials and one test trial. During the familiarization trials, an agent sat at a window in the back wall of an apparatus, centered behind a wide and a narrow toy (side was counterbalanced across trials); each toy stood in front of a placemat. Each toy consisted of a tall handle attached to either a wide or a narrow rectangular base; the toys differed in color and pattern, but the top portions of their handles were identical. In each trial, an experimenter's gloved hand grasped the handle of the nearer toy, shook it, placed it on its placemat, and then repeated these actions with the farther toy; the wide toy jingled when shaken, but the narrow toy did not. After the hand completed its actions and left, the agent grasped the handle of the wide toy, shook it, and then returned it to its placemat. The familiarization trials thus served to establish that the agent preferred the wide, jingling toy and reached for it wherever it happened to be. In the test trial, the agent was initially absent: Her window was closed. The wide toy now stood in front of a tall wide green box, and the narrow toy stood in front of a tall narrow blue box; the wide toy could fit only into the wide box. The hand shook the narrow toy and placed it in its box, repeated these actions with the wide toy, and then left. Because only the identical tops of the toys' handles protruded from the boxes, it was impossible to determine via visual inspection alone which box held which toy. Next, the agent opened her window, grasped the handle of either the wide (old-object event) or the narrow (new-object event) toy, and paused.

The infants looked reliably longer if shown the new- as opposed to the old-object event, suggesting that (1) during the familiarization trials, they attributed to the agent a preference for the wide, jingling toy, and (2) during the test trial, they expected the agent to infer which box held the wide toy, based on the physical knowledge that wide objects cannot fit inside narrow

containers (infants as young as 4 months of age possess this knowledge; Wang et al., 2004).

Support for this conclusion came from three additional conditions. In the *ignorance* condition, the two tall boxes were both wide; infants looked equally at the two test events, suggesting that they realized that the agent had no basis for inferring which box held the wide toy. The *memory* condition was identical to the ignorance condition except that the agent was present throughout the test trial and thus saw the gloved hand hide the toys; infants now looked reliably longer at the new-object event, suggesting that they expected the agent to remember which box held her preferred toy. Finally, the *control* condition was identical to the inference condition, with two exceptions: the agent was present throughout the test trial, and the wide toy no longer jingled (as though broken) when shaken by the gloved hand. Infants looked equally whether the agent grasped the handle protruding from the narrow or the wide box, suggesting that they realized that the agent might no longer hold a preference for the wide, silent toy.

Together, these results support two conclusions. First, they provide additional evidence that early psychological reasoning is mentalistic in nature: Even young infants recognize, non-egocentrically, that an agent may not know something that they themselves know. Thus, the infants in the ignorance condition expected the *agent* not to know where the wide toy was hidden, even though *they* knew where it was (as confirmed by the results of the memory condition). Second, these results indicate that early mentalistic reasoning goes beyond simple perception-action reasoning: Even young infants recognize that others may attain knowledge through inference. Thus, the infants in the inference condition expected the agent to correctly guess which box held the wide toy, even though she had not witnessed the hiding of the toys.

4. Epistemic states and the consistency principle

The evidence reviewed in the last two sections indicates that even young infants engage

in mentalistic reasoning and attribute epistemic as well as motivational states to agents. Since agents are essentially entities that detect their environment (epistemic states) and pursue goals within it (motivational states), the ability to represent these two types of mental states allows infants to begin to make sense of agents' actions.

We saw earlier that infants apply the consistency principle to motivational states: Agents should act in a manner consistent with their goals or dispositions. Do infants also apply the consistency principle to epistemic states? Two-object tasks such as those of Luo and Johnson (2009) and He and Baillargeon (2012) cannot shed light on these questions, because the results are insufficient to determine whether infants are applying the consistency principle to the agent's motivational or epistemic states. To illustrate, consider once again the memory condition of He and Baillargeon: We cannot know for certain whether infants detected a violation in the new-object test event because the agent's actions were inconsistent with (1) her epistemic states—she had just seen where the gloved hand hid the wide toy, so why was she reaching for the wrong toy? or (2) her motivational states—she had shown a robust preference for the wide toy, so why was she now reaching for the narrow toy? In short, to examine whether infants expect an agent to act in a manner consistent with her epistemic states, we need tasks where actions reflect primarily these states. Of particular relevance here are tasks in which infants are presented with either a *reliable* agent who acts in accordance with her epistemic states or an *unreliable* agent who does not. We summarize some of the findings obtained with such tasks below.

Older infants. There is mounting evidence that infants in the second year of life notice when an agent fails to act in a manner consistent with her knowledge. For example, in an *unreliable-labeler* task, 16-month-olds looked longer at an agent who provided incorrect labels for familiar objects if she faced the objects than if she faced away from them (Koenig & Echols,

2003; for related work with preschoolers, see Koenig, Clement, & Harris, 2004; Koenig & Harris, 2005). Building on these results, Poulin-Dubois and her colleagues tested 14- to 16-month-olds in a series of *unreliable-looker* tasks: These examined whether infants, after watching an agent act in a manner either consistent or inconsistent with her epistemic states in a first context, would hold expectations about her behavior in a second context (Chow et al., 2008; Chow & Poulin-Dubois, 2009; Poulin-Dubois et al., 2011). In the first context, the agent expressed excitement ("Wow!") when looking inside a bucket that either contained a toy (reliable-looker condition) or was empty (unreliable-looker condition). The second context was adapted from prior tasks and varied across experiments, but in each case infants held expectations for the actions of the reliable but not the unreliable looker. Thus, infants were more likely to activate a light-box with their foreheads after watching the reliable looker perform this action (Poulin-Dubois et al., 2011; adapted from Gergely et al., 2002); they were more likely to peek around a barrier after watching the reliable looker express excitement as she looked behind the barrier (Chow et al., 2008; adapted from Moll & Tomasello, 2004); and they were more likely to detect a violation when the reliable looker searched for an object in the incorrect location (Chow & Poulin-Dubois, 2009; adapted from Onishi & Baillargeon, 2005).

Together, these results suggest that by the second year of life infants expect an agent to act in a manner consistent with her epistemic states. If the agent fails to do so (e.g., exclaims over a bucket she can plainly see is empty), the psychological-reasoning system cannot generate a well-formed, rational explanation for her actions, and infants hold no expectation about her subsequent behavior. These results extend those of various control conditions discussed earlier in this chapter: Recall that after watching habituation events in which an agent failed to act efficiently (e.g., performed an unnecessary detour), infants held no expectation about the agent's

actions in the test events. A final report with an *unreliable-user* task suggests that, in addition to consistency and efficiency, infants consider normality (the third corollary of the rationality principle; see section 1-I-B) when assessing agents' actions. Zmyj, Buttelmann, Carpenter, and Daum (2010) first showed 14-month-olds events in which a male agent used everyday objects in either a normal manner (reliable-user condition; e.g., putting sunglasses on his nose, telephoning with his ear, using a toothbrush to brush his teeth) or an abnormal manner (unreliable-user condition; e.g., putting sunglasses on his foot, telephoning with the top of his head, using a toothbrush to brush his hand). Next, the agent activated a light-box with his forehead. The infants were reliably more likely to imitate this action when performed by the reliable user. These results suggest that by 14 months, infants have begun to accumulate a repertoire of knowledge about how everyday objects are used; if an agent repeatedly deviates from this normal usage, infants view him as irrational and interpret his subsequent actions accordingly (for an interesting exception showing infants' sensitivity to humor cues, see Hoicka & Wang, 2011). Adults, of course, make similar normality judgments: We would be alarmed if Aunt Lucinda arrived at a family reunion wearing underwear on her head, and we would be highly unlikely to follow her tips for hot stocks.

  Younger infants. Csibra and Volein (2008) examined whether young infants also expect an agent to act in a manner consistent with what she sees. In their experiment, 8- and 12-month-olds saw test events in which an agent sat behind and between two large screens. In each event, the agent first looked intently behind one of the screens; next, a curtain was lowered in front of her, and the two screens were lifted to reveal a toy behind one screen (full location) and no toy behind the other screen (empty location). At both ages, infants looked reliably longer at the empty location when the agent had first looked at that location (as though wondering what the

agent could have been looking at) than when she had first looked at the other location. Thus, like the older infants in the experiments discussed above, these younger infants were puzzled when the agent looked intently at an empty location.

1-IV. IS EARLY PSYCHOLOGICAL REASONING NON-REPRESENTATIONAL OR REPRESENTATIONAL?

The evidence reviewed in section 1-III indicated that, from an early age, psychological reasoning is mentalistic rather than teleological in nature. When watching an agent act in a scene, young infants are not egocentric, "mindblind" observers who can produce only "reality-based" interpretations of her actions. Rather, young infants attribute both epistemic and motivational states to the agent; moreover, if her knowledge about the scene is less complete than their own, they correctly use her knowledge to interpret her actions. The next issue that needed to be addressed was whether infants' mentalistic reasoning was qualitatively similar to that of older children and adults or differed from it in critical respects. This discussion centered chiefly on children's ability to represent *counterfactual* states such as *false beliefs*.

Beginning with the seminal work of Wimmer and Perner (1983), initial investigations of children's false-belief understanding used *elicited-response* tasks in which the child is asked a direct question about the likely behavior of an agent who holds a false belief about a scene (e.g., Baron-Cohen, Leslie, & Frith, 1985; Wellman & Bartsch, 1988). For example, in the classic *Sally-Anne* task (Baron-Cohen et al., 1985), children listen to a story enacted with props: Sally hides a marble in a basket and then leaves; in her absence, Anne moves the marble to a nearby box; Sally then returns, and children are asked where she will look for her marble. Beginning at about age 4 or 5, children typically answer correctly and point to the basket (where Sally falsely believes the marble is); in contrast, most 3-year-olds point to the box (where the marble actually

is), suggesting that they do not yet understand that Sally holds a false belief about the marble's location. This developmental pattern (from below-chance to above-chance performance) was confirmed with elicited-response tasks testing different false beliefs, such as false beliefs about expected contents (e.g., Gopnik & Astington, 1988; Perner, Leekam, & Wimmer, 1987). This developmental pattern was also observed in countries around the world, although its timing varied somewhat, with above-chance performance not being attained until age 6 or 7 in some countries (e.g., Callaghan et al., 2005; Liu, Wellman, Tardif, & Sabbagh, 2008; Vinden, 1999; Wellman, Cross, & Watson, 2001). Broadly speaking, two very different accounts were proposed for these results, and we refer to them as the non-representational and the representational view.

According to the *non-representational* view, these developmental findings signal a qualitative shift in mentalistic reasoning during the preschool years. Initially, mentalistic reasoning is non-representational: For young children, beliefs are not yet *representations* that can be inaccurate. Agents may know something about a scene, or they may lack some knowledge about a scene, but they cannot hold false beliefs about a scene. Representational mentalistic reasoning emerges between about 4 and 7 years of age, as a result of significant conceptual, linguistic, and/or executive-function advances—several different accounts have been offered for this transition (e.g., Carlson & Moses, 2001; de Villiers & Pyers, 2002; Gopnik & Wellman, 1994; Perner, 1995).

According to the *representational* view, pioneered by Alan Leslie (1987, 1994a, 2000), children are capable of representational mentalistic reasoning from an early age. This view rests on three key assumptions. First, the ability to understand pretense and the ability to understand false beliefs both involve a *decoupling* mechanism, which makes it possible to hold in mind two distinct versions of a scene: the true and pretend versions in the case of pretense, and the true and

false versions in the case of false beliefs (this decoupling mechanism is distinct from the simpler masking mechanism discussed earlier, which applies to a single version of a scene and merely blocks off those portions that are unavailable to the agent). When adults are tested in the Sally-Anne task, for example, their decoupling mechanism generates a separate version of the scene that incorporates Sally's false belief and enables them to predict where Sally will search for her marble when she returns (e.g., Senju, Southgate, White, & Frith, 2009). To put it another away, adults' decoupling mechanism allows them to temporarily put aside or decouple from their own perspective on the scene in order to adopt Sally's perspective instead (fMRI findings suggest that the medial prefrontal cortex plays a key role in this decoupling process; e.g., Mason & Just, 2009; Tamir & Mitchell, 2010). Second, because 15- to 16-month-olds understand simple pretend scenarios (e.g., Bosco, Friedman, & Leslie, 2006; Onishi, Baillargeon, & Leslie, 2007), it follows that the decoupling mechanism is already operational by the second year of life. Third, preschoolers' failures at elicited-response false-belief tasks must therefore reflect performance limitations, such as inhibition difficulties[11] (e.g., Bloom & German, 2000; Leslie & Polizzi, 1998; Roth & Leslie, 1998; Russell, Mauthner, Sharpe, & Tidswell, 1991). In line with this third assumption, several experiments found that 3-year-olds' performance in elicited-response false-belief tasks improved when inhibition demands were reduced through various means; however, children's performance was often no better than chance (instead of being below chance), providing only weak support for the representational view (e.g., Kovács, 2009; Lewis & Osborne, 1990; Mitchell & Lacohée, 1991; Yazdi, German, Defeyter, & Siegal, 2006).

A critical new research direction began with the discovery that 3-year-olds gave evidence of false-belief understanding in the Sally-Anne task when the experimenter delivered the standard test question (e.g., "Where will Sally look for her marble?") as a self-addressed prompt,

rather than as a direct question: Upon hearing the prompt, children spontaneously looked at the marble's original location, thus correctly anticipating where Sally's false belief would lead her to search (e.g., Clements & Perner, 1994; Ruffman, Garnham, Import, & Connolly, 2001). These positive results suggested that (1) tasks that did not require children to answer direct questions about mistaken agents, or *non-elicited-response* tasks, provided a more sensitive test of early false-belief understanding, and (2) children under age 3 might also demonstrate false-belief understanding if tested with such tasks.

1. Can infants succeed at a non-elicited-response false-belief task?

The first experiment to examine whether infants could attribute a false belief to an agent, using a non-elicited-response task, was conducted by Onishi and Baillargeon (2005). This experiment used a novel VOE task and tested 15-month-olds. The infants received three familiarization trials, one belief-induction trial, and one test trial. At the start of the first familiarization trial, a toy watermelon slice rested on an apparatus floor between two boxes, one yellow and one green; the boxes' openings faced each other and were covered with fringe. An agent (wearing a visor) opened two doors in the back wall of the apparatus, played with the toy briefly, hid it inside the green box, and then paused, with her hand inside the box, until the trial ended. During the second and third familiarization trials, the agent opened the back doors, reached inside the green box (as though to grasp the toy she had previously hidden there), and then paused until the trial ended. During the belief-induction trial, an event occurred that resulted in the agent holding either a true or a false belief about the toy's location; there were four versions of this trial, yielding four conditions. In the *knowledge-green* condition, the upper halves of the back doors were open to create a window; the agent watched through this window as the yellow box moved toward the green box and then returned to its original position (the agent observed no change in the

toy's location and should know it remained in the green box). In the *false-belief-green* condition, the agent was absent—the window remained closed—as the toy moved from the green box into the yellow box (the agent should falsely believe the toy was still in the green box). In the *knowledge-yellow* condition, the agent watched through the window as the toy moved from the green box into the yellow box (the agent should know the toy was now in the yellow box). Finally, in the *false-belief-yellow* condition, the agent watched through the window as the toy moved from the green box into the yellow box; she then closed the window, and the toy returned to the green box (the agent should falsely believe the toy was still in the yellow box). In the test trial, the agent opened the back doors, reached into one of the boxes, and paused until the trial ended. For half the infants in each condition, the agent reached into the green box (green-box event); for the other infants, she reached into the yellow box (yellow-box event).

In each condition, infants expected the agent to act on the information available to her, whether this information was true or false, and they detected a violation when she did not. Thus, the infants in the knowledge-green and false-belief-green conditions expected the agent to reach into the green box, and they looked reliably longer if shown the yellow- as opposed to the green-box event; in contrast, the infants in the knowledge-yellow and false-belief-yellow conditions expected the agent to reach into the yellow box, and they looked reliably longer if shown the green- as opposed to the yellow-box event. These results suggested that 15-month-olds already realize that an agent can hold and act on a false belief, supporting the representational view.

2. Additional evidence of early false-belief understanding

Since the publication of Onishi and Baillargeon (2005), over 15 reports—all using non-elicited-response tasks—have provided additional evidence of false-belief understanding in infants in the second year of life and in toddlers in the third year of life. These reports can be divided into

two groups: those using spontaneous-response tasks and those using prompted-action tasks.

Spontaneous-response false-belief tasks. In spontaneous-response tasks, an agent holds a false belief about some aspect of a scene, and investigators measure children's spontaneous responses to the unfolding scene. The spontaneous-response tasks that have been used to date vary along several dimensions, summarized below (an exhaustive review of these tasks is beyond the scope of this chapter; see Baillargeon, Scott, & He, 2010, for a review).

First, spontaneous-response tasks vary in the *measure* they use. In addition to VOE tasks (e.g., Song, Onishi, Baillargeon, & Fisher, 2008; Träuble, Marinović, & Pauen, 2010; Yott & Poulin-Dubois, 2012), spontaneous-response tasks include: anticipatory-looking tasks, which record where children look as they anticipate which location a mistaken agent will approach to find a goal object (e.g., He, Bolz, & Baillargeon, 2012; Senju, Southgate, Snape, Leonard, & Csibra, 2011; Southgate, Senju, & Csibra, 2007; Thoermer, Sodian, Vuori, Perst, & Kristen, 2012); anticipatory-pointing tasks, which record whether children point to inform a mistaken agent that her goal object has been moved to a new location (e.g., Knudsen & Liszkowski, 2012a) or that an aversive object has been placed at the location she falsely believes holds her goal object (e.g., Knudsen & Liszkowski, 2012b); and preferential-looking tasks, which measure whether children look preferentially at a matching over a non-matching picture as they listen to a story about a mistaken agent who is searching for a goal object (the matching picture depicts the agent searching the object's initial location, and the non-matching picture depicts the agent searching the object's current location; Scott, He, Baillargeon, & Cummins, 2012).

Second, spontaneous-response tasks vary in the *scene* they present. Although many tasks have used a scene in which a human agent searches for a goal object and holds a false belief about its location, as in the Sally-Anne task, a few tasks have used a non-human agent (e.g.,

Meristo et al., 2012; Surian, Caldi, & Sperber, 2007; Surian & Geraci, 2012), and other tasks have varied the counterfactual state the agent holds about the scene: There is now evidence that older infants and toddlers can represent not only false beliefs about location but also false perceptions (Song & Baillargeon, 2008), false beliefs about identity (Scott & Baillargeon, 2009; Scott, Richman, & Baillargeon, 2013), false beliefs about expected contents (He, Bolz, & Baillargeon, 2011b), and false beliefs about non-obvious properties (Scott, Baillargeon, Song, & Leslie, 2010). This last task also varied the goal pursued by the agent in the scene. Instead of searching for an object, the agent watched an experimenter demonstrate that a target object rattled when shaken, and then selected one of two test objects to produce the same effect. One test object was identical to the target object, and the other test object differed from it in color and pattern; infants (but not the agent) knew that only the different object rattled when shaken. The infants expected the agent to falsely assume that the identical object would rattle, on the general expectation that perceptually identical objects typically share non-obvious properties.

Third, spontaneous-response tasks vary in their *linguistic requirements*. Although most tasks have been entirely non-verbal, a few have used simple language (e.g., He et al., 2011b; Scott et al., 2010; Song et al., 2008), and others—given only to toddlers—have been highly verbal, with linguistic demands comparable to those of elicited-response tasks (He et al., 2012; Scott et al., 2012). For example, 2.5-year-olds in a VOE task watched a typical Sally-Anne scene along with an adult "subject" who was then asked where Sally would look for her toy when she returned. The children looked reliably longer when the "subject" responded incorrectly and pointed to the container that currently held the toy.

The various spontaneous-response false-belief tasks cited above have yielded positive results with Western children between 13-15 months (e.g., Song & Baillargeon, 2008; Träuble et

al., 2010; Surian et al., 2007) and 32 months (e.g., He et al., 2011b, 2012; Scott et al., 2012).

Recently, positive results have also been obtained with 23- to 40-month-olds from three

traditional non-Western societies: a Salar community in northwest China, a predominantly Shuar

community in southeastern Ecuador, and a Yasawan community in northwest Fiji (Barrett et al.,

2013). Together, these results provide robust evidence that infants and toddlers can represent and

reason about false beliefs, supporting the representational view of early psychological reasoning.

Prompted-action false-belief tasks. Prompted-action tasks represent a hybrid between

elicited- and spontaneous-response tasks. As in elicited-response tasks, children are given a

verbal prompt; however, this prompt only indirectly taps their representation of the agent's false

belief (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Southgate, Chevallier, & Csibra, 2010).

In the task of Buttelmann et al. (2009), for example, an experimenter showed 18-month-olds two

lidded boxes and demonstrated how to lock and unlock them; the boxes were left unlocked.

Next, a male agent entered the room, hid a toy in one of the boxes, and then left. While he was

gone, the experimenter moved the toy to the other box and locked both boxes. When the agent

returned, he tried to open the box where he had hidden the toy, without success, and then sat

centered behind the boxes. When prompted to help the agent (e.g., "Go on, help him!"), most

infants approached the other box (i.e., the one the agent did not act on), suggesting that they

realized the agent wanted to retrieve his toy and falsely believed it was still in its original

location. To respond correctly when prompted, the infants only needed to consider the agent's

goal; they did not have to directly tap their representation of his false belief, as they would have

if asked which box he would approach to retrieve his toy. So far, positive results have been

obtained in prompted-action tasks with 17- to 18-month-olds, providing converging evidence

that infants in the second year of life are capable of representational mentalistic reasoning.

3. Are young infants also capable of false-belief understanding?

Can infants in the first year of life also attribute false beliefs to others? Evidence that they do not would suggest that (1) the decoupling mechanism that enables infants to understand pretense, false beliefs, and other counterfactual states does not become operational until the second year of life, and (2) there is an important shift after children reach their first birthday from non-representational to representational mentalistic reasoning. On the other hand, evidence that infants in the first year of life can already attribute false beliefs would cast doubt on the existence of such a shift (or at least suggest that it occurs earlier).

To date, there have been three reports of false-belief understanding in the first year of life, with infants ages 7 to 11 months (He & Baillargeon, 2013; Kovács, Téglás, & Endress, 2010; Luo, 2011a). To illustrate, consider a VOE experiment conducted with 10- to 11-month-olds (He & Baillargeon, 2013). The infants in the *false-belief* condition first received four familiarization trials. In each trial, a female agent sat at a window in the back wall of an apparatus, and a female experimenter knelt at a window in the right wall; on the apparatus floor were two short open containers and a tall toy dog. The agent played with the dog briefly, returned it to the apparatus floor, and then hid herself by lifting a large cloth that filled her window. The experimenter placed the dog in one of the containers and then signaled the agent to return ("Ok!"). At that point, the agent lowered her cloth, grasped the dog's head, and paused until the trial ended. Across trials, different containers were used, and the dog was placed in the left or right container (order was counterbalanced). The familiarization trials thus served to establish that the agent wanted the dog and reached for it wherever the experimenter happened to place it. Next, the infants received two test trials involving a tall and a short container, each closed with a lid; the dog was taller than the short but not the tall container. As before, the agent

played with the dog and then hid behind her cloth. Next, the experimenter silently shortened the dog (its body could be collapsed by pressing firmly on its head), placed it in the short container, and then closed the container. When the agent returned, she grasped the lid of either the tall (tall-container event) or the short (short-container event) container and then paused until the trial ended. The infants looked reliably longer at the short- than at the tall-container event, suggesting that they expected the agent (1) to falsely believe that the dog was still tall and hence (2) to falsely infer that the dog was hidden in the tall container (infants as young as 7.5 months realize that a tall object cannot be hidden inside a short container; e.g., Hespos & Baillargeon, 2006). These results (which were replicated with a second group of infants) provided evidence that infants in the first year of life can already reason about false beliefs.

Support for this interpretation came from two additional conditions. In the *knowledge* condition, the agent watched all of the experimenter's actions through a large hole in her cloth. In the test trials, the infants expected the agent to reach for the short container, where she had seen the experimenter place the shortened dog, and they detected a violation when she reached for the tall container instead. The *ignorance* condition was identical to the false-belief condition except that, before the agent hid behind her cloth in the test events, she saw the experimenter shorten the dog. Because the shortened dog could be hidden in either container, the agent had no basis for determining its location, and the infants looked about equally whether she reached for the tall or the short container.

In a final *tall-block* condition, infants saw events identical to those in the false-belief condition, with one exception: In a pretest trial administered prior to the test trials, the experimenter demonstrated that the tall container was in fact a tall block topped with a lid (control data indicated that infants realized that no object could be inserted into this block, in

accordance with the solidity principle). The infants looked reliably longer at the short-container than at the tall-block event, suggesting that they expected the agent (1) to falsely believe that the dog was still tall, (2) to falsely assume that the tall block was a tall container, and hence (3) to falsely infer that the dog was hidden in this tall container. The infants were thus able to represent *three* distinct false beliefs.

The results described above, together with those of Kovács et al. (2010) and Luo (2011a), support two conclusions. First, false-belief understanding is already present in the first year of life, supporting the view that representational mentalistic reasoning emerges early in infancy. Second, early false-belief understanding is already quite sophisticated: The infants tested by He and Baillargeon (2013) attributed to the agent a causally coherent set of motivational, epistemic, and counterfactual mental states that included as many as two (false-belief condition) or three (tall-block condition) interlocking false beliefs.

1-V. WHY DO YOUNG CHILDREN FAIL AT ELICITED-RESPONSE FALSE-BELIEF TASKS?

In section 1-IV, we reviewed extensive evidence, from various non-elicited-response tasks, that children in the first few years of life can already represent false beliefs. Although the focus of this chapter is on infancy, in this section we take a moment to discuss the glaring question raised by these findings: If young children are capable of false-belief understanding, why do they consistently fail at standard elicited-response tasks (e.g., Callaghan et al., 2005; Liu et al., 2008; Wellman et al., 2001)? This question is the topic of intense debate (for example, see the special issue (volume 30, issue 1) published in 2012 by the *British Journal of Developmental Psychology*). As a full review of this debate is beyond the scope of this chapter, we limit our discussion to two relevant accounts: the processing-load and the explicit-reasoning account. The

processing-load account assumes that non-elicited- and elicited-response tasks tap the *same* form of false-belief understanding and looks elsewhere for an explanation of young children's difficulties with elicited-response tasks. In contrast, the explicit-reasoning account assumes that elicited-response tasks tap a *different* form of false-belief understanding than do non-elicited-response tasks.

1. Processing-load account

According to the processing-load account, elicited-response tasks are challenging because, in addition to false-belief understanding, they involve executive-function processes (in terms of the terminology introduced in Footnote 10, the processing-load account is an expression as opposed to an emergence account). Below, we describe our own version of this account (e.g., He et al., 2012; Scott & Baillargeon, 2009), which builds on prior versions by other researchers.

Response-selection and inhibition. We assume that at least two executive-function processes, response selection and inhibition, play critical roles in elicited-response false-belief tasks. When children are asked the test question in a standard elicited-response task, a *response-selection* process is activated: Children must interpret the test question, choose to answer it, and generate a response (e.g., Scott & Baillargeon, 2009; see also Mueller, Brass, Waszak, & Prinz, 2007; Saxe, Schulz, & Jiang, 2006). Under some conditions, the response-selection process inadvertently leads to a prepotent but context-inappropriate response. In the Sally-Anne task, for example, the test question "Where will Sally look for her marble?" triggers a "reality bias": Because agents usually look for an object where it is, a prepotent response is that Sally will look for the marble in its current location. Thus, instead of—or in addition to—tapping their representation of Sally's false belief, children tap their own knowledge about the marble's current location. As a result, children cannot succeed at the task unless their *inhibition* skills are

sufficiently mature to allow them to suppress the prepotent but context-inappropriate response generated by the reality bias (e.g., Birch & Bloom, 2003; Carlson & Moses, 2001; Hala, Hug, & Henderson, 2003; Leslie & Polizzi, 1998; Mitchell & Lacohée, 1991; Russell et al., 1991).

The preceding analysis allows us to readily make sense of the positive results of the non-elicited-response tasks reported earlier. For example, in spontaneous-response tasks where the question "Where will Sally look for her marble?" is delivered as a self-addressed prompt (He et al., 2012) or is addressed to an adult "subject" (Scott et al., 2012), the response-selection process is not activated, leaving children free to tap their representation of Sally's false belief. In prompted-action tasks, the response-selection process is activated, but the prompts do not trigger prepotent responses (e.g., "Go on, help him!"; Buttelmann et al., 2009), allowing children to respond based on their representation of the scene, which includes the agent's false belief.

Young children experience similar response-selection/inhibition difficulties with other prepotent cognitive biases, such as the "gravity bias" (Hood, 1995). Lee and Kuhlmeier (2013) examined 26-month-olds' spontaneous and elicited responses in a one-tube task adapted from Hood (1995). Children were first familiarized with a vertical frame that had three openings at the top and three opaque cups at the bottom, one below each opening. Next, a flexible opaque tube was attached to an opening and to a non-aligned cup, creating an S-shaped configuration; an experimenter dropped a ball into the opening and then encouraged children to retrieve the ball from one of the cups. Following familiarization, children watched videotaped events in which a hand dropped the ball into one of the openings, with the tube arranged into various S-shaped configurations; two seconds after the ball was dropped, children were asked "Where is the ball?" The authors measured both where children looked immediately after the ball was dropped (spontaneous response) and where children pointed following the test question (elicited

response). Strikingly, one group of toddlers looked spontaneously at the correct cup, but tended to point to the cup directly below the opening where the ball was dropped. These toddlers clearly understood that the opaque tube constrained the path of the ball, as indicated by their looking responses; when they were asked the test question, however, the response-selection process led them to first consider where the ball had been dropped, and this in turn triggered a prepotent gravity bias: Because an object typically falls in a straight line, they tended to point to the cup directly below the opening used by the hand. Only children who had sufficient inhibitory skills to inhibit this prepotent response succeeded at the task.

Thus, it appears that (1) response-selection processes can inadvertently trigger cognitive biases that mask young children's underlying competences, (2) these biases are not limited to the domain of psychological reasoning, and (3) different biases are overcome at different ages, no doubt as relevant everyday experiences gradually help strengthen inhibitory skills, making it easier to suppress context-inappropriate prepotent responses.

Working-memory limitations. If young children's difficulties with elicited-response false belief tasks were entirely due to their inability to inhibit prepotent responses generated by the reality bias, we might expect them to succeed at elicited-response tasks designed to circumvent this bias. One such task is the "undisclosed-location" task: Instead of transferring Sally's marble from the basket to the box, for example, Anne takes it away to an undisclosed location. Since children do not know where the marble is, we might expect the reality bias to have little or no effect, leaving them free to answer the test question based on their representation of Sally's false belief. This is not the case, however: Young children typically perform at chance in standard undisclosed-location tasks (e.g., Bartsch, 1996; Wellman et al., 2001; Wimmer & Perner, 1983), suggesting that additional difficulties undermine their performance. We have suggested that some

of these difficulties involve *working-memory* limitations. When children are asked the test question, the response-selection process is activated; although in undisclosed-location tasks no prepotent response is generated, children must still interpret the test question and produce a response. The joint demands created by the false-belief-representation and response-selection processes tend to overwhelm young children's limited information-processing resources, leading to chance performance. In other words, young children cannot easily hold in mind the agent's false belief *and* answer a direct question about this belief.

The preceding analysis predicts that, if response-selection demands could be reduced in an undisclosed-location task (thus lessening the overall demands of the task), young children might succeed in tapping their representation of the agent's false belief. Recent findings support this prediction: 2.5-years-olds succeeded at an elicited-response undisclosed-location task as long as they first received two response-selection practice trials highly similar in form to the test trial (Setoh, Scott, & Baillargeon, 2011). In these experiments, children heard a story accompanied by a picture-book; in each of six story trials, an experimenter turned a page of the picture-book to reveal a new picture and recited a line of the story. Briefly, the story introduced Emma who found an apple in one of two containers (e.g., a box), moved it to the other container (e.g., a bowl), and then went outside to play with her ball; in her absence, her brother Ethan found the apple and took it away; finally, Emma returned to look for her apple. In the test trial, children saw two pictures depicting the box and the bowl, and they were asked where Emma would look for her apple. Interspersed among the story trials were two response-selection practice trials: In one, children saw two pictures depicting an apple and a banana, and they were asked where was Emma's apple; in the other, children saw two pictures depicting a ball and a frisbee, and they were asked where was Emma's ball (in each case, children were required to point to the appropriate picture). Children

performed reliably above chance in the test trial, pointing to the container Emma falsely believed held her apple. Additional experiments confirmed these results and also revealed that children failed if they received only one practice trial (Setoh et al., 2011) or if the practice questions ("Which one is Emma's apple?") differed in form from the test question ("Where will Emma look for her apple?") (Scott & Setoh, 2012). Together, these results indicate that (1) young children' working-memory or information-processing resources can be overwhelmed by the joint demands of representing an agent's false belief and answering a test question about this belief, even in an undisclosed-location task, and (2) when response-selection demands are reduced by simple practice trials, children succeed in tapping their representation of the agent's false belief to answer the test question correctly.

2. The explicit-reasoning account

In psychological research, the term *implicit* has several distinct meanings. One such meaning (akin to *sham*) refers to processes that mimic more advanced, explicit processes but do not, in fact, involve the same concepts and computations. For example, if an infant who saw three objects being placed in a box represented that "an object and another object and another object" were in the box, we would grant the infant an implicit, but not an explicit, understanding of the concept "three"; this last concept would not figure in any way in her reasoning about the event. Some researchers have recently argued that all of psychological reasoning prior to about age 4—the age at which children begin to succeed at standard elicited-response false-belief tasks—is implicit in this first (sham) sense (e.g., Perner, 2010; Ruffman et al., 2012). In this view, young children are "mindblind" and incapable of attributing *any* mental states to others; instead, they rely on statistical rules, derived from everyday experience, to predict or respond to others' actions. These statistical rules (1) describe how agents tend to act in specific situations

(e.g., agents who are searching for an object typically search for it where they saw it last) and as such (2) "provide an implicit theory of mind [in that] they capture the workings of the mind without mentioning the mind" (Perner, 2010, p. 259). Later in childhood, through learning processes that remain to be elucidated, children begin to explicitly understand agents' actions in terms of underlying mental states, as evidenced by their correct responses and justifications in standard elicited-response false-belief tasks. Such a view, we believe, is not supported by the extensive evidence reviewed in this chapter that infants attribute mental states to agents (for related arguments, see Carruthers, 2013; Jacob, in press).

Another meaning of the term implicit (akin to *procedural*) refers to knowledge that has not yet been transformed into more advanced knowledge that is general and accessible and that can be verbally described—in other words, explicit knowledge (e.g., Karmiloff-Smith, 1992; Low, 2010). In line with this distinction, Low (2010) proposed that, whereas implicit false-belief understanding is sufficient for success in non-elicited-response tasks, explicit false-belief understanding is necessary for success at elicited-response tasks. In a series of experiments, Low administered a battery of tasks to 3- and 4-year-olds and obtained two main findings. First, performance in a non-elicited-response false-belief task (a verbal anticipatory-looking task) was related to performance in various elicited-response false-belief tasks (for similar longitudinal findings, see Thoermer et al., 2012). Second, performance in a language task and an executive-control task was related to performance in the elicited- but not the non-elicited-response false-belief tasks. Low speculated that, with the development of language and executive control, implicit false-belief understanding gradually metamorphoses "into a higher-order format that supports conscious and verbally correct false-belief judgments...[with] reiterated rounds of internal redescription of implicit knowledge yielding stronger representations of explicit

knowledge" (p. 612).

How might this redescription occur? Low (2010) suggested that, early in life, infants occasionally encounter situations where an agent acts in an unexpected way (e.g., searches for an object in the wrong place), and they correctly infer that the agent is acting on a false belief (they may then reveal this implicit understanding in spontaneous anticipatory glances, pointing gestures, helping actions, and so on). However, because such inferences are uncommon—agents do not routinely act on false beliefs—this early false-belief understanding remains for several years local and piecemeal. Over time, these isolated inferences continue to accumulate and are eventually integrated into a general and explicit format that supports correct responses in elicited-response false-belief tasks. In this view, then, children succeed at elicited-response tasks when false-belief understanding becomes a part of their "folk" theory of mind—in other words, when they become explicitly aware that individuals sometimes hold false beliefs about the world.

3. Is explicit false-belief understanding necessary for success at elicited-response tasks?

Does the processing-load account, the explicit-reasoning account, or some other account best explain young children's difficulties with elicited-response false-belief tasks? This issue remains highly controversial (e.g., Butterfill & Apperly, in press; Carruthers, 2013; Jacob, in press; Low & Watts, 2013; Perner & Roessler, 2012; Scott et al., 2013), and additional research is needed to settle it. For what it's worth, our intuition is that developing an explicit understanding of false belief is unlikely to be necessary for success at elicited-response tasks, for two reasons. First, if limited executive-function skills already explain young children's failure at these tasks, why *also* assume that this failure reflects a lack of explicit false-belief understanding? Why posit two explanations when one suffices? Second, it is unclear how an explicit-reasoning account could explain the evidence that reducing executive-function demands

leads to significant improvements in performance. For example, recall that 2.5-year-olds succeeded at an elicited-response undisclosed-location task after receiving response-selection practice trials (e.g., "Where is Emma's apple?"; Setoh et al., 2011); it is highly unlikely that these trials induced an explicit understanding of false belief. Note that we are not questioning whether children eventually develop a folk theory of mind: There is little doubt that they do (e.g., Gopnik & Wellman, 1994; Leslie, 2000), just as they develop folk theories in other domains such as biology and astronomy (e.g., Carey, 1985; Vosniadou & Brewer, 1992). Rather, what we are questioning is the notion that developing an explicit understanding of false belief as part of one's folk theory of mind is necessary for success at elicited-response false-belief tasks.

1-VI. SUMMARY

The evidence reviewed in the first part of this chapter suggests four main conclusions. First, infants are equipped with a psychological-reasoning system that provides them with a skeletal causal framework for reasoning and learning about the intentional actions of human and non-human agents. Second, a key part of this framework is the principle of rationality, with its corollaries of consistency and efficiency. Third, early psychological reasoning is mentalistic and representational in nature: Even infants in the first year of life can attribute to agents motivational states (goals, dispositions), epistemic states (knowledge, ignorance), and counterfactual states (false beliefs, pretend beliefs). Fourth, whether young children succeed at a given psychological-reasoning task depends on the total demands of the task; standard elicited-response false-belief tasks, which involve false-belief reasoning as well as inhibition and other executive-function processes, are particularly challenging for young children.

Together, these various points suggest that psychological reasoning is an evolved adaptation that is present early in life and that enables humans to make sense of, and to rapidly

learn about, the actions of fellow humans and of non-human animals.

PART 2: EARLY SOCIOMORAL REASONING

As it became clear that infants could make sense, at least in simple situations, of the intentional actions of a single agent in a scene, researchers were naturally led to ask whether infants might also possess expectations about social interactions among two or more agents. These questions led to a new focus on early sociomoral expectations. As we discuss this work in Part 2, readers will notice that many findings provide converging evidence for conclusions reached in Part 1 (e.g., infants can attribute goals to non-human agents). To avoid repetition, however, we will focus mainly on new conclusions concerning early sociomoral reasoning.

For adults, sociomoral expectations do not apply to agents in general, but rather to a subset of agents we will refer to as *individuals*. Agents, as we saw, are entities who detect their environment and produce intentional actions; individuals are agents who can have social interactions with other agents (e.g., Jackendoff, 2007). Henceforth, we thus frame our discussion in terms of infants' expectations about how individuals should act toward each other.

Two traditional approaches have dominated the study of the development of morality: One approach focuses on how parental practices and other socialization processes help children internalize and conform to moral norms, whereas the other approach adopts a more cognitive stance and focuses on how children gradually construct moral concepts and norms through their interactions with peers and other individuals (for reviews, see Eisenberg, Fabes, & Spinrad, 2006; Turiel, 2006). Despite their marked differences, both approaches generally assume that sensitivity to moral norms does not emerge until the preschool years. This assumption has been called into question by a new approach that springs from multiple disciplines within cognitive science—including anthropology, economics, linguistics, philosophy, and psychology—and

views morality as an *evolved adaptation* (e.g., Brewer, 1999; Dupoux & Jacob, 2007; Dwyer, 2006; Fiske, 2004; Gintis, Henrich, Bowles, Boyd, & Fehr 2008; Greene, 2005; Haidt, 2001; Jackendoff, 2007; Mikhail, 2007; Premack, 2007; Sigmund, Fehr, & Novak, 2002; Sripada & Stich, 2006). Among proponents of this approach, there is a great deal of disagreement about how best to characterize our human capacity for intuitive sociomoral reasoning. For example, this capacity has been variously ascribed to emotional intuitions (e.g., Haidt, 2001), relational models (e.g., Fiske, 2004), norm-detection mechanisms (e.g., Sripada & Stich, 2006), and sociomoral principles (e.g., Premack, 2007). In this chapter, we adopt this last, principled-based conception of sociomoral reasoning. This conception is, of course, consistent with the principled-based conceptions of physical and psychological reasoning discussed earlier in this chapter.

The principle-based conception of sociomoral reasoning rests on several key assumptions (e.g., Brewer, 1999; Dupoux & Jacob, 2007; Dwyer, 2009; Haidt, 2008; Jackendoff, 2007; Premack & Premack, 2003; Sigmund et al., 2002). First, humans are born with a set of abstract sociomoral principles that evolved during the millions of years our ancestors lived in small groups of hunter-gatherers, where survival depended on cooperation. Second, the principles determine not what is virtuous but rather what is obligatory, permissible, and forbidden in social interactions. As such, the principles have far-ranging applications: They affect expectations about how individuals are likely to act, interpretations of why individuals acted as they did, evaluations of individuals and their actions, and so on. Third, the principles are elaborated and rank-ordered in different ways by different cultures, resulting in the diverse moral landscape that exists in the world today.

How many principles guide human sociomoral reasoning? How should each principle be defined? How do the principles interact? Infancy research can make a unique contribution to the

ongoing debate concerning these questions, because infants' sociomoral expectations are less affected by experience and as such closer to our human "starting state". Uncovering infants' expectations about how individuals should act toward each other will thus shed significant light on the basic cognitive architecture that underlies intuitive sociomoral reasoning.

To date, infancy researchers have begun to explore four candidate sociomoral principles: *reciprocity*, *fairness*, *ingroup*, and *authority* (findings about a fifth principle, *no-harm*, are too preliminary to discuss). To investigate these principles, researchers typically use either first- or third-party tasks. In *first-party* tasks, infants participate in social interactions with others, and various spontaneous and elicited responses are measured. In *third-party* tasks, infants observe others interact (but do not participate in these interactions themselves), and their spontaneous responses are measured. As will become clear, findings from the two types of tasks are not always consistent, largely because self-interest tends to prevail in first-party tasks. These findings underscore the importance of the socialization processes that help children gradually overcome their self-interest and align their behavior with their sociomoral expectations (e.g., Eisenberg et al., 2006; Premack, 2007).

2-I. DO INFANTS ASSIGN A VALUE TO SOCIAL ACTIONS?

Many sociomoral expectations presuppose a fundamental ability to assess the *value* of actions directed at others. The value of an action is typically defined in terms of its *valence* (positive, negative, or neutral) and its *magnitude* (e.g., Jackendoff, 2007; Premack, 1990). As might be expected, positive actions are those that have a beneficial effect on others (e.g., helping, comforting, sharing), whereas negative actions are those that have a detrimental effect (e.g., hitting, hindering, stealing). Our review of early sociomoral reasoning thus begins with experiments that explored infants' ability to discriminate between positive and negative actions.

1. Do infants distinguish between positive and negative actions?

     Third-party findings. In a seminal experiment, Premack and Premack (1997) habituated 12-month-olds to one of four computer-animated events involving interactions between two circles, one gray and one black. In the *caress* event, the gray circle gently rubbed the top of the black circle; in the *hit-vertical* event, the grey circle hit the top of the black circle, causing it to momentarily deform; in the *help* event, after watching the black circle attempt unsuccessfully to reach a high opening in a vertical barrier, the gray circle pushed the black circle through the opening; finally, in the *hinder* event, after passing through the opening in the barrier, the gray circle prevented the black circle from doing the same by pushing it away from the opening. Within each habituation trial, the event was repeated continuously until the trial ended. Following habituation, all infants received three identical test trials involving a novel *hit-horizontal* event: The gray circle hit the left side of the black circle, causing it to momentarily deform. During the habituation trials, looking times at the four events declined across trials, with no reliable differences among the events. During the test trials, however, the infants habituated to the caress and help events showed greater recovery in looking time than did those habituated to the hit-vertical and hinder events. The infants thus (1) viewed caressing and helping as positive actions and hitting and hindering as negative actions, and (2) dishabituated when they detected a valence change (from positive to negative) in the actions directed at the black circle.

     These results suggested three broad conclusions. First, by one year of age, infants already assign valences to actions toward others. Second, infants can determine whether actions are beneficial or detrimental along multiple dimensions: For example, hitting and hindering were both viewed as negative actions, despite their marked superficial differences. Finally, infants do not generally expect individuals to act positively—or to refrain from acting negatively—toward

others: The infants tended to look equally during the habituation trials whether the gray circle acted positively or negatively toward the black circle.

First-party findings. Additional evidence that infants discriminate between positive and negative actions came from experiments in which 9-, 12, and 18-month-olds interacted with a female experimenter who sat across from them at a table and was at times unwilling and at times unable to give them a toy (Behne, Carpenter, Call, & Tomasello, 2005). For example, in *teasing* trials, the experimenter offered the infant a ball but teasingly pulled it back when the infant reached for it; in *clumsy* trials, she attempted to give the infant a ball but clumsily dropped it in such a way that it rolled back toward her. Detailed analyses of the infants' reactions suggested that they detected the experimenter's positive or negative intentions. At all ages, the infants reached reliably more for the ball in the teasing than in the clumsy trials. Moreover, the 9-month-olds banged on the table reliably more in the teasing trials, and the 18-month-olds looked away from the experimenter reliably more in the teasing trials. The authors concluded that the infants responded with impatience or frustration when the experimenter was unwilling to give them the ball, but not when she was merely unable to do so.

2. Do infants prefer individuals who produce positive actions?

The evidence reviewed in the last section indicated that infants aged 9 months and older distinguish positive and negative actions produced by the *same* individual. In ground-breaking experiments, Kuhlmeier, Hamlin, Wynn, and Bloom created scenes where *one* individual acted positively and a *different* individual acted negatively, in order to examine whether infants would prefer, and would expect others to prefer, the individual who acted positively (e.g., Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007, 2010; Hamlin, Wynn, Bloom, & Mahajan, 2011; Kuhlmeier, Wynn, & Bloom, 2003).

Choice and VOE tasks. In one experiment, for example, Hamlin et al. (2007) habituated 6- and 10-month-olds to live events involving a steep irregular hill, which rose from the lower right to the upper left of the apparatus, and three different wooden blocks with prominent eyes, which served as the climber, helper, and hinderer. The infants saw a help and a hinder event on alternate habituation trials. In the *help* event, the climber climbed to the first plateau in the hill, "danced" briefly, and then attempted unsuccessfully to climb to the plateau at the top of the hill; on its third attempt, the climber was aided by the helper, who entered the apparatus at the bottom of the hill, moved up the hill, and bumped the climber twice until it finally reached the top plateau. While the helper moved back down the hill and exited the apparatus, the climber "danced" briefly on the top plateau and then paused until the trial ended. The *hinder* event was similar except that the hinderer entered the apparatus at the top of the hill, moved down the hill, and bumped the climber twice, causing it to tumble down the hill end-over-end. The hinderer moved back up the hill and exited the apparatus, and the climber paused at the bottom of the hill until the trial ended. In line with the findings reported by Premack and Premack (1997), the infants looked about equally at the final paused scenes of the help and hinder events (Hamlin, pers. comm., 08/14/2012; see also Hamlin et al., 2011).

Following habituation, the infants received two tasks in counterbalanced order: a *choice* task that measured whether they preferred the helper over the hinderer, and a *VOE* task that assessed whether they expected the climber to show the same preference. In the choice task, the infants were presented with the helper and the hinderer on a board and were encouraged to choose one ("Would you like to pick a toy?"). At both ages, the infants were reliably more likely to touch the helper first, suggesting that they preferred it over the hinderer. This preference was eliminated in a control experiment in which the climber was replaced with an inanimate object

(without eyes or self-propelled motion) that was simply pushed up or down the hill. Similar preferences were observed with 9-, 5-, and even 3-month-olds (using a preferential-looking measure) with various positive and negative actions (e.g., helping a puppet retrieve a toy from a box versus hindering the puppet by slamming the box shut; returning a ball to a puppet who had dropped it versus stealing the ball; Hamlin et al., 2010; Hamlin & Wynn, 2011).[12]

In the VOE task, the steep hill was replaced with a shallow symmetrical hill; the climber stood on the top of the hill, and the helper and hinderer stood on either side of the hill. In the test events, the climber approached and stopped next to the helper (*approach-helper* event) or the hinderer (*approach-hinderer* event) and then paused until the trial ended. At 6 months, infants tended to look equally at the two events (this negative result could mean either that infants this age do not yet expect others to share their preference for individuals who produce positive actions, or that this preference is still weak and can only be revealed by a forced-choice measure). At 10 months, however, infants looked reliably longer at the approach-hinderer than at the approach-helper event, suggesting that they expected the climber to share their preference for the helper over the hinderer. Using an anticipatory-looking measure, Fawcett and Liszkowski (2012) found that 12-month-olds anticipated that the climber would approach the helper as opposed to the hinderer, providing converging evidence that older infants expect individuals to prefer helpers over hinderers.[13]

Together, the preceding results were important for three reasons. First, they provided additional evidence that infants distinguish between positive and negative actions.[14] Second, the results provided further evidence that infants do not generally expect individuals to act positively toward others: Across experiments, infants tended to look equally at the positive and negative actions shown in the habituation trials. Third, and most critically, the results indicated that

infants (1) kept track of which individual acted positively and which individual acted negatively in the habituation trials and (2) preferred the former over the latter when asked to choose between them. What was the basis of this preference? How did infants evaluate the two individuals? We next consider these questions.

Sociomoral versus affiliative evaluations. When it comes to social evaluations, it is helpful to distinguish between *sociomoral* evaluations, which are based on whether individuals act in accordance with sociomoral principles, and *affiliative* evaluations, which are based on whether individuals possess properties one deems desirable in affiliates. As might be expected, similarity in appearance and attitudes plays a large role in affiliative evaluations (birds of a feather flock together; e.g., Byrne, 1971; Sunnafrank, 1983), but so do more arbitrary or accidental properties. To illustrate, 3- to 6-year-olds prefer children of their own gender when asked with whom they want to be friends (e.g., Shutts, Pemberton Roben, & Spelke, 2013); they prefer children who have experienced a lucky event (e.g., found money on a sidewalk) over children who have experienced an unlucky event (e.g., gotten splashed by a passing car) (e.g., Olson, Dunham, Dweck, Spelke, & Banaji, 2008); and, in a phenomenon called *evaluative contagion*, they prefer associates of lucky children over those of unlucky children (e.g., Olson, Banaji, Dweck, & Spelke, 2006).

Keeping in mind the distinction between sociomoral and affiliative evaluations, let us return to the findings reviewed in the last section. Because the infants tended to look equally at the help and hinder events presented during the habituation trials, it seems unlikely that their evaluations of the helper and hinderer were based on sociomoral grounds. Had helping been perceived as obligatory, or hindering been perceived as forbidden, the infants would have looked reliably longer at the hinder event (we will see later that infants do hold sociomoral expectations

about help and hinder events in some contexts). Thus, the infants most likely preferred the helper over the hinderer on affiliative grounds; a tendency toward helping behavior is a highly desirable property in an affiliate, whereas a tendency toward hindering behavior is not.

Other findings support the notion that infants readily form and act on affiliative evaluations. In experiments by Mahajan and Wynn (2012), 11.5-month-olds first chose between crackers and green beans (snacks experiment) or between orange and yellow mittens (mittens experiment). Next, two animated puppets indicated their own preferences; the *similar* puppet expressed a liking for the infant's choice and a dislike for the other choice, and the *dissimilar* puppet expressed the opposite attitudes. Finally, the infants were offered the two puppets and encouraged to choose one. In both experiments, the infants reliably preferred the similar puppet, suggesting that they had formed similarity-based affiliative evaluations of the puppets.

Yet other findings indicate that, for infants as for older children and adults, affiliative preferences can also arise through evaluative contagion. In additional snacks experiments (Hamlin, Mahajan, Liberman, & Wynn, 2013), 9- and 14-month-olds chose a snack, watched as the puppets indicated their preferences, and then saw scenarios in which new characters acted toward one of the puppets either positively (i.e., returned a ball the puppet had dropped) or negatively (i.e., stole a ball the puppet had dropped). Infants who saw actions directed at the similar puppet preferred the character who returned the puppet's ball over the character who stole the puppet's ball; infants who saw actions directed at the dissimilar puppet showed the opposite preference. Likewise, Hamlin et al. (2011) first showed 8- and 19-month-olds help and hinder scenarios, followed by return-ball and steal-ball scenarios in which new characters acted toward the helper or the hinderer (these new characters were absent during the initial scenarios, so that sociomoral considerations of third-party retaliation or justice were unlikely). Infants who

saw actions directed at the helper preferred the character who returned the ball, whereas infants who saw actions directed at the hinderer preferred the character who stole the ball. Together, these results suggest that, even for infants, affiliative evaluations (whatever their bases) are highly contagious and spread in predictable ways: A friend of a friend is a friend, but so is an enemy of an enemy.

2-II. DO INFANTS SHOW SENSITIVITY TO THE RECIPROCITY PRINCIPLE?

Researchers from diverse fields—including anthropology, economics, linguistics, psychology, and sociology—have argued that reciprocity is one of the fundamental principles guiding human social interactions (e.g., Fehr & Gachter, 2000; Fiske, 1992; Fry, 2006; Gouldner, 1960; Jackendoff, 2007; Premack, 1990; Trivers, 2006). Although researchers interested in reciprocity often focus on its role in the development of stable long-term social relations, here we are concerned mainly with the effects of reciprocity in short-term interactions.

The principle of reciprocity specifies how an individual should act *in response* to another individual's action: If a first individual acts in some way toward a second individual, who chooses to respond, then the second individual's action should match that of the first individual in value (e.g., Jackendoff, 2007; Premack, 1990). Thus, an individual should act positively (or *reciprocate*) in response to a positive action, but should act negatively (or *retaliate*) in response to a negative action. Importantly, although reciprocal actions are expected to match initial actions in value, they need not match in form: If Mr. Smith kindly shoveled the driveway of elderly neighbor Miss Bee, we would deem it acceptable for her to offer him home-baked cookies in return. In the same vein, Trivers (2006) argued that "tit-for-tat" or "responding in kind" is "often the appropriate response in reciprocal relations. Do unto others what they have just done unto you. If someone is nice, be nice; if not so nice, not so nice" (p. 69).

First-party findings. Do infants follow the principle of reciprocity when responding to others' actions? To address this question, Dunfield and Kuhlmeier (2010) examined (in a task adapted from Behne et al., 2005) whether 21-month-olds would be more likely to give a toy to an individual who had demonstrated positive intentions toward them than to an individual who had not. During the familiarization phase, the infant stood across from two women, E1 and E2, who sat at a low table that sloped toward them; the Es took turns offering the infant a toy, and each E offered four toys in all. In one condition, *clumsy* E1 was unsuccessful at giving the infant the toy because it always rolled down the slanted tabletop, whereas *teasing* E2 always removed the toy when the infant reached for it. In the test phase, an experimenter placed a new toy on the edge of the table, between the two Es, in such a way that it fell to the floor; both Es then leaned forward and stretched an arm toward it (so they were not looking at the infant). The experimenter encouraged the infant to pick up the new toy and give it to one of the Es. In accordance with the reciprocity principle, infants who were willing to give away the new toy were reliably more likely to give it to clumsy E1, who had demonstrated positive intentions toward them.

In another condition, infants were equally likely to give the new toy to clumsy E1 or to *successful* E2, who succeeded in giving them toys in the familiarization phase. In yet another condition, infants were reliably more likely to give the new toy to successful E1 than to *indifferent* E2, who explored but then abandoned each toy in the familiarization phase, leaving the infant free to pick it up. Together, these results make clear that what infants cared most about were the Es' intentions: They favored Es who intended to give them toys over Es who did not, whether or not they actually received toys from them.

Third-party findings. The results of Dunfield and Kuhlmeier (2010) suggested that 21-month-olds are sensitive to reciprocity, but alternative interpretations were possible: Perhaps the

infants were mainly signaling an affiliative preference for the experimenter who had demonstrated positive intentions toward them by approaching her (e.g., Fawcett & Liszkowski, 2012; Hamlin et al., 2007) or by imitating her intended actions (e.g., Bellagamba & Tomasello, 1999; Meltzoff, 1995). To avoid these interpretive difficulties, He, Jin, Baillargeon, and Premack (2013) conducted VOE experiments to examine infants' expectations about reciprocation and retaliation in third-party scenes involving very different initial and reciprocal actions.

In the first experiment, 15-month-olds first received three familiarization trials in which they saw live events involving two unfamiliar women, E1 and E2; E1 sat at a window in the right wall of the apparatus, and E2 sat at a window in the back. In the *give* condition, E1 had two cookies and E2 had none; E1 gave a cookie to E2, and then they ate the cookies. Next, E1 exited the apparatus, closing her window behind her, and E2 looked down and paused until the trial ended. Familiarization trials in the *steal* condition were similar except that E1 and E2 each had a cookie; E1 stole E2's cookie, stuffed it in her mouth, and left with her own cookie. Next, the infants received a pretest trial in which E2 was absent (her window was closed); in front of E1 were a pile of five small pieces of paper and a pile of five stickers, and next to her was a colorful box. During the trial, E1 placed a sticker on a piece of paper, admired it briefly, and then placed it in the box; she then repeated these actions with a second sticker (this trial served to make clear to the infants that E1's goal was to prepare and store stickers). In the test trial, E2 returned and watched E1 prepare and store two more stickers. Finally, after E1 placed the fifth and last sticker on the last piece of paper, a bell rang, and E1 said, "Oh! I have to go! I'll be back!"; she then exited, leaving her last sticker on the apparatus floor. Next, either E2 put the sticker in the box, thus helping E1 by completing her actions (*store* event), or she tore up the sticker and dropped the pieces on the apparatus floor (*tear* event); in either case, after finishing her actions, E2

looked down and paused as before.

During the familiarization trials, the infants in the give and steal conditions tended to look equally, thus confirming prior results that infants do not expect individuals to always act positively toward others (e.g., Hamlin et al., 2007, 2011; Premack & Premack, 1997). During the test trial, however, the infants in the two conditions responded differently: The infants in the give condition looked reliably longer at the final paused scene if shown the tear as opposed to the store event, whereas the infants in the steal condition had the opposite looking pattern. These results suggested that the infants expected E2's actions toward E1 to follow the principle of reciprocity: When E1 had acted positively toward E2, they detected a violation if E2 chose to act negatively toward E1; conversely, when E1 had acted negatively toward E2, they detected a violation if E2 chose to act positively toward E1.

However, another possible interpretation of the results was that the infants were merely responding to low-level novelty effects: After seeing the give (steal) event in the familiarization trials, perhaps the infants viewed the tear (store) event as perceptually more novel. To rule out this interpretation, He at al. (2013) ran a second experiment identical to the first except that in the test trial E2 entered the apparatus only *after* E1 had exited; as before, E2 picked up the sticker left on the apparatus floor and either stored it or tore it up. The infants in the give and steal conditions now looked about equally at the store and tear events: Because E2 did not know that the sticker belonged to E1, E2's actions could not be perceived as reciprocal actions knowingly produced in response to E1's initial actions. The principle of reciprocity no longer applied, and the infants had no specific expectations about E2's actions.

Together, these results supported several broad conclusions. First, the results confirmed that infants evaluate actions toward others and assign a positive valence to beneficial actions

(sharing, helping) and a negative valence to detrimental actions (stealing, destroying). Second, the results provided further evidence that infants do not expect initial actions toward others to be uniformly positive: The infants in both experiments tended to look equally whether E1 shared her cookies with E2 or stole E2's cookie. Third, and most importantly, the results indicated that by 15 months infants expect reciprocal actions to match initial actions in valence, in accordance with the reciprocity principle. This expectation of tit-for-tat is quite abstract: It applies to scenes involving either reciprocation or retaliation, and it holds even when infants witness initial and reciprocal actions that differ markedly in form (e.g., stealing cookies and tearing up stickers). Finally, this expectation of reciprocity is unlikely to be due to low-level factors, because it is observed only when reciprocal actions are knowingly produced in response to initial actions.

2-III. DO INFANTS SHOW SENSITIVITY TO THE FAIRNESS PRINCIPLE?

A common theme in the experiments described in the preceding sections was that, although infants showed expectations about how individuals should act following others' actions (e.g., approach those who produced positive actions, retaliate in response to negative actions), they had no specific expectations about individuals' *initial* actions—that is, about how individuals should act toward others in the first place. Across experiments, infants tended to look equally in the habituation or familiarization trials whether individuals helped or hindered others, shared their possessions with others or stole from them, and so on (e.g., Hamlin et al., 2007, 2011; He et al., 2013; Premack & Premack, 1997). One interpretation of these negative results was that infants are generally agnostic about initial actions: They expect neither consistently positive nor consistently negative initial actions toward others. Another interpretation was that infants do possess expectations about initial actions, but that the scenes used in these experiments simply happened *not* to tap any of the principles that would have given rise to such

expectations. Recent research on infants' sensitivity to the principles of *fairness* and *ingroup* provides evidence for the second of these interpretations; we focus on the fairness principle in this section and on the ingroup principle in the next section.

1. Do infants expect individuals to allocate resources fairly?

One context that is commonly studied in research on fairness in young children is the *allocation of windfall resources*. Here again, researchers have used both first-party tasks, where the children tested are potential recipients, and third-party tasks, where they are not. Perhaps not surprisingly, given young children's pervasive difficulty in curbing their self-interest, the two types of tasks have produced markedly inconsistent results. Thus, although 3- to 4-year-olds show sensitivity to fairness when tested with third-party tasks, where self-interest cannot intrude (e.g., Olson & Spelke, 2008; Rochat et al., 2009), they show no such sensitivity when tested with first-party tasks, where self-interest typically prevails (e.g., Benenson, Pascoe, & Radmore, 2007; Blake & Rand, 2010; Damon, 1975; Fehr, Bernhard, & Rockenbach, 2008; Gummerum, Hanoch, Keller, Parsons, & Hummel, 2010; Rochat et al., 2009). In one first-party experiment, for example, 3- to 4-year-olds chose how sweets should be shared between themselves and an anonymous child (Fehr et al., 2008). Across conditions, children chose between an allocation of one sweet for themselves and one sweet for their partner (1,1) and an allocation of (1,0), (1,2), or (2,0). Children chose randomly in the first two conditions (they received one sweet either way and did not much consider what their partner would get), and they chose the (2,0) allocation in the last condition, to maximize their own gain. Similarly, Brownell, Svetlova, and Nichols (2009) found that 2-year-olds who could share treats with an experimenter tended to choose randomly between a (1,1) and a (1,0) allocation, unless the experimenter explicitly expressed her desire for a treat. As might be expected given these results, investigations of infants' concern

about fairness in the allocation of resources have relied mainly on third-party tasks (e.g., Geraci & Surian, 2011; Schmidt & Sommerville, 2011; Sloane, Baillargeon, & Premack, 2012).

In one VOE experiment, Sloane et al. (2012) showed 19-month-olds live events in which an experimenter divided two objects between two identical animated puppet giraffes. The infants saw an equal allocation (*equal* event) and an unequal allocation (*unequal* event) on alternate trials for three pairs of trials. At the start of each trial, the two giraffes protruded from openings in the back wall of the apparatus; in front of each giraffe was a small placemat. The giraffes "danced" until the experimenter entered the apparatus, brought in a tray with two identical objects (e.g., edible cookies), and announced, "I have cookies!"; the giraffes then responded excitedly, "Yay, yay!" (in two distinct voices). Next, the experimenter placed either one object in front of each giraffe (equal event) or both objects in front of the same giraffe (unequal event). Finally, the experimenter left, and the two giraffes looked down at their placemats and paused until the trial ended. The infants looked reliably longer at the final paused scene in the unequal than in the equal event, suggesting that, by 19 months, infants expect a distributor to divide resources fairly between two similar recipients. This conclusion was supported by two control conditions. In one, the giraffes were inanimate (they never moved or talked), and the infants looked about equally at the two test events. In the other, instead of bringing in and distributing the two objects in each trial, the experimenter removed covers resting over the giraffes' placemats to reveal the objects; the infants again looked equally at the two test events, suggesting that they did not merely expect similar individuals to have similar numbers of items.

Geraci and Surian (2011) showed 16-month-olds computer-animated events involving five different characters: There were two distributors (a bear and a lion), two recipients (a donkey and a cow), and one observer (a chicken). The infants saw an *equal* event in two trials

and an *unequal* event in two trials. At the start of the equal event, the fair distributor stood at the center of the computer monitor, near the two disks to be allocated. The observer entered, brought the two disks closer to the distributor, and then watched the scene from the bottom of the monitor. Next, the two recipients entered and took up positions in the top two corners of the monitor. Finally, the distributor gave one disk to each recipient, and the five characters then paused until the trial ended. The unequal event was identical except that it involved the unfair distributor, who gave both objects to the same recipient. At the end of the testing session, the infants also received a choice task (modeled after Hamlin et al., 2007) in which they were presented with pictures of the two distributors on a tray and encouraged to pick up the picture they wanted. Results indicated that, although the infants tended to look equally at the equal and unequal events (perhaps due to the sheer number of characters and actions in each event), they were reliably more likely to reach for the fair than the unfair distributor. This preference was eliminated in a control experiment where the recipients were replaced with inanimate artifacts. Together, these results indicated that by 16 months of age, infants already show sensitivity to fairness, and that this sensitivity guides their sociomoral evaluations.[15]

Finally, Schmidt and Sommerville (2011) showed 15-month-olds videotaped events in which a female distributor divided either crackers or milk between two female recipients. In the cracker scenario, for example, the distributor brought in a bowl with four crackers, the recipients slid their empty plates toward her, and the distributor then placed crackers on the plates, though the outcomes of her actions were blacked out (a screen appeared on the computer monitor occluding the bowl and plates). Next, the infants saw two still frames (in counterbalanced order) in which the empty bowl rested in front of the distributor and the two recipients looked down at their plates; in the *equal* frame, each recipient had two crackers, and in the *unequal* frame, one

recipient had one cracker and one recipient had three crackers. The infants looked reliably longer at the unequal than at the equal frame, suggesting that they expected the distributor not only to give some resources to each recipient, but to divide them equally between the two recipients. This interpretation was supported by the results of additional trials in which the infants saw the same still frames except that the distributor and recipients were absent; the infants now looked equally at the two frames.

Together, the preceding results suggest that, by 15-19 months of age, infants expect a distributor to allocate resources fairly between two similar recipients: Infants detect a violation when one recipient is given resources and the other recipient is given none, or when both recipients are given resources but in unequal amounts. Do infants younger than 15 months also expect distributors to divide resources fairly? Although initial investigations using the last two procedures described above yielded negative results with 10- and 12-month-olds (Geraci & Surian, 2011; Sommerville, Schmidt, Yun, & Burns, 2013), a more recent investigation using the (arguably simpler) procedure of Sloane et al. (2012) has yielded positive results with 9-month-olds (Sloane & Baillargeon, 2012a; see also Meristo, Strid, & Surian, 2012). Sensitivity to fairness thus appears to already be present in the first year of life.

2. Do infants expect individuals to dispense rewards fairly?

Another context that is commonly studied in research on fairness in young children is the *dispensation of rewards for merit*. Investigations of 3- to 5-year-olds' responses in this context have yielded the same mixed results as in the resource-allocation context: When tested with first-party tasks, preschool children show little sensitivity to merit (they keep more rewards than they should or divide them equally regardless of merit; e.g., Hook & Cook, 1979; Lane & Coon, 1972; Kanngiesser & Warneken, 2012; Lerner, 1974); when tested with third-party tasks,

however, children perform better (e.g., Baumard, Mascaro, & Chevalier, 2012; Leventhal, Popp, & Sawyer, 1973; Thomson & Jones, 2005; Tsutsu, 2010). In one third-party experiment, for example, Baumard et al. (2012) told 3- and 4-year-olds a simple story illustrated with pictures: Amélie and Hélène began to bake cookies together, but Hélène soon stopped working and began to play, leaving Amélie to finish on her own. In one condition, children were then shown a large and a small cookie, and they were asked who should receive the large cookie; at each age, children were reliably more likely to give the large cookie to Amélie (the worker) than to Hélène (the slacker). In another condition, children were shown three cookies; at each age, children were reliably more likely to give two cookies to the worker than to the slacker.

Can infants also take merit into account when reasoning about the distribution of rewards in a third-party task? To find out, Sloane et al. (2012) showed 21-month-olds live events in which a female experimenter asked two women to put away toys and then rewarded them equally whether both had worked to complete the chore (*equitable* event) or one of the women had done all the work while the other played (*inequitable* event). At the start of each event, the two women knelt at open windows in the right and left walls of the apparatus. Next to each woman was an open transparent box, and at the center of the floor was a pile of toys. Each woman played with the toys until the experimenter opened a window at the back of the apparatus and asked the women to put away the toys. In the *explicit* condition, the experimenter told the women they would receive a sticker if they complied; in the *implicit* condition, the experimenter did not mention the stickers beforehand. In both conditions, the experimenter exited after delivering her instructions, closing her window behind her. In the equitable event, both women then worked at putting away the toys, each placing half in her box. In the inequitable event, one woman (the slacker) continued to play while the other woman (the worker) placed all of the toys

in her box. In both events, the women then closed their boxes, and the experimenter returned. After inspecting each box (because the boxes were transparent, the experimenter could determine who had worked in her absence), the experimenter gave a sticker to each woman and then left. In the final scene of each event, the women played with their stickers until the trial ended.

The infants in the explicit condition looked reliably longer at the final scene if shown the inequitable as opposed to the equitable event, suggesting that they detected a violation when the worker and the slacker were rewarded equally. The infants in the implicit condition showed the same looking pattern, indicating that a prior explicit contract is not necessary for infants to hold expectations about the dispensation of rewards. These conclusions were supported by a control condition identical to the explicit condition except that the women's boxes were not transparent. For half the infants, the boxes were completely opaque (painted beige); for the other infants, the boxes retained a clear window at the front through which the infants (but not the experimenter) could see their contents. In either case, the infants now looked equally at the two test events. Thus, infants held clear expectations about the experimenter's actions only when she could determine who had worked and who had not; when she could not see the boxes' contents, infants no longer detected a violation when she rewarded the worker and the slacker equally.

Together, the results presented in this and the preceding sections suggest three main conclusions. First, infants in the second (and even the first) year of life expect unfamiliar individuals to produce fair distributions. Second, this expectation is unlikely to be due to low-level factors, because the same behavior on the part of the experimenter—giving one item to each of two individuals—is viewed as expected when the experimenter is allocating resources in a windfall situation, but as unexpected when the experimenter is dispensing rewards to a worker and a slacker. Third, this last result makes clear that infants' sense of fairness is not based on a

simple concept of equality (or inequality aversion), but rather on a concept of *equity*. Infants do not simply expect all individuals to be treated equally: Had this been the case, the infants in all three conditions above would have expected the experimenter to give a reward to each woman in both test events. Rather, infants expect individuals to *receive the treatment they deserve*, whether it is a fair share of a windfall or a fair reward for the work they have performed.

To bolster the notion that infants consider equitable or just deserts, it will be important to explore infants' intuitions about fairness in the context of *negative* actions (as opposed to positive actions such as the distribution of resources and rewards). First, consider *justifiable* negative actions such as punishments for misdeeds. If infants reason in terms of just deserts, then they should expect an experimenter who is meting out punishments in a scene to punish only those who committed misdeeds. Next, consider *unjustifiable* negative actions such as theft or vandalism. Here, infants should hold no expectation about fairness: As no one deserves the theft or wanton destruction of their possessions, it does not make sense to expect antisocial actions to be performed "fairly" with respect to all potential victims involved. Thus, if two individuals had two cookies each, infants should look about equally whether a thief stole two cookies from the same individual or one cookie from each individual. Such a negative result, coupled with the positive results reviewed above, would provide robust evidence that for infants, fairness is not a matter of treating individuals the same way but a more subtle and context-sensitive matter of treating individuals the way they deserve to be treated.

Finally, research is also needed to explore how and when infants begin to identify the various factors that, in specific contexts, render acceptable outcomes that would otherwise be perceived as fairness violations. Everyday family life is filled with pertinent experiences for learning about some of these factors: For example, daddy always receives a bigger scoop of ice

cream, big sister always gets to stay up much later, baby is never allowed to play with knives, earrings, pills, lipstick, or paper money, and only the dog is allowed to eat from its bowl. Over time, infants no doubt gradually learn about the many factors that affect who gets what in different situations within their family life as well as within their broader social environment.

2-IV. DO INFANTS SHOW SENSITIVITY TO THE INGROUP PRINCIPLE?

Most adults belong to multiple social groups, including family, friends, co-workers, neighbors, and so on. According to Tajfel, Bilig, Bundy, and Flament (1971), "A network of intergroup categorizations is omnipresent in the social environment…the articulation of an individual's social world in terms of its categorization into groups becomes a guide for his conduct…An undifferentiated social environment makes very little sense and provides no guidelines for action" (p. 153). In the same vein, Brewer (1999) argued that social boundaries can "shift from person to person or context to context so as to be more or less inclusive depending on local conditions or individual needs. Wherever drawn, however, ingroup-outgroup distinctions shape social interactions" (p. 432).

Part of what guides social interactions is the *ingroup* principle: Members of a social group should act in ways that sustain the group (e.g., Brewer, 1999; Haidt & Joseph, 2007; Tajfel et al., 1971). The ingroup principle has two corollaries, *loyalty* and *support*, each of which carries a rich set of expectations. Thus, ingroup loyalty dictates that in situations involving ingroup and outgroup individuals, one should (1) *prefer* and *align with* ingroup as opposed to outgroup individuals, (2) *protect* ingroup individuals who are threatened by outgroup aggressors, and (3) display *favoritism* toward ingroup over outgroup individuals (e.g., when allocating resources). Ingroup support dictates that when interacting with ingroup individuals, one should (1) engage in prosocial actions such as *helping* ingroup members in need of assistance,

comforting ingroup members in distress, and *sharing* resources with ingroup members, and (2) limit negative interactions within the ingroup by *refraining* from unprovoked negative actions, *curbing* retaliatory actions, and engaging in *social acting*, the everyday social pretense that adults produce—in the form of white lies, false cheer, and so on—to avoid hurtful or awkward interactions with ingroup members (e.g., Baillargeon et al., 2013; Yang & Baillargeon, in press).

Two additional comments may be helpful. First, several of the expectations listed above could also be described as interactions between the ingroup principle and other sociomoral principles. For example, defending an ingroup member who is threatened by an outgroup aggressor represents an interaction between the ingroup and reciprocity principles: The commitment to the ingroup leads to co-retaliation against the outgroup aggressor. Second, in contexts where the ingroup principle and other principles provide *diverging* guidelines for action, different cultures may make different choices about which principle should be ranked first. For example, some cultures may rank fairness above ingroup in some contexts (and thus frown upon nepotism and other forms of ingroup favoritism), whereas other cultures may adopt the reverse ranking. Children in each culture must learn these "preference orderings among competing . . . moral values" (Dupoux & Jacob, 2007, p. 377).

Do infants act, and expect others to act, in accordance with the ingroup principle? Below, we review recent findings on ingroup loyalty and then ingroup support.

A. Do Infants Show Sensitivity to Ingroup Loyalty?

1. Do infants prefer their ingroup and align their choices with those endorsed by their ingroup?

In a seminal first-party experiment, Kinzler, Dupoux, and Spelke (2007) examined whether infants might *prefer* a speaker of their native language over a foreign speaker and might

demonstrate this preference by choosing a toy offered by the native speaker over an identical toy offered by the foreign speaker. In the experiment, 10-month-olds from monolingual English families in Boston and from monolingual French families in Paris sat at a table in front of a computer monitor and received four test trials. Each trial had a speech phase and a toy-offering phase. During the *speech* phase, the infants watched, in alternation, videotaped events depicting a woman who spoke native infant-directed English and a woman who spoke native infant-directed French. During the *toy-offering* phase, the two women stood side by side, each silent, smiling, and holding an identical toy animal. In synchrony, the two women lowered their toy forward until it disappeared from the monitor (as though offering it to the infant); at that point, real-life versions of the toys appeared on the table, below the still images of the women. In each country, infants were reliably more likely to reach for the toy offered by the woman who spoke their native language.

In the preceding experiments, the two women held identical toys; in subsequent experiments, the women held different foods or toys, and Shutts, Kinzler, and their colleagues examined whether infants would *align their choice* with that of the woman who spoke their native language (e.g., Kinzler, Dupoux, & Spelke, 2012; Shutts, Kinzler, McKee, & Spelke, 2009). In an experiment on food choices (Shutts et al., 2009), 12-month-olds from English-speaking households sat at a table in front of a computer monitor and received two familiarization trials and one test trial. In one familiarization trial, the infants watched a videotaped event in which a woman tasted a food (e.g., plum sauce in a purple cup) and spoke about it in English, with positive affect; at the end of the event, the woman offered the food to the infant, and a real-life replica of the food appeared on the table so that the infant could sample it. The other familiarization trial was identical except that it involved another woman, who tasted

a different food (e.g., applesauce in a green bowl) and spoke about it in French (the food and language paired with each woman were counterbalanced). In the test trial, the two women stood side by side and silently offered their foods; real-life examples of the foods appeared on the table, and the investigators measured which container the infant touched first. The infants reliably chose the food offered by the English speaker over the food offered by the French speaker, suggesting that they aligned their food selection with that of the speaker of their native language. In an experiment on toy choices (Kinzler et al., 2012), 10-month-olds from English-speaking families received four test trials. Each trial had a speech phase, in which the infants heard, in alternation, a woman who spoke English and a woman who spoke French, followed by a toy-modeling phase, in which the two women stood side by side, each silent, smiling, and holding a different toy animal (a frog or a cow, counterbalanced); real-life replicas of the two toys rested on a table below the computer monitor throughout the trials. Following the toy-modeling phase, the infant was wheeled closer to the table to select one of the toys. Across trials, the infants reliably chose the toy held by the English speaker, suggesting that simply seeing her hold the toy (without offering it) was sufficient to guide their toy selection.

Together, the preceding results indicate that by 10-12 months of age infants (1) prefer speakers of their native language over foreign speakers and (2) align their food and toy choices with those endorsed by speakers of their native language. What should we make of these results? One interpretation is that the infants' responses were driven mainly by *familiarity*: Specifically, the infants preferred the woman whose language was familiar and, by association, the food or toy in her possession (e.g., Dunham, Baron, & Banaji, 2008; Kinzler, Shutts, & Correll, 2010). Another, richer interpretation is that in these minimal contexts contrasting two unfamiliar women from different speech communities, the infants extended (temporarily) *ingroup status* to the

woman from their speech community and, as a result, tended to prefer her and to align their choices with hers, in accordance with the principle of ingroup loyalty. This second interpretation would suggest that, even before they begin to speak, infants use language information to assign unfamiliar individuals to social groups and to guide their own responses to the individuals.

To decide between the two interpretations above, evidence is needed that cannot be attributed to simple familiarity-based preferences. One approach might be to test (e.g., using VOE or anticipatory-looking tasks) infants' expectations about third-party interactions involving individuals from *unfamiliar* social groups (whether real or arbitrary). For example, would infants expect an adult dressed in an outlandish national costume to approach an individual wearing the same costume as opposed to an individual wearing a different costume? Positive evidence would suggest that infants not only demonstrate an ingroup bias in their own responses, but expect individuals from other social groups to do the same, in accordance with the ingroup principle.

One final cautionary comment might be in order: The proposal that the infants in the preceding experiments extended ingroup status to the speaker of their native language should not be taken to suggest that infants routinely view *all* members of their speech community as members of their ingroup. For infants as for adults, ingroup boundaries may be somewhat fluid and context-dependent. Just as an American tourist may momentarily extend ingroup status to another American tourist encountered in a hotel in Paris, infants facing one individual who speaks their language and one who does not may temporarily extend ingroup status to the former but not the latter. Infants who meet unfamiliar adults from their speech community in a more typical situation where language does not provide a salient basis for social categorization (e.g., in a park, supermarket, or waiting room) may well treat the adults as outgroup individuals.[16]

2. Do infants expect individuals to co-retaliate against outgroup aggressors?

When an individual's ingroup member is mistreated by an outgroup aggressor, do infants expect the individual to co-retaliate against the aggressor, in accordance with the principle of ingroup loyalty? He, Baillargeon, and Premack (2011a) conducted a series of third-party experiments to address this question. Rather than using real-life social categories (e.g., categories based on age, race, gender, or language), He et al. built on the well-established finding from social psychology that almost any salient basis for categorization, however minimal or arbitrary, can lead adults to categorize individuals into distinct social groups and foster preferential treatment of ingroup members (e.g., Ashbrun-Nardo, Voils, & Monteith, 2001; Brewer, 1999; Gregg, Seibt, & Banaji, 2006; Tajfel et al., 1971; for similar results with older children, see Dunham, Baron, & Carey, 2011; Master & Walton, 2013; Patterson & Bigler, 2006; Vaughan, Tajfel, & Williams, 1981). Accordingly, He et al. created two novel social groups, using highly distinctive outfits as group markers, and asked whether 13-month-olds would detect a violation if a target individual helped an outgroup perpetrator who had stolen a toy from the target individual's ingroup member.

In the first experiment, 13-month-olds watched live events involving three women: One (A1) wore a bright orange hooded sweatshirt and purple-framed eyeglasses, and the other two (B1 and B2) wore black turtlenecks and tiger-patterned ears and chokers. The infants in the *outgroup-provokes-ingroup* condition first received four familiarization trials. At the start of each trial, all three women sat around the apparatus: A1 sat at a window on the right, B1 sat at a window on the left, and B2 (the target individual) sat at a window at the back, facing the infant. During the first two trials, a plastic egg and a marble stood in front of B1; while A1 and B2 watched, B1 put the marble inside the egg and then shook it, causing it to rattle. At the start of the next two trials, the marble rested across the apparatus from B1, close to A1; after B1

unsuccessfully attempted to retrieve the out-of-reach marble (e.g., Warneken & Tomasello, 2007), A1 picked up the marble and quickly left with it, closing her window as she went. In the next, pretest trial, B1 and her toys were absent, and her window was closed; in front of A1 were now a pole and a tray filled with colorful rings of varying widths. While B2 watched, A1 stacked the rings on the pole, from largest to smallest. The test trial began in the same way as the pretest trial except that the smallest ring now rested across the apparatus from A1. After A1 had stacked the available rings on the pole, she tried unsuccessfully to retrieve the out-of-reach ring and then exited, closing her window. Next, B2 grasped the smallest ring and either placed it in the tray so that A1 could stack it when she returned (*help* event) or deliberately dropped it to the floor (*hinder* event). B2 then paused, looking down at the apparatus floor, until the trial ended.

The infants looked reliably longer at the final paused scene if shown the help as opposed to the hinder event, suggesting that (1) they assigned A1 to a different social group than B1 and B2, and (2) after seeing A1 act negatively toward B1, they viewed it as unexpected and disloyal for B2 to help A1, in accordance with the ingroup principle. This interpretation was supported by the results of two additional conditions. One examined the possibility that the infants in the outgroup-provokes-ingroup condition detected a violation when B2 helped A1 merely because A1 belonged to the outgroup. In this *outgroup-does-not-provoke* condition, the infants did not receive the last two familiarization trials and thus never saw A1 steal B1's marble. The infants now looked about equally at the help and hinder events, indicating that, when A1 committed no provocative act, the infants viewed it as equally expected or acceptable for B2 to help or hinder A1. The other condition assessed the possibility that the infants in the outgroup-provokes-ingroup condition detected a violation in the help event out of a sense of justice (rather than loyalty), because they expected any perpetrator—whether outgroup or ingroup—to be punished.

This *ingroup-provokes-outgroup* condition was identical to the outgroup-provokes-ingroup condition except that the target individual wore an orange sweatshirt and purple eyeglasses and thus belonged to the same social group as A1 (the perpetrator) instead of B1 (the victim); we refer to the target individual in this condition as A2. The infants now looked reliably longer if shown the hinder as opposed to the help event, indicating that, although A1 had again stolen B1's marble, the infants had no expectation that A1 should be punished. A1 needed help, and ingroup A2 was expected to provide it; the fact that A1 had mistreated outgroup B1 was apparently ignored or dismissed as of little import.

In a second experiment, additional 13-month-olds were tested using the procedure of the outgroup-provokes-ingroup condition, with one exception: In the last two familiarization trials, B2 exited the scene either *after* (*knowledgeable-B2* condition) or *before* (*ignorant-B2* condition) A1 stole the marble from B1. In the test trial, all infants saw the help event. As expected, the infants in the knowledgeable-B2 condition looked reliably longer at the final paused scene than did those in the ignorant-B2 condition. Thus, (1) the infants in the knowledgeable-B2 condition responded as did as the infants in the outgroup-provokes-ingroup condition of the first experiment (B2 knew that A1 had stolen B1's marble and so should not help A1), and (2) the infants in the ignorant-B2 condition responded as did the infants in the outgroup-does-not-provoke condition of the first experiment (B2 did not know that A1 had stolen B1's marble and so could help A1).

Finally, in a third experiment, 29-month-old toddlers were tested using the same three conditions as in the first experiment. The two novel social groups were marked with nonsense labels, instead of distinctive outfits (prior findings indicate that infants and toddlers interpret naming phrases such as "a dax" as referring to distinct categories; e.g., Balaban & Waxman,

1997). Prior to the familiarization trials, the toddlers received two labeling trials in which A1 said, "I'm a jaybo!", B1 said "I'm a topid", and B2 said "I'm a topid too!". These labels were repeated twice in each labeling trial, and then the procedure continued as in the first experiment (though slightly abbreviated for these older children). Even though the toddlers in this experiment had to remember to which group each individual belonged, results were identical to those from the first experiment.

Together, these results support three main conclusions. First, like older children and adults, infants and toddlers (1) form "minimal groups" when provided with salient bases—such as outfits or labels—for categorizing individuals, and (2) immediately hold expectations about social interactions within and between these groups. Second, one such expectation concerns *loyalty*: Infants and toddlers detect a violation when a target individual helps an outgroup perpetrator who has mistreated a member of the target individual's ingroup. In our discussion of the reciprocity principle, we saw that infants understand retaliation: After E1 hinders E2, infants deem it unexpected if E2 helps E1, but not if she hinders E1 (He et al., 2013). The present results suggest that infants and toddlers also understand *co-retaliation* against an outgroup perpetrator: Although A1 stole the toy of B1, not B2, the children viewed it as unexpected if B2 helped A1, but not if she hindered A1. This effect did not come about because the children were confusing B1 and B2, or treating them as interchangeable tokens of the same social group: Recall that the infants held clear expectations about B2's actions only when she had witnessed A1's poor treatment of B1. In sum, the children detected a violation when B2 acted positively toward perpetrator A1, but they detected no violation when B2 loyally co-retaliated against her, in accordance with the ingroup principle.

Third, although infants and toddlers have no particular expectation about whether

outgroup individuals who cannot reach a goal object should be helped or hindered, the same is not true of *ingroup* individuals. When out-group A1 needed assistance, the children viewed it as equally expected if B2 helped or hindered A1 (as long as A1 had not mistreated B1 in B2's presence). In contrast, when in-group A1 needed assistance, the children viewed it as unexpected if A2 hindered as opposed to helped A1 (apparently, A1's negative treatment of B1 was dismissed as irrelevant; A1 needed help, and A2 was expected to provide it). These results are consistent with Brewer's (1999) proposal that "ingroup love is not a necessary precursor of outgroup hate" (p. 442). According to Brewer, "discrimination between ingroup and outgroups is a matter of relative favoritism toward the ingroup and the absence of equivalent favoritism toward outgroups" (p. 434). Viewed from this perspective, the present results also suggest a possible interpretation of the many negative findings reviewed earlier showing that infants looked about equally (in habituation or familiarization trials) when an individual acted positively or negatively toward another individual (e.g., Hamlin et al., 2007; He et al., 2013; Premack & Premack, 1997). Specifically, it could be that, because the infants lacked sufficient evidence to infer that the individuals in the events belonged to the *same* social group, they tended to view the positive and (mild) negative actions performed as equally expected.

3. Do infants expect individuals to display ingroup favoritism when allocating resources?

We saw earlier that when a distributor divides resources between two similar recipients, infants expect the distributor to do so fairly (e.g., Schmidt & Sommerville, 2011; Sloane et al., 2012). What happens when the two recipients are dissimilar, with one belonging to the same social group as the distributor and the other belonging to a different social group? In this situation, the fairness and ingroup principles lead to diverging expectations: Fairness dictates that each recipient should get an equal share, whereas ingroup loyalty suggests that the ingroup

recipient should receive at least a larger share. How do infants expect the distributor to act in this situation? To examine this issue, Sloane, Baillargeon, and Premack (2013) tested 19-month-olds in a third-party resource-allocation task (adapted from Sloane et al., 2012) involving animated puppets from two distinct groups, monkeys and giraffes. The infants in the *dissimilar-recipients* condition received three test trials in which a distributor divided two objects (edible cookies, toy ducks, or toy cars) between two recipients, one from the same group as the distributor (ingroup recipient) and one from the other group (outgroup recipient). The distributor was a monkey for half the infants and a giraffe for the other infants, so that which recipient was ingroup and which was outgroup varied across infants. Across trials, the infants saw one of three test events: an *equal* event, in which the distributor gave one object to each recipient (e.g., the monkey distributor gave one object each to the monkey and giraffe recipients); a *favors-ingroup* event, in which the distributor gave both objects to the ingroup recipient (e.g., the monkey distributor gave both objects to the monkey recipient); or a *favors-outgroup* even, in which the distributor gave both objects to the outgroup recipient (e.g., the monkey distributor gave both objects to the giraffe recipient). The infants (a) looked reliably longer if shown the equal event or the favors-out-group event than if shown the favors-in-group event and (b) looked about equally at the equal and favors-out-group events. The infants thus expected the distributor to favor the ingroup recipient (favors-ingroup event), and they detected a violation when the distributor gave half (equal event) or all (favors-outgroup event) of the resources at its disposal to the outgroup recipient.

Support for this conclusion came from two additional conditions. In the *similar-recipients* condition, both recipients belonged to the different group than the distributor, who divided the objects either equally (equal event) or unequally (unequal event) between them (e.g., the monkey

distributor gave either one object each to the two giraffe recipients or both objects to one giraffe recipient). As in previous experiments, the infants looked reliably longer if shown the unequal as opposed to the equal event, ruling out the possibility that the infants in the dissimilar-recipients conditions simply assumed that fairness did not apply to outgroup recipients. All other things being equal, the infants did expect fairness; if for some reason a distributor had to divide resources between two outgroup recipients, then the distributor should do so fairly.

In the *group-induction* condition, the two recipients were again dissimilar, as in the dissimilar-recipients condition. Prior to the test trials, however, the infants now saw the monkey and giraffe recipients engage in a friendly game of pat-a-cake, while the (monkey or giraffe) distributor watched. In the test trials, the infants saw an equal or a favors-ingroup event. Because coordinated activity is generally held to be a marker of group membership (e.g., Barragan & Dweck, 2012; Premack & Premack, 1995), the infants were expected to view the two recipients as belonging to the same social group, despite their different appearances. As predicted, the infants who saw the favors-ingroup event now looked reliably longer than those who saw the equal event. This result ruled out low-level interpretations of the findings of the dissimilar-recipients conditions (i.e., the equal event elicited reliably different responses across conditions) and also confirmed that coordinated activity speaks louder than appearance when it comes to determining group membership.

The preceding experiment was conducted in a Midwestern American town, and the results suggest that by 19 months American infants (and perhaps Western infants more generally) have learned to rank ingroup loyalty above fairness in resource-allocation situations: They expect the former to prevail over the latter when the two are in conflict. Further experiments using women dressed in distinctive outfits (princesses versus hippies) confirmed

this finding and also provided evidence that the ordering of ingroup loyalty above fairness is learned at about 18 to 19 months (Sloane & Baillargeon, 2013). How might this ordering be learned? It could be that around this age, as a result of well-documented developments in the sense of self (e.g., Amsterdam, 1972; Kärtner, Keller, & Chaudhary, 2010), the typical infant comes to identify as a member of a particular social group: a family or household. From that point on, as they go about their social environment, infants begin to notice that each family has its own stuff (e.g., house, car, furniture, clothes, dishes, toys, bicycles, and so on), which is not shared equally with other families. Through these everyday observations, infants gradually abstract the generalization that when ingroup loyalty and fairness both apply in a resource-allocation situation, ingroup loyalty is given priority: Resources are reserved for the ingroup, not squandered on the outgroup.

As time goes on, children no doubt acquire a more nuanced and context-sensitive understanding of the relationship between fairness and ingroup loyalty in their social environment. As a case in point, consider an experiment by Olson and Spelke (2008) with 3.5-year-olds. The children were presented with five dolls; one was identified as the protagonist and the other four were identified as the protagonist's siblings, as the protagonist's friends, or as strangers. The children were asked to help the protagonist distribute resources to the other dolls (e.g., divide two to four stickers among two siblings and two strangers). When there were fewer than four items to distribute, the children favored siblings and friends over strangers (in accordance with ingroup loyalty). When there were four items to distribute, however, the children divided them equally among the four dolls regardless of how they were identified (in accordance with fairness). This finding contrasts with the result above that 19-month-olds expect a monkey distributor who is dividing two objects between a monkey and a giraffe recipient to

give *both objects* to the monkey recipient. One possibility is that, at some point between 19 months and 3.5 years, American toddlers learn via admonitions from parents and teachers that fairness should be ranked above ingroup loyalty in typical play situations. Cross-cultural research is needed to examine whether children in more collectivist cultures show similar or different developmental trajectories in their responses to ingroup loyalty and fairness.

B. Do Infants Show Sensitivity to Ingroup Support?

According to the principle of ingroup support, individuals act prosocially toward their ingroup: *help* ingroup members in need of assistance, *comfort* ingroup members in distress, *share* resources with ingroup members, and so on. The study of the development of prosocial behavior, which began in the 1970s, has generated a very large body of work (for reviews, see Radke-Yarrow, Zahn-Waxler, & Chapman, 1983; Eisenberg et al., 2006). In recent years, there has been particular interest in infants' helping and comforting responses, as reviewed below. Sharing has been more difficult to study using first-party tasks because infants and toddlers rarely rectify unequal distributions in their favor: If they happen to receive more resources (e.g., toys or crackers) than an unlucky partner, they make no effort to equate the shares or reduce the disparity between them (e.g., Hamann, Warneken, Greenberg, & Tomasello, 2011; LoBue, Nishida, Chiong, DeLoache, & Haidt, 2011). Nevertheless, young children do share their resources under some conditions: for example, if their mother instructs them to do so, or if their partner explicitly requests resources (e.g., Dunfield, Kuhlmeier, O'Connell, & Kelley, 2011; Hay, Caplan, Castle, & Stimson, 1991; Levitt, Weber, Clark, & McDonnell, 1985).

1. Helping someone in need of assistance

In the first part of the chapter, we saw that infants in the second year of life produce various actions intended to help others: For example, they point to update an ignorant agent

about a toy's location, and they retrieve a toy for a mistaken agent who is searching for it in the wrong location (e.g., Buttelmann et al., 2009; Knudsen & Liszkowski, 2012a; Liszkowski et al., 2006, 2008). In an influential series of experiments, Warneken and his colleagues set out to explore the conditions under which infants engage in helping behavior (e.g., Warneken, 2013; Warneken & Tomasello, 2006, 2007, 2008; Warneken, Hare, Melis, Hanus, & Tomasello, 2007).

Infants as helpers. In the first experiment, Warneken and Tomasello (2006) built on earlier results by Rheingold (1982) and presented 18-month-olds with 10 scenarios involving a male experimenter; each scenario had an experimental version where the experimenter required help (e.g., he accidentally dropped a marker on the floor and unsuccessfully reached for it) and a control version where he did not (e.g., he deliberately threw the marker on the floor and did not reach for it). The scenarios were grouped into four categories: objects out of reach, as in the preceding example; obstacles (e.g., the experimenter could not open the doors of a cabinet because his hands were full); failed attempts (e.g., the experimenter placed a book on top of a stack, but the book slipped); and misconceived attempts (e.g., the experimenter used a small instead of a large opening to retrieve an object from a box). Each infant saw all 10 scenarios, five in the experimental version and five in the control version. Positive results were obtained with six of the scenarios (three from the out-of-reach category and one from each of the other categories), with reliably more infants helping the experimenter in the experimental than in the control version. Moreover, most infants helped in at least one of the five experimental scenarios, and in almost all cases they did so spontaneously, before the experimenter looked or spoke to them. Similar, though weaker, results were subsequently obtained in experiments that tested 18-month-olds with fewer scenarios. Thus, Svetlova, Nichols, and Brownell (2010) used three out-of-reach scenarios and found that 44% of the infants helped at least once, and Dunfield et al.

(2011) used a single out-of-reach scenario and reported that 33% of the infants helped. Clearly, the more scenarios are used, the more likely infants are to help at least once, as they become more comfortable with the task and the experimenter.

Warneken and Tomasello (2007) found that even 14-month-olds were capable of helping, appropriately and spontaneously, in the simpler, out-of-reach scenarios. Other experiments using these same scenarios showed that 18-month-olds helped even when they had to circumvent obstacles in order to do so (Warneken et al., 2007), or when they had to temporarily stop playing with an interesting toy in order to help the experimenter across the room (Warneken & Tomasello, 2008). Yet other experiments revealed a marked improvement at about 2 years of age in the ability to infer, in the absence of explicit cues, that an adult requires help to achieve a goal (Warneken, 2013).

On the basis of their results, Warneken and Tomasello (2009) concluded that infants' tendency to help others reflects a natural predisposition for altruism. Moreover, because "infants help rather indiscriminately, without taking into account if the beneficiary is a relative or a stranger" (p. 465), they proposed that "children start out as rather indiscriminate altruists who become more selective as they grow older" (pp. 466). This proposal differs from the framework adopted here, which ties helping and other prosocial behaviors to a predisposition for ingroup support, rather than altruism, and thus predicts that infants should be selective helpers: Specifically, they should be more likely to help ingroup than outgroup individuals.

Are infants indiscriminate or selective helpers? The experiments reviewed above could not provide a definite answer to this question, because they included a warm-up phase in which the experimenter played with the infant for several minutes; this coordinated activity could have induced the infant to extend *ingroup status* to the experimenter, thus facilitating prosocial

actions, in accordance with the principle of ingroup support. This possibility suggested that infants might be less likely to help an experimenter they did not perceive as a member of their ingroup (for additional arguments on selectivity in infants' helping efforts, see Wynn, 2009).

Evidence consistent with this suggestion comes from a striking experiment by Over and Carpenter (2009). They examined 18-month-olds' helping behavior in a single scenario where an experimenter accidentally dropped sticks to the floor. Prior to the test phase, the infants were introduced to the experimenter but did not play with her; instead, an assistant showed them eight color photographs, one at a time. Each photograph depicted an everyday object in the foreground (e.g., a teapot) and a prime in the background. The infants saw one of four different primes, depending on condition. The *affiliative* prime involved two small dolls facing each other in close proximity; the other primes were *non-affiliative* and involved the same dolls standing back to back, one doll standing alone, or two stacks of blocks about the same size as the dolls. Following the affiliative prime, about 60% of the infants spontaneously helped the experimenter pick up the sticks in the test phase; following the other primes, however, only about 20% of the infants did so. These results thus suggest that (1) most infants do not spontaneously help an unfamiliar adult, and (2) following exposure to stimuli that prime an *ingroup mindset*, infants are reliably more likely to engage in prosocial actions.

Barragan and Dweck (2012) reported preliminary evidence consistent with these conclusions. They found that 14- to 24-month-olds were reliably more likely to help a male experimenter in out-of-reach scenarios if during a 6-minute familiarization phase they interacted with the experimenter in *coordinated* play (e.g., the experimenter and the infant took turns rolling a ball or stacking rings) as opposed to *parallel* play (e.g., the experimenter played with his own set of toys, a few feet away from the child). In both play conditions, the experimenter

interacted positively with the child: Every 60 seconds, the experimenter looked at the child, smiled, and made friendly statements (e.g., "Toys are fun!"). Nevertheless, the infants in the coordinated-play condition produced significantly more helping actions, suggesting that the coordinated activities were critical for inducing an ingroup mindset. In a similar vein, Carpenter, Uebel, and Tomasello (in press) found that 18-month-olds were reliably more likely to help a female experimenter if during a 6-minute familiarization phase she promptly mimicked every action they produced, as opposed to performing a different action (most infants realized they were being mimicked and engaged in testing behaviors that resulted in turn-taking sequences similar to those in coordinated play).

The evidence that infants are more likely to help individuals from the ingroup is also consistent with some of the third-party findings reviewed earlier. Recall that, across several experiments, infants looked about equally at the help and hinder events when the individuals involved either belonged to different social groups or had undifferentiated group memberships; in contrast, infants looked reliably longer at the hinder event when the individuals involved were clearly identified as members of the same social group (e.g., Hamlin et al., 2007; He et al., 2011a, 2013; Premack & Premack, 1997). Research is needed to examine whether infants would view not only hindering but also *failing to help* an ingroup member as unexpected or unacceptable. For example, would infants detect a violation if a "helper" failed to bring an out-of-reach toy closer to the ingroup member who needed it? From the perspective of ingroup support, helping an ingroup member requiring assistance may be *obligatory*, at least under some conditions (e.g., Brewer,1999), and it will be interesting to explore whether even infants possess a sense of prosocial obligation to the ingroup.

Finally, one aspect of the results of Over and Carpenter (2010) is worthy of comment.

Readers may be surprised that infants primed with affiliative photos were then willing to help an unfamiliar experimenter. As Brewer (1999) observed, helping ingroup members does not have to be limited to familiar or well-liked associates: "clear group boundaries provide one mechanism for achieving the benefits of cooperative interdependence without the risk of excessive costs…By limiting aid to mutually acknowledged ingroup members, total costs and risks of non-reciprocation can be contained. Thus, ingroups can be defined as bounded communities of mutual trust and obligation that delimit mutual interdependence and cooperation. An important aspect of this mutual trust is that it is *depersonalized . . .*, extended to any member of the ingroup whether personally related or not" (p. 433, italics added). The results of Over and Carpenter fit well with this analysis: When primed to adopt an ingroup mindset, infants are willing to help even a near stranger.

2. Comforting someone in distress

Findings from first-party tasks. Over the past few decades, there has been a great deal of research on infants' and toddlers' responsiveness to distress (for reviews, see Dunfield & Kuhlmeier, 2013; Eisenberg et al., 2006). Comforting tasks in the laboratory have used distress scenarios from two main categories: sadness over a broken or lost toy (e.g., an adult "accidentally" breaks a toy and simulates sadness) and pain due to a minor physical injury (e.g., an adult "accidentally" bumps her knee and simulates pain). Here we focus on prosocial responses that went beyond expressions of concern, such as providing physical or verbal comfort (e.g., hugging), attempting to remedy the situation (e.g., offering an alternative object), or seeking help from a parent.

Very different levels of prosocial responding have been reported across comforting tasks (e.g., from 0% in some tasks to 85% in others), making it difficult to form a coherent picture of

early comforting behavior. One factor that clearly affects levels of responding is age: Comforting responses increase steadily in frequency between ages 1 and 3 (e.g., Dunfield & Kuhlmeier, 2013; Svetlova et al., 2010; Zahn-Wexler, Radke-Yarrow, & Wagner, 1992). Another (no doubt related) factor is that infants in comforting tasks may not always understand why the adult is distressed or what intervention might be appropriate. When discussing helping tasks in the last section, we saw that infants perform particularly well in out-of-reach scenarios (e.g., Warneken & Tomasello, 2006, 2007): As they watch the experimenter's unsuccessful efforts, they can easily determine what goal he is trying to achieve (retrieving the object) and what intervention is needed to help him attain his goal (giving him the object). In contrast to these transparent scenarios, scenarios in comforting tasks may often be more opaque for infants, leaving them unsure why the adult is distressed or what they can do about it.[17]

With these factors in mind, let us first consider comforting tasks that used scenarios in which an adult is sad because a toy is broken or lost. For example, in an experiment by Bischöf-Kohler (1991), 16- to 24-month-olds were first familiarized with a female experimenter, and then they received a test session in which the experimenter played with a teddy bear until she "accidentally" broke it, causing her to sob moderately for a distress period of up to 2.5 minutes. Results indicated that 31% of the children attempted to comfort the experimenter during the distress period. Using a similar scenario and procedure, Kärtner et al. (2010) tested 19-month-olds in Berlin and Delhi, and found that 28% of the infants in Berlin and 35% of the infants in Delhi produced prosocial responses. Finally, Vaish, Carpenter, and Tomasello (2009) tested 18- and 25-month-olds using a scenario in which a familiarized experimenter "accidentally" lost a helium balloon and expressed sadness for up to 2 minutes. Prior to this test phase, the children received a sympathy-induction phase. In one (*harm*) condition, the experimenter was the victim

of repeated negative actions by an assistant. In another (*neutral*) condition, the interactions of the experimenter and assistant were neutral. In the neutral condition, 38% of the children produced prosocial responses when the experimenter lost her balloon; in the harm condition, 66% of the children did so. This difference was reliable and indicated that the children's sympathy for the victimized experimenter induced them to act more prosocially toward her.

In the experiments above, levels of responding to the sad adult ranged from 28% to 38% under typical conditions. In contrast, much *lower* levels of responding have generally been found with scenarios involving an adult in pain. In a series of experiments, Dunfield and her colleagues tested 18-month-olds to 4.5-year-olds with scenarios in which an experimenter hit her knee against a table or slammed her finger in a door; she then simulated pain, rubbed the affected area, and vocalized about it. At 18 and 24 months, no infant showed prosocial responding during the 10 seconds that followed the injury (Dunfield et al., 2011). Responding was still slight at 2.5 years of age; it was reliably higher at ages 3.5 and 4.5 and consisted almost exclusively of verbal reassurances (Dunfield & Kuhlmeier, 2013; for related naturalistic observations in daycare settings, see e.g., Howes & Farver, 1987; Lamb & Zakhireh, 1997). The results of Dunfield et al. are similar to earlier findings by Zahn-Wexler et al. (1992). Two samples of 2-year-olds were brought to the laboratory on three separate occasions, spaced one month apart. The children saw a simple pain scenario on each occasion; one scenario involved their mothers and the other two scenarios involved female experimenters. Comforting responses to the experimenters were infrequent and ranged from 4% to 15% across samples and experimenters. In marked contrast, comforting responses to the mother ranged from 78% to 85%.

Together, the preceding results suggest that, although spontaneous comforting responses do occur in the second year of life, they are somewhat less common than spontaneous helping

responses. Identifying an appropriate prosocial intervention in a helping task may be easier for several reasons: Helping tasks often involve multiple trials, whereas comforting tasks do not; problems in helping tasks often have obvious solutions (e.g., an object out of reach), whereas comforting tasks do not (e.g., a broken toy); standard interventions in helping tasks (e.g., moving an object within reach) may feel right with a recently familiarized experimenter, whereas standard interventions in comforting tasks (e.g., hugging) may not; and facing an adult who needs instrumental help to attain a goal may be less novel and/or upsetting for infants than facing an adult who sobs or moans for some time.

    Findings from third-party tasks. Given the difficulties noted above, third-party tasks provide a valuable alternative for exploring infants' expectations about comforting responses. In seminal experiments, Johnson and her colleagues began to examine infants' expectations about caregivers' responsiveness to a crying baby in third-party tasks (Johnson, Dweck, & Chen, 2007; Johnson et al., 2010). In one VOE experiment, 13-month-olds were habituated to a computer-animated event involving a large red oval (the "adult") and a small blue oval (the "baby"). To start, the adult and baby stood at the bottom of a steep incline. The adult moved halfway up the incline to a small plateau, and the baby began to bounce and cry (the soundtrack used a recorded human infant cry). The animation then paused until the trial ended. Following habituation, the infants saw two test events. Each began where the habituation event had ended, with the adult on the plateau and the baby crying. In the *responsive* event, the adult returned and stopped next to the baby; in the *unresponsive* event, the adult continued up the incline and stopped at the top. In either case, the animation then paused until the trial ended. After the infants participated in this experiment, their attachment status (secure or insecure) was assessed using the Strange Situation procedure (Ainsworth, Blehar, Waters, & Wall, 1978). The securely attached infants looked

reliably longer at the unresponsive than at the responsive event, suggesting that they interpreted the events as an adult-baby interaction, expected the adult to attend to the crying baby, and detected a violation when the adult ignored the baby instead. In contrast, the insecurely attached infants showed a non-significant tendency to look longer at the responsive event.

In another experiment, 13-month-olds were habituated to two events similar to the responsive and unresponsive events from the previous experiment except that two different "adults" (a red triangle and a green square) were used in the events. The test events were modeled after those of Kuhlmeier et al. (2003): The hill was removed, the two adults stood in the top corners of the computer monitor, and the baby approached either the responsive or the unresponsive adult. The securely attached infants looked reliably longer when the baby approached the unresponsive as opposed to the responsive caregiver; the insecurely attached infants tended to show the reverse pattern, though this difference was again not significant.

Together, these results suggest that securely attached infants hold expectations about how adults should respond to a crying baby and that they can apply these expectations even when shown minimal adult-baby interactions. The responses of the insecurely attached infants are more difficult to interpret: The infants could have held different expectations about caregivers' responsiveness, or they could have had more difficulty processing the minimal events they were shown (maternal sensitivity also affected infants' performance in a physical-reasoning task with minimal computer-animated events; Hohenberger et al., 2012). This second interpretation suggested that more robust expectations about responsiveness to distress might be uncovered with more natural events.

With this possibility in mind, Jin et al. (2012) tested infants using realistic videotaped events. At the start of the responsive test event, one woman (e.g., a blonde) folded towels at a

table on the left side of a room; at the back of the room were a chair with additional towels and a large stroller (one could not see whether there was a baby inside the stroller). Next, a baby began to cry; the woman walked to the stroller and bent over it, as though attempting to comfort the crying baby. The unresponsive event was similar except that it involved a different woman (e.g., a brunette), who walked to the chair to pick up more towels, ignoring the crying baby. In a VOE experiment, 12-month-olds looked reliably longer at the unresponsive than at the responsive event; this effect was eliminated when the baby laughed (instead of cried) in the recorded soundtrack. In another experiment, 8-month-olds were tested with the infant-triggered-video (ITV) method: They faced two computer monitors, one depicting a still picture of the responsive woman bent over the stroller, and one depicting a still picture of the unresponsive woman bent over the chair. The infants were first shown that touching each picture triggered the corresponding event. Next, the infants chose the event they wanted to see by touching the appropriate picture. The infants were allowed three triggers, and they chose the unresponsive event reliably above chance; this effect was again eliminated when the baby laughed.

The third-party experiments summarized in this section indicate that 8- to 13-month-olds generally expect an adult to attend to a crying infant. Additional research is needed to examine whether this finding reflects only a local expectation abstracted from everyday experience (e.g., moms attend to crying babies), or whether it reflects a more general expectation. For example, would infants expect responsiveness to a distressed woman, as in the comforting experiments reviewed in the previous section? Finally, research is needed to examine the role of group membership in infants' expectations about responsiveness to distress. If expectations about prosocial behaviors are generally guided by the principle of ingroup support, as we suggested, then comforting responses should be viewed as obligatory in the ingroup, but as merely

permissible in the outgroup.

2-V. DO INFANTS SHOW SENSITIVITY TO THE PRINCIPLE OF AUTHORITY?

When discussing social groups so far, we have always spoken about them as though they were amorphous, unstructured entities. In truth, however, social groups often have a hierarchical structure, and researchers have suggested that expectations about the actions of more dominant and more subordinate individuals within the hierarchy are governed by a principle of *authority* (e.g., Fiske, 2004; Haidt & Joseph, 2007). Authority is to be understood as "a two-way street: Subordinates must show respect and deference, but superiors must then protect them from external threats and maintain order within the group. This pro-social side of authority seems to go unrecognized…norms and virtues govern the behaviors of superiors (e.g., impartiality, magnanimity, fatherliness) and subordinates (e.g., respect, deference)" (Haidt & Joseph, 2007, p. 384). Would infants be able to represent a social hierarchy? Would they hold differential expectations about the actions of more dominant and more subordinate individuals within the hierarchy? As a first step in addressing these questions, researchers have begun to examine whether infants can represent a *dominance* relation between two individuals. Initial results suggested that, when two individuals, A and B, have conflicting goals in a situation, 12- to 13-month-olds expect A to defer to B (1) if B is much larger than A (Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011), or (2) if B previously prevailed over A in a similar conflict situation (Mascaro & Csibra, 2012).

Building on these results, Mascaro and Csibra (2012) went on to examine whether 15-month-olds would expect the dominance relation between two individuals to be stable across *different* conflict situations (results with 12-month-olds were negative, pointing to more limited, situation-bound expectations). The infants watched a computer-animated familiarization event

involving two small geometric figures, "subordinate" A and "dominant" B; both figures had prominent eyes and differed in color and shape. To start, A was alone and repeatedly entered a small enclosure, to establish that this was its goal. Next, B arrived and monopolized the enclosure by repeatedly pushing A away. The familiarization event thus served to demonstrate that B prevailed over A in this first conflict situation. The test events were designed to assess whether the infants would expect B to also prevail over A in a second conflict situation. First, a small object fell from above, and one of the two figures entered from one side of the computer monitor, collected the object, and left. Next, a second object fell, and the other figure entered from the opposite side of the monitor, collected the object, and left. Finally, a third object fell, both figures approached it, and then either A (*unexpected* event) or B (*expected* event) collected the object. Each test event was preceded by the familiarization event in a familiarization-test-familiarization-test sequence, and different figures were used in the first two and last two events of the sequence. The infants looked reliably longer at the unexpected than at the expected event, suggesting that they (1) noticed which figure prevailed in the first conflict situation, (2) expected this dominance relation to be maintained in the second conflict situation, and hence (3) detected a violation when A took the third object instead of letting B have it. This result was eliminated if a new figure, C, replaced A in the test events. Thus, after seeing B be dominant over A in one conflict situation, infants generalized this asymmetric relation to a novel conflict situation, but not to a novel individual: They did not expect B to be dominant over C in the test events.

Could infants expect one individual to be dominant over another, even if they had not seen the two individuals interact before, through transitive inference? If in one situation C prevailed over B, and B prevailed over A, then in a simple social hierarchy, with ordered, transitive, relations, C should be dominant over *both* B and A. To find out whether infants could

use transitivity to infer dominance relations, Mascaro and Csibra (2012) conducted another experiment with 15-month-olds. In the familiarization event, A was pushed away from the enclosure by B, who was in turn pushed away by C. The test events employed the object-collection conflict situation; different groups of infants saw expected and unexpected test events involving A and B, B and C, or (to assess transitivity) A and C. Replicating the results of the last experiment, the infants who saw test events involving the A-B or the B-C pair expected the dominant figure in the pair to again prevail. In contrast, the infants who saw test events involving the novel A-C pair looked about equally at the expected and unexpected events, suggesting that they had no particular expectation as to whether A or C should prevail and collect the third object. Thus, although 15-month-olds can represent two separate dominance relations, they seem unable to order these relations into a hierarchy.

What might be the source of infants' difficulty with transitive dominance relations? One possibility is that infants in the second year of life still lack the cognitive resources necessary to order dominance relations, perhaps due to neurological immaturity. Another possibility is that infants can order dominance relations, but only when all of the individuals involved (e.g., A, B, and C) clearly belong to the same social group. Future research can evaluate these possibilities. In addition, research is needed to assess whether infants hold general expectations consistent with the authority principle about dominant and subordinate individuals. For example, would infants expect the dominant individual within a group to main order via third-party justice?

2-VI. DO INFANTS POSSESS A SENSE OF JUSTICE AND EXPECT INDIVIDUALS WHO VIOLATE SOCIOMORAL PRINCIPLES TO BE PUNISHED?

If infants expect individuals to act in accordance with principles of reciprocity, fairness, and ingroup loyalty and support, as we have argued, then what happens when infants witness

actions that do not adhere to the principles? If the principles specify what is obligatory or forbidden in social interactions, it should be the case that deliberate violations carry social costs. At least two sets of issues are pertinent here. One set concerns infants' *sociomoral evaluation* of wrongdoers. After seeing a wrongdoer produce an unacceptable action in one situation, do infants attribute to the wrongdoer a "bad" disposition and expect further unacceptable actions in other situations? Moreover, if the wrongdoer belongs to a novel social group, are infants likely to form a negative stereotype and generalize this attribution to other members of the group (e.g., Bigler & Liben, 2006, 2007)? Another set of issues concerns infants' expectations about the punishment of wrongdoers by third parties—in other words, about *third-party justice*. After seeing a wrongdoer mistreat a victim, do infants view negative actions directed at the wrongdoer by a third party as permissible? If yes, what factors modulate infants' expectations about who can exact punishments, what are appropriate punishments, and so on?

Researchers are just beginning to explore these issues with infants (e.g., Hamlin et al., 2011; Sloane & Baillargeon, 2012b; Vaish, Missana, & Tomasello, 2011). For example, Hamlin et al. (2011) built on results by Vaish, Carpenter, and Tomasello (2010) showing that 3-year-olds were more likely to give a toy to a bystander than to an antisocial adult. Hamlin et al. tested 19- to 23-month-olds using a procedure that included a training phase, a familiarization phase, and a test phase. During the training phase, an experimenter trained the infants to give "treats" (foam blocks) to five stuffed dogs by placing a treat in each dog's bowl. During the familiarization phase, the infants saw live help and hinder events on alternate trials for a total of six trials. For half the infants, a tiger puppet attempted unsuccessfully to retrieve a toy from a closed transparent box, while two dog puppets (one in a yellow shirt and one in a blue shirt) looked on; the helper dog (e.g., the one in the yellow shirt) helped the tiger open the box, and the hinderer

dog (e.g., the one in the blue shirt) jumped on the box. For the other infants, the tiger played with a ball and then dropped it; the helper dog returned the ball to the tiger, and the hinderer dog stole the ball. Finally, in the test phase, the experimenter from the training phase returned. She placed the two dogs on the apparatus floor, with a bowl in front of each dog. In the *give* condition, she handed the infants one treat, told them it was the last one left, and asked them to give it to one of the dogs. In the *take* condition, she placed a treat in each dog's bowl, introduced a new stuffed dog, and then asked the infants to take a treat away from one of the other dogs and give it to the new dog. In the give condition, the infants were reliably more likely to give the last remaining treat to the helper than the hinderer; in the take condition, the infants were reliably more likely to take a treat away from the hinderer than the helper. Because it is unclear whether the infants' responses reflected sociomoral evaluations (e.g., they sought to punish the hinderer for its unacceptable actions) or affiliative evaluations (e.g., they liked the helper more), we cannot draw firm conclusions about early expectations about justifiable punishment or third-party justice. Nevertheless, these results suggest a helpful method for exploring these issues in future research.

2-VII. SUMMARY AND FUTURE DIRECTIONS

The evidence reviewed in the second part of this chapter is consistent with the suggestions that (1) our human capacity for intuitive moral reasoning is an evolved adaptation, and (2) this capacity depends on a small set of abstract sociomoral principles that specify how individuals should act toward others in a variety of situations. In this chapter, we reviewed preliminary evidence that, at least by the second year of life, principles of reciprocity, fairness, ingroup, and authority affect infants' interpretations of individuals' actions as well as (in a few cases at least) infants' own responses to the individuals producing these actions.

Although tantalizing, the evidence we have reviewed on early sociomoral reasoning

remains critically limited in several ways. First, additional evidence is needed to define each principle more precisely and to better understand the expectations that it conveys. Second, as researchers develop paradigms for examining each principle in older infants and toddlers, it will be important to adapt these paradigms to test infants in the first year of life. Evidence that young infants share, or do not share, the same sociomoral expectations as older infants will provide a useful starting point for exploring the contributions of experience to the emergence and development of these expectations.

Third, although our review was largely silent on this issue, there can be no doubt that emotions contribute in varied ways to sociomoral reasoning, even in infants. For example, some scenes may innately evoke strong emotional responses (e.g., infants may perceive harm directed at the ingroup as aversive); emotions expressed by participants in a scene may influence infants' interpretation of the scene (e.g., if A laughs when B hits him, infants may view B's action as playful rather than as harmful); individual differences in empathic concern may affect infants' responses to distressed or victimized individuals; and socio-cultural differences may arise early in development as emotional responses are shaped by each culture to reinforce its rank-orderings of principles and other sociomoral choices.

Finally, it will be interesting to explore how early sociomoral expectations contribute to the socialization processes that help young children conform to societal norms. To illustrate, consider the disciplinary practice of induction, in which children are given explanations or reasons for changing their behavior (e.g., Hoffman, 2000). Inductions are thought to be particularly effective because they "focus children's attention on the consequences of their behavior for others, thereby capitalizing on children's capacity to empathize and experience guilt" (Eisenberg et al., 2006, p. 667). Another possibility, however, is that inductions are

effective because they capitalize on children's own sociomoral expectations and, in essence, remind them of what they know (e.g., keeping all the cookies for themselves is unfair to others).

Beyond what it reveals about this fascinating aspect of human development, the ongoing research on early sociomoral reasoning should have a significant impact on two fields of study. One is clinical psychology: If infants possess skeletal expectations about how ingroup members should interact, then marked deviations from these expectations (e.g., in neglectful or abusive households) might be particularly salient and, as such, might help explain the long-term consequences of early negative experiences. The other field is social psychology: Understanding early sociomoral expectations should help constrain theoretical models of adult moral cognition.

References

Aguiar, A., & Baillargeon, R. (2002). Developments in young infants' reasoning about occluded objects. *Cognitive Psychology*, *45*, 267-336.

Ainsworth, M. D., S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Hillsdale, NJ: Erlbaum.

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, *5*, 297-305.

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology*, *81*, 789-799.

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3,* 2-13.

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development, 62,* 1227-1246.

Baillargeon, R., He, Z., Setoh, P., Scott, R. M., Sloane, S., & Yang, D. Y.-J. (2013). False-belief understanding and why it matters: The social-acting hypothesis. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 88-95). New York, NY: Oxford University Press.

Baillargeon, R., Li, J., Gertner, Y., & Wu, D. (2011). How do infants reason about physical events? In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 11-48). Oxford, England: Blackwell.

Baillargeon, R., Li, J., Ng, W., & Yuan, S. (2009a). An account of infants' physical reasoning. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 66-116). New York,

NY: Oxford University Press.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*, 110-118.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition, 20,* 191-208.

Baillargeon, R., Stavans, M., Wu, D., Gertner, Y., Setoh, P., Kittredge, A., & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative account. *Language Learning and Development, 8,* 4-46.

Baillargeon, R., Wu, D., Yuan, S., Li, J., & Luo, Y. (2009b). Young infants' expectations about self-propelled objects. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 285-352). Oxford, England: Oxford University Press.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3-26.

Baldwin, D. A., & Moses, L. J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development, 10,* 309-329.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *13*, 513-552.

Baron-Cohen, S. (1995). *Mind blindness: An essay on autism and theory of mind*. Cambridge, MA : MIT Press/Bradford Books.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37-46.

Barragan R. C., & Dweck, C. S. (2012, June). *Is a norm of reciprocity necessary for*

young children to help? Paper presented at the Biennial International Conference on Infant Studies, Minneapolis, MN.

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences, 19,* 107-154.

Barrett, H. C., Broesch, T., Scott, R. M., He, Z., Baillargeon, R., Wu, D., Bolz, M., Henrich, J., Setoh, P., Wang, J., & Laurence, S. (2013). Early false-belief understanding in traditional non-Western societies. *Proceedings of the Royal Society, B (Biological Sciences), 280,* 20122654.

Bartsch, K. (1996). Between desires and beliefs: Young children's action predictions. *Child Development, 67,* 1671-1685.

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*, 492-498.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology*, *41*, 328–337.

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12-and 18-month-olds. *Infant Behavior and Development*, *22*, 277-282.

Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior*, *28*, 168-175.

Bigler, R. S., & Liben, L. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol., 34, pp. 39-89). San Diego, CA: Elsevier.

Bigler, R. S., & Liben, L. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science,*

16, 162-166.

Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science, 14,* 283-286.

Bíró, S., Csibra, G., & Gergely, G. (2007). The role of behavioral cues in understanding goal-directed action in infancy. *Progress in Brain Research, 164,* 303-322.

Bíró, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science, 10,* 379-398.

Bíró, S., Verschoor, S., & Coenen, L. (2011). Evidence for a unitary goal concept in 12-month-old infants. *Developmental Science, 14,* 1255-1260.

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

Blake, P. R., & Rand, D. G. (2010). Currency value moderates equity preference among young children. *Evolution and Human Behavior*, *31*, 210-218.

Bloom, P. & German, T. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77,* B25-B31.

Bosco, F. M., Friedman, O., & Leslie, A. M. (2006). Recognition of pretend and real actions in play by 1-and 2-year-olds: early success and why they fail. *Cognitive Development, 21*, 3-10.

Bower, T. G. R. (1977). *The perceptual world of the child.* Cambridge, MA: Harvard University Press.

Brandone, A. C., & Wellman, H. M. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science*, *20,* 85-91.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action

understanding: Inferential processes versus action simulation. *Current Biology, 17,* 2117-2121.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup Hate? *Journal of Social Issues*, *55*, 429-444.

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38,* 958-966.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*, 535-543.

Brownell, C. A., Svetlova, M., & Nichols, S. (2009). To share or not to share: When do toddlers respond to another's needs? *Infancy, 14,* 117-130.

Buresh, J. S., & Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition, 104,* 287-314.

Butler, S. C., Caron, A. J., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development*, *1*, 359-377

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false-belief understanding in an active helping paradigm. *Cognition*, *112*, 337-342.

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in-group over out-group members in 14-month-olds infants. *Child Development, 84,* 422-428.

Butterfill, S. A., & Apperly, I. A. (in press). How to construct a minimal theory of mind. *Mind and Language*.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55-72.

Byrne, D. E. (1971). *The attraction paradigm*. New York, NY: Academic Press.

Callaghan, T. C., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., Tapanya, S., & Singh, S. (2005). Synchrony in the onset of mental-state reasoning: Evidence from five cultures. *Psychological Science*, *16,* 378-384.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

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

Caron, A. J., Kiel, E. J., Dayton, M., & Butler, S. C. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition and Development*, *3*, 445-464.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*, 315-330.

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science, 8,* F13-F20.

Carpenter, M., Nagell, K., Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*, 1-143.

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

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language, 28,* 141-172.

Chen, M. L., & Waxman, S. R. (2013). "Shall we *Blick*?": Novel words highlight actors' underlying intentions for 14-month-old infants. *Developmental Psychology, 49,* 426-431.

Cheries, E. W., Wynn, K., & Scholl, B. J. (2006). Interrupting infants' persisting object representations: An object-based limit? *Developmental Science, 9,* F50-F58.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chow, V., Poulin-Dubois, D., & Lewis, J. (2008). To see or not to see: Infants prefer to follow the gaze of a reliable looker. *Developmental Science*, *11*, 761-770.

Chow, V., & Poulin-Dubois, D. (2009). The effect of a looker's past reliability on infants' reasoning about beliefs. *Developmental Psychology, 45*, 1576-1582.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377-395.

Cloutier, J., Gabrieli, J. D. L., O'Young, D., & Ambady, N. (2011). An fMRI study of violations of social expectations: When people are not who we expect them to be. *NeuroImage, 57,* 583-588.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *10*7, 705-717.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27,* 111-133.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13,* 148-153.

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition, 72,* 237–267.

Csibra, G. & Volein, A. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26,* 1-11.

Damon, W. (1975). Early conceptions of positive justice as related to the development of logical operations. *Child Development*, *46*, 301-312.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-months-old infants. *Developmental Psychology, 47,* 1499-1503.

Dennett, D. C. (1987). *The intentional stance.* Cambridge, MA: MIT Press.

D'Entremont, B., & Morgan, R. (2006). Experience with visual barriers and its effects on subsequent gaze-following in 12- to 13-month-olds. *British Journal of Developmental Psychology*, *24*, 465-475.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17,* 1037-1060.

Dunfield, K., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological science*, *21*, 523-527.

Dunfield, K. A., & Kuhlmeier, V. A. (in press). Classifying prosocial behavior: Children's responses to instrumental need, emotional distress, and material desire. *Child Development.*

Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, *16*, 227-247.

Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences*, *12*, 248-253.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, *82*, 793-811.

Dunphy-Lelii, S., & Wellman, H. M. (2004). Infants' understanding of occlusion of others' line-of-sight: Implications for an emerging theory of mind. *European Journal of Developmental Psychology*, *1*, 49-66.

Dupoux, E., & Jacob, P. (2007). Universal moral grammar: A critical appraisal. *Trends in Cognitive Sciences, 9*, 373-378.

Dwyer, S. (2006). How good is the linguistic analogy? In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Vol. 2. Culture and cognition* (pp. 237–256). Oxford, England: Oxford University Press.

Dwyer, S. (2009). Moral dumbfounding and the linguistic analogy: Methodological implications for the study of moral judgment. *Mind & Language, 24,* 274-296.

Egyed, K., Király, I., & Gergely, G. (in press). Communicating shared knowledge in infancy. *Psychological Science*.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 646-718). New York, NY: Wiley.

Fawcett, C., & Liszkowski, U. (2012). Infants anticipate others' social preferences. *Infant and Child Development*, *21*, 239-249.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*, 1079-1083.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives, 14,* 159-181.

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

Fiske, A. P. (2004). Relational models theory 2.0. In N. Haslam (Ed.), *Relational models theory: A contemporary overview* (pp. 3-25). Mahwah, NJ: Erlbaum.

Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

Fry, D. P. (2006). Reciprocity: The foundation stone of morality. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 399-422). Mahwah. NJ: Erlbaum.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science, 14,* 79-106.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183-209.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14,* 1012-1020.

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

Gergely, G. & Csibra, G. (1997). Teleological reasoning in infancy: The infant's naïve theory of rational action—A reply to Premack and Premack. *Cognition, 63,* 227-233.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in Cognitive Sciences*, *7*, 287-292.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science, 10,* 139-146.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56,* 165-193.

Gintis, H., Henrich, J., Bowles, S., Boyd, R., & Fehr, E. (2008). Strong reciprocity and the roots of human morality. *Social Justice Research*, *21*, 241-253.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26-37.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). New York, NY: Cambridge University Press.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review, 25,* 161-178.

Greene, J. (2005). Cognitive neuroscience and the structure of the moral mind. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Vol. 1. Structure and contents* (pp. 338-352). Oxford, England: Oxford University Press.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*, 1-20.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage, 21,* 744-750.

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

Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., & Hummel, A. (2010). Preschoolers' allocations in the dictator game: The role of moral emotions. *Journal of Economic Psychology*, *31*, 25-34.

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences USA, 107,* 9066-9071.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.

Haidt, J. (2008). Morality. *Perspectives on Psychological Science, 3*, 65-72.

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind*: *Vol. 3. Foundations and the future* (pp. 367-391). Oxford, England: Oxford University Press.

Hala, S., Hug, S., & Henderson, A. (2003). Executive function and false-belief understanding in preschool children: Two tasks are harder than one. *Journal of Cognition and Development, 4*, 275-298.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature, 476,* 328-331.

Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn K. (2013). Not like me = bad: Infants prefer those who harm dissimilar others. *Psychological Science, 24,* 589-594.

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development, 26,* 30-39.

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

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*, 923-929.

Hamlin, J. K., Wynn K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences USA, 108,* 19931-19936.

Hamlin, J. K., Wynn, K., & Bloom, P., (2012). The case for social evaluation in infants.

Response to Scarf et al. (2012). *PLoS ONE, 7,* e42698.

Hay, D. F., Caplan, M., Castle, J., & Stimson, C. A. (1991). Does sharing become increasingly "rational" in the second year of life? *Developmental Psychology, 27,* 987-993.

He, Z., & Baillargeon, R. (2012, June). *Young infants attribute to others knowledge based on memory and on inference.* Paper presented at the Biennial International Conference on Infant Studies, Minneapolis, MN.

He, Z., & Baillargeon (2013). Attributing false beliefs in the first year of life. Manuscript in preparation.

He, Z., Baillargeon, R., & Premack, D. (2011a, March). *Expectations about co-retaliation in infants and toddlers*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada.

He, Z., Bolz, M., & Baillargeon, R. (2011b). False-belief understanding in 2.5-year-olds: Evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*, *14*, 292-305.

He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, *30*, 14-29.

He, Z., Jin, K., Baillargeon, R., & Premack, D. (2013, April). *Infants' expectations about reciprocation and retaliation: Valence matters but form does not.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Seattle, WA.

Henderson, A. M. E., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental Science, 15,* 641-652.

Hernik, M., & Southgate, V. (2012). Nine-month-old infants do not need to know what the agent prefers in order to reason about its goals: On the role of preference and persistence in infants'

goal-attribution. *Developmental Science, 15,* 714-722.

Hespos, S. J., & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition*, *99,* B31-B41.

Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infants' perception of goal-directed actions performed by a mechanical device. *Infant Behavior and Development, 28,* 466-480.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York, NY: Cambridge University Press.

Hohenberger, A., Elsabbagh, M., Serres, J., de Schoenen, S., Karmiloff-Smith, A. & Aschersleben, G. (2012). Understanding goal-directed human actions and physical causality: The role of mother-infant interaction. *Infant Behavior and Development, 35,* 898-911.

Hoicka, E., & Wang, S. (2011). Fifteen-month-old infants match vocal cues to intentional actions. *Journal of Cognition and Development, 12,* 299-314.

Hood, B. M. (1995). Gravity rules for 2- to 4-year-olds? *Cognitive Development, 10,* 577-598.

Hook, J. G., & Cook, T. D. (1979). Equity theory and the cognitive ability of children. *Psychological Bulletin*, *86*, 429-445.

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

Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: MIT Press.

Jacob, P. (in press). A puzzle about belief-ascription. In B. Kaldis (Ed.), *Mind and Society: Cognitive Science Meets the Philosophy of the Social Sciences*. Berlin, Germany: Synthese Philosophy Library, Springer.

Jin, K., Houston, J., Baillargeon, R., Roisman, G. I., Sloane, S., & Groh, A. M. (2012, June). *8-*

month-olds expect an adult to respond to a crying but not to a laughing infant. Paper presented at the Biennial International Conference on Infant Studies, Minneapolis, MN.

Johnson, S. C. (2003). Detecting agents. *Philosophical Transactions of the Royal Society*, *358*, 549-559.

Johnson, S. C. (2005). Reasoning about intentionality in preverbal infants. In P. Carruthers, S. Laurence, & S. Stich. (Eds.), *The innate mind: Vol. 1. Structure and contents* (pp. 254-271). Oxford, England: Oxford University Press.

Johnson, S. C., Bolz, M., Carter, E., Mandsanger, J., Teichner, A., & Zettler, P. (2008). Calculating the attentional orientation of an unfamiliar agent in infancy. *Cognitive Development, 23*, 24-37.

Johnson, S. C., Booth, A., & O'Hearn, K. (2001). Inferring the goals of a nonhuman agent. *Cognitive Development*, *16*, 637-656.

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

Johnson, S. C., Dweck, C. S., Chen, F. S., Stern, H. L., Ok, S.-J., & Barth, M. E. (2010). At the intersection of social and cognitive development: Internal working models of attachment in infancy. *Cognitive Science, 34,* 807-825.

Johnson, S. C., Shimizu, Y. A. & Ok, S.-J. (2007). Actors and actions: The role of agent behavior in infants' attribution of goals. *Cognitive Development, 22*, 310-322.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development*, *20*, 303-320.

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources

with others. *PloS ONE*, *7*, e43979.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

Kärtner, J., Keller, H., & Chaudhary, N. (2010). Cognitive and social influences on early prosocial behavior in two sociocultural contexts. *Developmental Psychology, 46,* 905-914.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology, 15,* 483-524.

Kim, M., & Song, H. (2008). Korean 9- and 7-month-old infants' ability to encode the goals of others' pointing actions. *The Korean Journal of Developmental Psychology, 21,* 41-61.

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

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2012). "Native" objects and collaborators: Infants' object choices and acts of giving reflect favor for native over foreign speakers. *Journal of Cognition & Development*, *13*, 67-81.

Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology, 40,* 581-592.

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

Knudsen, B., & Liszkowski, U. (2012a). Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science, 15,* 113-122.

Knudsen, B., & Liszkowski, U. (2012b). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy, 17,* 672-691.

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true

and false statements. *Psychological Science*, *15*, 694-698.

Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, *87*, 179-208.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*, 1261-1277.

Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science, 12,* 48-54.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*, 1830-1834.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological science*, *14*, 402-408.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21,* 1134-1140.

Lamb, S., & Zakhireh, B. (1997). Toddler's attention to the distress of peers in a daycare setting. *Early Education & Development, 8,* 105–118.

Lane, I. M., & Coon, R. C. (1972). Reward allocation in preschool children. *Child Development*, *43*, 1382-1389.

Lee, V., & Kuhlmeier, V. A. (2013). Young children show a dissociation in looking and pointing behavior in falling events. *Cognitive Development, 28,* 21-30.

Lempers, J. D., Flavell, E. R., & Flavell, J. H. (1977). The development in very young children of tacit knowledge concerning visual perception. *Genetic Psychology Monographs, 95,* 3-53.

Lerner, M. J. (1974). The justice motive: "Equity" and "parity" among children. *Journal of Personality and Social Psychology*, *29*, 539-550.

Leslie, A. M. (1984). Infant perception of a manual pick-up event. *British Journal of Developmental Psychology, 2,* 19-32.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review, 94,* 412-426.

Leslie, A. M. (1994a). Pretending and believing: Issues in the theory of ToMM. *Cognition, 50,* 211-238.

Leslie, A. M. (1994b). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-148). New York, NY: Cambridge University Press.

Leslie, A. M. (2000). How to acquire a "representational theory of mind." In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 197-223). Oxford, England: Oxford University Press.

Leslie, A. M., & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science, 1,* 247-253.

Leventhal, G. S., Popp, A. L., & Sawyer, L. (1973). Equity or equality in children's allocation of reward to other persons? *Child Development*, *44*, 753-763.

Levitt, M. J., Weber, R. A., Clark, M. C., & McDonnell, P. (1985). Reciprocity of exchange in toddler sharing behavior. *Developmental Psychology*, *21*, 122-123.

Lewis, C., & Osborne, A. (1990). Three-year-olds' problems with false belief: Conceptual deficit or linguistic artifact? *Child Development, 61,* 1514-1519.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7,* 297-307.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). Twelve- and 18-month-olds

point to provide information for others. *Journal of Cognition and Development, 7,* 173-187.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Pointing out new news, old news, and absent referents at 12 months of age. *Developmental Science, 10,* F1-F7.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition, 108,* 732-739.

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

LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S., & Haidt, J. (2011). When getting something good is bad: Even three-year-olds react to inequality. *Social Development, 20*, 154-170.

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development*, *81*, 597-615.

Low, J. & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science, 24,* 305-311.

Luo, Y. (2010). Do 8-month-old infants consider situational constraints when interpreting others' gaze as goal-directed action? *Infancy, 15,* 392-419.

Luo, Y. (2011a). Do 10-month-old infants understand others' false beliefs? *Cognition*, *121*, 289-298.

Luo, Y. (2011b). Three-month-old infants attribute goals to a non-human agent. *Developmental Science, 14,* 453-460.

Luo, Y., & Baillargeon, R. (2005a). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16,* 601-608.

Luo, Y. & Baillargeon, R. (2005b). When the ordinary seems unexpected: Evidence for incremental physical knowledge in young infants. *Cognition, 95*, 297-328.

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, *105*, 489-512.

Luo, Y., & Beck, W. (2010). Do you see what I see? Infants' reasoning about others' incomplete perceptions. *Developmental Science, 13,* 134-142.

Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science, 12,* 142-149.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about events involving inert and self-propelled objects. *Cognitive Psychology, 58,* 441-486.

Luo, Y., Markson, L., Hennefield, L., Mou, Y., & vanMarle, K. (2013). Understanding agents' preferences at 8.5 months of age. Manuscript in preparation.

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

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

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development, 45,* 357-366.

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

Mason, R. A., & Just, M. A. (2009). The role of the theory-of-mind cortical network in the comprehension of narratives. *Language and Linguistics Compass, 3,* 157-174.

Master, A. & Walton, G. M. (2013). Minimal groups increase young children's motivation and learning on group-relevant tasks. *Child Development, 84,* 737-751.

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology, 24*, 470-476.

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. (2002). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 6-25). Oxford, England: Blackwell.

Meltzoff, A. N. (2005). Imitation and other minds: The "like me" hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science* (Vol. 2, pp. 55-77). Cambridge, MA: MIT Press.

Meltzoff, A. N., & Brooks, R. (2001). "Like Me" as a building block for understanding other minds: bodily acts, attention, and intention. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp.171-191). Cambridge, MA: MIT Press.

Meltzoff, A. N., & Brooks, R. (2007). Eyes wide shut: The importance of eyes in infant gaze-following and understanding other minds. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze following: Its development and significance* (pp. 217-241). Mahwah, NJ: Erlbaum.

Meristo, M., Morgan, G., Geraci, A., Iozzi, L., Hjelmquist, E., Surian, L., & Siegal, M. (2012). Belief attribution in deaf and hearing infants. *Developmental Science, 15*, 633-640.

Meristo, M., Strid, K., & Surian, L. (2012, June). *Preverbal infants expect agents to perform egalitarian distributions of resources*. Paper presented at the Biennial International Conference on Infant Studies, Minnesota, MN.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in cognitive sciences*, *11*, 143-152.

Mitchell, P., & Lacohée, H. (1991). Children's early understanding of false belief. *Cognition, 39,* 107-127.

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*, F1-F9.

Morissette, P., Ricard, M., & Décarie, T. G. (1995). Joint visual attention and pointing in infancy: A longitudinal study of comprehension. *British Journal of Developmental Psychology, 13,* 163-175.

Moses, L. J. (2001). Executive accounts of theory of mind development. *Child Development, 72,* 688-690.

Mueller, V. A., Brass, M., Waszak, F., & Prinz, W. (2007). The role of the preSMA and the rostral cingulated zone in internally selected actions. *NeuroImage, 37,* 1354-1361.

Needham, A. (1998). Infants' use of featural information in the segregation of stationary objects. *Infant Behavior and Development, 21,* 47-76.

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition, 47,* 121-148.

Newcombe, N., Huttenlocher, J., & Learmonth, A. (1999). Infants' coding of location in continuous space. *Infant Behavior and Development, 22,* 483-510.

Olineck, K. M., & Poulin-Dubois, D. (2005). Infants' ability to distinguish between intentional and accidental actions and its relation to internal state language. *Infancy*, *8*, 91-100.

Olineck, K. M., & Poulin-Dubois, D. (2009). Infants' understanding of intention from 10 to 14 months: Interrelations among violation of expectancy and imitation tasks. *Infant Behavior*

and Development, *32*, 404-415.

Olson, K. R., Banaji, M. R., Dweck, C. S., & Spelke, E. S. (2006). Children's biased evaluations of lucky versus unlucky people and their social groups. *Psychological Science, 17,* 845-846.

Olson, K. R., Dunham, Y., Dweck, C. S., Spelke, E. S., & Banaji, M. R. (2008). Judgments of the lucky across development and culture. *Journal of Personality and Social Psychology, 94,* 757-776.

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

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

Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica, 124*, 106-128.

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

Patterson, M. M., & Bigler, R. S. (2006). Preschool children's attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development*, *77*, 847-860.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011a). Bridging the gap between the other and me: The functional role of motor resonance and action effects in infants' imitation. *Developmental Science, 14*, 901-910.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011b). Imitation in infancy: Rational or motor resonance? *Child Development, 82*, 1047-1057.

Paulus, M., Hunnius, S., van Wijngaarden, C., Vrins, S., van Rooij, I., & Bekkering, H. (2011c).

The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Developmental Psychology, 47*, 976-983.

Perner, J. (2005). The many faces of belief: Reflections on Fodor's and the child's theory of mind. *Cognition, 57,* 241-269.

Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P.A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science* (Vol. 1, pp. 241-261). Hove, England: Psychology Press.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125-137.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences, 16,* 519-525.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, *98*, 137-155.

Phillips, A. T., Wellman, H. A., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition, 85,* 53-78.

Piaget, J. (1954). *The construction of reality in the child.* New York, NY: Routlege.

Pinkham, A. M., & Jaswal, V. K. (2011). Watch and learn? Infants privilege efficiency over pedagogy during imitative learning. *Infancy*, *16*, 535-544.

Poulin-Dubois, D., Brooker, I., & Polonia, A. (2011). Infants prefer to imitate a reliable person. *Infant Behavior and Development*, *34*, 303-309.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36,* 1-16.

Premack, D. (2007). Foundations of morality in the infant. In O. Vilarroya & F. I. Argimon

(Eds.), *Social brain matters: Stances on the neurobiology of social cognition* (pp. 161-167). Amsterdam, The Netherlands: Rodopi.

Premack, D., & Premack, A. J. (1995). Origins of human social competence. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 205-218). Cambridge, MA: MIT Press.

Premack, D., & Premack, A. J. (1997). Infants attribute value +/- to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience*, *9*, 848-856.

Premack, D., & Premack, A. J. (2003). *Original intelligence: Unlocking the mystery of who we are*. New York, NY: McGraw-Hill.

Radke-Yarrow, M. R., Zahn-Waxler, C., & Chapman, M. (1983). Children's prosocial dispositions and behavior. In P. H. Mussen (Series Ed.) & E. M. Hetherington (Vol. Ed.), *Carmichael's manual of child psychology: Vol.4. Socialization, personality, and social development* (4th ed., pp. 469-546). New York, NY: Wiley.

Repacholi, B. M. (1998). Infants' use of attentional cues to identify the referent of another person's emotional expression. *Developmental Psychology*, *34*, 1017-1025.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33,* 12-21.

Repacholi, B. M., Meltzoff, A. N., & Olsen, B. (2008). Infants' understanding of the link between visual perception and emotion: "If she can't see me doing it, she won't get angry." *Developmental Psychology*, *44*, 561-574.

Rheingold, H. L. (1982). Little children's participation in the work of adults, a nascent prosocial behavior. *Child Development*, *53*, 114-125.

Robson, S. J., & Kuhlmeier, V. A. (2013, April). *Selectivity promotes 9-month-old infants to encode the goals of others.* Paper presented at the Biennial Meeting of the Society for

Research in Child Development, Seattle, WA.

Rochat, P., Dias, M. D. G., Liping, G., Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*, 416-442.

Roth, D., & Leslie, A. M. (1998). Solving belief problems: Toward a task analysis. *Cognition, 66,* 1-31.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, *80*, 201-224.

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology*, *30*, 87-104.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9,* 331-349.

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience, 1,* 284-298.

Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science, 16,* 995-1001.

Saxe, R., Tzelnic, T., & Carey, S. (2006). Five-month-old infants know humans are solid, like inanimate objects. *Cognition, 101,* B1-B8.

Saxe, R., Tzelnic, T., & Carey, S. (2007). Knowing who-dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology, 43,* 149-158.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia, 43,* 1391-1399.

Scarf, D., Imuta, K., Colombo, M., & Hayne, H. (2012). Social evaluation or simple association? Simple associations may explain moral reasoning in infants. *PloS ONE*, *7*, e42698.

Schlottmann, A., & Ray, E. D. (2010). Goal attribution to schematic animals: Do 6-month-olds perceive biological motion as animate? *Developmental Science, 13,* 1-10.

Schlottmann, A., Ray, E. D., & Surian, L. (2012). Emerging perception of causality in action-and-reaction sequences from 4 to 6 months of age: Is it domain-specific? *Journal of Experimental Child Psychology, 112,* 208-230.

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, *6*, e23223.

Schwier, C., van Maanen, C., Carpenter, M., & Tomasello, M. (2006). Rational imitation in 12-month-old infants. *Infancy*, *10*, 303-311.

Scott, R. M. & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development, 80,* 1172-1196.

Scott, R. M., & Baillargeon, R (2013). Do infants really expect others to act efficiently? A critical test of the rationality principle. *Psychological Science, 24,* 466-474.

Scott, R. M., Baillargeon, R., Song, H. & Leslie, A. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology*, *61*, 366-395.

Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science*, *15*, 181-193.

Scott, R. M., Richman, J., & Baillargeon, R (2013). Infants understand the art of deception:

New evidence for mentalistic reasoning in infancy. Manuscript under review.

Scott, R. M., & Setoh, P. (2012, June). *Nature of response practice affects 2.5-year-olds' performance in elicited-response false-belief tasks.* Paper presented at the Biennial International Conference on Infant Studies, Minneapolis, MN.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science, 22,* 878-880.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous Theory of Mind in Asperger syndrome. *Science, 325,* 883-885.

Setoh, P., & Baillargeon, R. (2010, March). *Infants generalize food preferences across novel animals based on shape.* Paper presented at the Biennial International Conference on Infant Studies, Baltimore, MD.

Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences, USA, 110,* 15937-15942.

Setoh, P., Scott, R. M., & Baillargeon, R. (2011, April). *False-belief reasoning in 2.5-year-olds: Evidence from an elicited-response low-inhibition task.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada.

Shimizu, Y. A., & Johnson, S. C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science, 7,* 425-430.

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*, 1-17.

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

35-62.

Sigmund, K., Fehr, E., & Novak, M. A. (2002). The economics of fair play. *Scientific American, 286*, 82-87.

Sloane, S., & Baillargeon, R. (2012a, June). *Expectations of fairness in 9-month-old infants*. Paper presented at the Biennial International Conference on Infant Studies, Minneapolis, MN.

Sloane, S., & Baillargeon, R. (2012b, June). *21-month-olds expect individuals to punish others fairly*. Paper presented at the Biennial International Conference on Infant Studies, Minneapolis, MN.

Sloane, S., & Baillargeon, R. (2013, October). *Charity Begins and Ends at Home: Developmental Changes in Ingroup Favoritism in Infancy*. Paper presented at the Biennial Meeting of the Cognitive Development Society, Memphis, TN.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*, 196-204.

Sloane, S., Baillargeon, R., & Premack, D. (2013). Infants expect ingroup favoritism to trump fairness. Manuscript in preparation.

Sodian, B., Schoeppner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents? *Infant Behavior and Development*, *27*, 31-41.

Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to goal-directed action. *Journal of Cognition and Development, 5,* 289-316.

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, *10*, 199-204.

Sommerville, J. A., Schmidt, M. F. H., Yun, J., & Burns, M. (2013). The development of

fairness expectations and prosocial behavior in the second year of life. *Infancy, 18,* 40-66.

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

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

Song, H., & Baillargeon, R. (2007). Can 9.5-month-old infants attribute to an agent a disposition to perform a particular action on objects? *Acta Psychologica, 124,* 79-105.

Song, H., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology*, *44,* 1789-1795.

Song, H., Baillargeon, R., & Fisher, C. (2005). Can infants attribute to an agent a disposition to perform a particular action? *Cognition, 98,* B45-B55.

Song, H., Baillargeon, R., & Fisher, C. (in press). The development of infants' use of novel verbal information when reasoning about others' actions. *PLoS ONE*.

Song, H., Onishi, K. H., Baillargeon, R., & Fisher, C. (2008) Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition*, 109, 295-315.

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*, 907-912.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology, 45,* 1794-1798.

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition*, *107*, 1059-1069.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false

belief by two-year-olds. *Psychological Science, 18*, 587-592.

Spaepen, E. & Spelke, E. (2007). Will any doll do? 12-month-olds' reasoning about goal objects. *Cognitive Psychology, 54,* 133-154.

Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition, 50,* 431-445.

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

Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44-78). Oxford, England: Clarendon Press.

Sripada, C., & Stich, S. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Vol. 2. Culture and cognition* (pp. 280-301). Oxford, England: Oxford University Press.

Sunnafrank, M. (1983). Attitude similarity and interpersonal attraction in communication processes: In pursuit of an ephemeral influence. *Communication Monographs, 50,* 273-284.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*, 580-586.

Surian, L., & Geraci, A. (2012). Where will the triangle look for it? Attributing false beliefs to geometric shapes at 17 months. *British Journal of Developmental Psychology, 30,* 30-44.

Svetlova, M., Nichols, S. R., & Brownell, C. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child Development*, *81*, 1814-1827.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*, 149-178.

Tamir, D. I., & Mitchell, J. P. (2010). Neural correlates of anchoring-and-adjustment during

mentalizing. *Proceedings of the National Academy of Sciences USA, 107,* 10827-10832.

Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology, 30,* 172-187.

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

Thomson, N. R., & Jones, E. (2005). Children's, adolescents', and young adults' reward allocations to hypothetical siblings and fairness judgments: Effects of actor gender, character type, and allocation pattern. *The Journal of Psychology*, *139*, 349-368.

Tomasello, M. (1999). Having intentions, understanding intentions, and understanding communicative intentions. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 63-75). Mahwah, NJ: Erlbaum.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675-690.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12-and 18-month-olds know what is new for other persons. *Developmental Psychology*, *39*, 906-912.

Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy, 15,* 434-444.

Trevarthen, C. (1998). The concept and foundations of infant intersubjectivity. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 15-46). Cambridge: Cambridge University Press.

Trivers, R. L. (2006). Reciprocal altruism: 30 years later. In P. M. Kappeler & C. P. van Schaik

(Eds.), *Cooperation in primates and humans: Mechanisms and evolution* (pp. 67-83). Berlin, Germany: Springer.

Tsutsu, K. (2010). Influences on the amount of the reward: How five-year-old children distribute rewards. *Shinrigaku kenkyu: The Japanese Journal of Psychology*, *81*, 201-209.

Turiel, E. (2006). The development of morality. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 789-857). New York, NY: Wiley.

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective-taking and its relation to prosocial behavior in toddlers. *Developmental Psychology, 45*, 534-543.

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development, 81,* 1661-1669.

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *British Journal of Developmental Psychology, 29,* 124-130.

Vaughan, G. M., Tajfel, H., & Williams, J. (1981). Bias in reward allocation in an intergroup and an interpersonal context. *Social Psychology Quarterly*, *44*, 37-42.

Vinden, P. G. (1999). Children's understanding of mind and emotion: A multi-culture study. *Cognition and Emotion, 13*, 19-48.

Vosniadou, S. & Brewer, W. F. (1992). Mental models of the earth: a study of conceptual change in childhood. *Cognitive Psychology, 24,* 535-585.

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition, 93*, 167-198.

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

Warneken, F., Hare, B., Melis, A. P., Hanus, D., & Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young Children. *PLoS Biology*, *5*, e184.

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

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*, 271-294.

Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology*, *44*, 1785-1788.

Warneken, F., & Tomasello, M. (2009). The roots of human altruism. *British Journal of Psychology*, *100*, 455-471.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30*, 239-277.

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

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition, 72,* 125-166.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103-128.

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

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful

behaviors. *Infant Behavior and Development, 22,* 145-160.

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science, 6,* 297-311.

Woodward, A. L. (2005). The infant origins of intentional understanding. In R. V. Kail (ed.) *Advances in Child Development and Behavior* (Vol. 33, pp. 229-262). Oxford, England: Elsevier.

Woodward, A. L. & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development, 17,* 1061-1084.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11,* 73-77.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Wynn, K. (2009). Constraints on natural altruism. *British Journal of Psychology*, *100*, 481-485.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology, 30,* 111-153.

Yang, D. Y., & Baillargeon, R. (in press). Difficulty in understanding social acting (but not false beliefs) mediates the link between autistic traits and ingroup relationships. *Journal of Autism and Developmental Disorders*.

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

Yott, J., & Poulin-Dubois, D. (2012) Breaking the rules: Do infants have a true understanding of false belief? *British Journal of Developmental Psychology, 30,* 156-171.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of

concern for others. *Developmental Psychology, 28*, 126-136.

Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology, 106,* 208-220.

Footnotes

1. We prefer the term psychological reasoning (also mentalizing, mindreading, or intuitive psychology) to the commonly used term "theory of mind" for three reasons. First, the implicit principles that guide infants' expectations about agents' actions are not tantamount to a folk theory of psychology (e.g., Leslie, 2000). Second, although children eventually acquire a folk theory of psychology (just as they acquire folk theories of biology and astronomy; e.g., Carey, 1985; Vosniadou & Brewer, 1992), it is as yet unclear what is the relationship between this late-developing, explicit folk theory of psychology and the implicit principles that from an early age guide psychological expectations. Finally, use of the term "theory of mind" has misled many researchers into viewing the acquisition of an explicit folk theory of psychology as the ultimate milestone of development in this domain. We question the wisdom of this assumption, just as we would question the notion that the acquisition of an explicit grammar is the ultimate goal of language acquisition.

2. In this chapter, we will use interchangeably the terms positive disposition, preference (in the sense of a habitual preference, as in a preference for jazz or ballroom dancing), predilection, and fondness to denote a relatively enduring tendency to favor a particular object or activity. The interpretation offered here is that, when an agent continually chooses object-A over object-B in a scene, infants take this unvarying choice information to signal a positive disposition toward object-A.

3. The preference-based interpretation advanced here predicts that, after watching an agent reach repeatedly for object-A as opposed to object-B in a scene, young infants should expect the agent to continue reaching for object-A in the test trials even if it was now paired with new object-C. Results from preliminary experiments by Robson and Kuhlmeier (2013) and from ongoing

experiments by Bian and Baillargeon support this prediction.

4. Research is also needed to determine how younger infants would perform in these one- and two-category tasks. Previous evidence indicates that infants in the first year of life do not spontaneously encode objects' taxonomic categories (e.g., Baillargeon et al., 2012; Xu & Carey, 1996); therefore, it is possible that early encodings of preferences refer to particular objects, rather than to taxonomic categories of objects. In line with this suggestion, Schlottmann and Ray (2009) obtained positive results with 6-month-olds in a two-object task with two distinct circles as the objects (in computer-animated events).

5. Two-object tasks in which the agent continually grasps, points to, or looks at object-A as opposed to object-B can also be described as demonstrating deviations from random sampling: Instead of randomly choosing between object-A and object-B, the agent always chooses object-A. To draw an analogy with the task of Kushnir et al. (2010), imagine that an agent is presented several times in a row with a box containing a duck and a frog, and each time removes the duck. In all of these cases, the deviation from random sampling signals a preference.

6. Csibra and his colleagues adapted their detour tasks to also explore infants' sensitivity to *consistency* constraints (e.g., Csibra et al., 2003; Southgate & Csibra, 2009). In one experiment, for example, 12-month-olds saw the same habituation event as in the last detour task (with the large red circle chasing the small yellow circle), but presented on a larger computer monitor (Csibra et al., 2003). In the test trials, the same event continued on with two different outcomes. In the *old-goal* event, the small circle stopped past the bars and the large circle came to rest against it; in the *new-goal* event, the small circle again stopped but the large circle now traveled past it off the monitor. The infants looked reliably longer at the new- than at the old-goal event, suggesting that (1) during the habituation trials, they attributed to the large circle the goal of

chasing the small circle, and (2) during the test trials, they viewed the new-goal event as unexpected because the large circle changed goal for no apparent reason, in violation of the consistency principle. These results extend those reviewed in section C. by showing that, by their first birthday, infants expect agents to act in a manner consistent with their goals as well as with their dispositions.

7. In the experiments described earlier on infants' use of equifinality information (e.g., Luo, 2011b; Song & Baillargeon, 2007), equifinality was equivalent to *context-dependent* variation: Contextual changes forced the agent to choose different means across trials to attain the same goal. Csibra's (2008) experiment, by contrast, used *choice-dependent* variation: Although the context did not change, the agent deliberately chose different means across trials to achieve the same goal. Research is needed to explore the links between these two forms of equifinality.

8. Given the results of Csibra (2008), readers may be puzzled as to why 6-month-olds failed in the detour task of Csibra et al. (1999), which was similar to that of Gergely et al. (1995), described in the text.. The most likely explanation is that the impoverished computer animations used in these early experiments did not provide sufficient visual cues (e.g., depth cues) to enable younger infants to correctly perceive the events.

9. From an adult perspective, these negative results may seem puzzling: Why didn't the infants simply conclude that the long-handle box, rod, arm-shaped screen, and mechanical claw were inanimate tools controlled by an unseen agent and used to signal a preference for object-A over object-B? There are indeed conditions under which even young infants seem to infer that they are facing objects manipulated by unseen agents, but this research lies beyond the scope of our chapter (e.g., Bíró, Csibra, & Gergely, 2007; Bíró & Leslie, 2007; Hofer, Hauf, & Aschersleben, 2005; Saxe, Tenenbaum, & Carey, 2005; Saxe, Tzelnic, & Carey, 2007).

10. Recent research on early biological reasoning provides converging evidence that self-propulsion and agency are distinct properties for infants. In a series of experiments, 8-month-olds expected a novel object that was *both* self-propelled and agentive—but not an object that lacked either of these properties—to have *insides*, suggesting that they viewed it as an animal and ascribed to it basic biological properties (Setoh, Wu, Baillargeon, & Gelman, 2013). Thus, self-propelled object, agent, and animal are all separate concepts for infants; each is embedded in a different explanatory system, and each carries a rich set of expectations.

11. When discussing the role of executive-function limitations in false-belief understanding, researchers often distinguish between two types of accounts (e.g., Carlson & Moses, 2001; Moses, 2001): *emergence* accounts, which fit within the non-representational view and assume that executive-function limitations prevent the emergence of false-belief understanding, and *expression* accounts, which fit within the representational view and assume that executive-function limitations prevent only the expression of children's false-belief understanding.

12. In another experiment using a choice task, Hamlin et al. (2007) found that at 6 and 10 months infants (1) were reliably *more* likely to prefer the helper than a neutral character whose actions were identical to those of the helper except that the climber remained at the bottom of the hill, and (2) were reliably *less* likely to prefer the hinderer than a neutral character whose actions were identical to those of the hinderer except that, here again, the climber remained at the bottom of the hill (at both ages, negative results were found in similar VOE tasks examining which of the two individuals infants expected the climber to approach). Hamlin et al. concluded that infants both preferred individuals who acted positively and avoided individuals who acted negatively. It would be interesting to see whether infants would still prefer the helper over a neutral character who interacted with the climber in some fashion (e.g., the neutral character and

climber took turns watching each other climb the hill). In the scene used by Hamlin et al., the neutral character ignored the climber, and its actions (as it moved back and forth on the hill) might have been perceived as odd or inefficient, leaving open the possibility that infants might value more equally positive and neutral characters whose actions were more comparable.

13. Surprisingly, Kuhlmeier et al. (2003) found that 12-month-olds looked reliably longer at the *approach-helper* than at the approach-hinderer event. This result is open to at least two possible interpretations. One is that these older infants not only detected the violation in the approach-hinderer event, but also detected a violation in the approach-helper event when the climber and helper stood side by side without interacting in any way. The other interpretation is that the test events used by Kuhlmeier at al. overwhelmed the infants' information-processing resources to such an extent that great effort was required to process even the expected, approach-helper event. Compared to the events used by Hamlin et al. (2007), those of Kuhlmeier et al. differed in several ways that might have rendered them harder to process: They were computer-animated rather than live; the helper, hinderer, and climber had no eyes; and there was no hill in the test events, making the scene less similar to that shown in the habituation events (at the start of each test event, the climber stood centered at the bottom of the computer monitor, and the helper and hinderer stood in the top corners; the climber approached and stopped next to either the helper or the hinderer).

14. Of course, infants will distinguish positive and negative actions only when they can infer the goals underlying these actions. Scarf, Imuta, Colombo, and Hayne (2012) reported data suggesting that the results of Hamlin et al. (2007) might be due to low-level factors, but they apparently failed to include in their replication many of the subtle cues Hamlin et al. had used to make clear to infants the goals of the climber, helper, and hinderer (see the response of Hamlin,

Wynn, and Bloom (2012) to the article of Scarf et al.).

15. Prior to the choice task, Geraci and Surian (2011) showed their 16-month-olds additional computer-animated events involving the observer and the two distributors. In the first two trials, the observer first stood at the bottom of the computer monitor at the entrance of a vertical Y-shaped tunnel; the observer moved out of view into the tunnel and came out through the left exit in one trial and through the right exit in the other trial. In the next two trials, the fair and unfair distributors stood at the exits; the observer entered the tunnel, exited next to one of the distributors, and stood near it until the trial ended. Although the infants failed to anticipate which distributor the observer would approach (perhaps because they had seen the observer approach both distributors during the equal and unequal events at the start of the testing session), they looked reliably longer when the observer stood near the *fair* (approach-fair) as opposed to the unfair (approach-unfair) distributor. Thus, as in the experiment of Kuhlmeier et al. (2003; see Footnote 12), the infants showed longer looking times at the expected event. Here again, this result is open to at least two interpretations. One is that these older infants detected a violation in the approach-fair event when the observer and fair distributor stood side by side without interacting in any way. The other interpretation is that, because the scene in the approach-fair and approach-unfair events differed markedly from that shown in the equal and unequal events, great effort was required for the infants to process even the expected, approach-fair event.

16. Evidence supporting these speculations comes from an experiment by Buttelmann, Zmyj, Daum, and Carpenter (2013). In each of four trials, 14-month-olds watched a videotaped event in which a bilingual man spoke in either their native language (German; ingroup condition) or a foreign language (Russian; outgroup condition), followed by another videotaped event in which the same man silently performed an unusual action on a novel object (e.g., activated a light-box

with his forehead; two imitation trials) or chose one of two toys (e.g., a cylinder or an octagon; two preference trials). The infants were then presented with real-life replicas of the object or toys, and their responses were recorded. A reliable difference between the two conditions was found in only one trial: 46% of the infants in the ingroup condition activated the light-box with their heads compared with 21% in the outgroup condition (the other infants did so with their hands). The level of imitation in the ingroup condition was rather low, however, compared to that typically found in this task (67-69%, as discussed earlier in the chapter). Together, these results suggest that simply watching an unfamiliar man speak their language (with no outgroup contrast) may not be sufficient to induce infants to view the man as an ingroup member. We return to the issue of infants' responses to unfamiliar as opposed to familiarized adults in our discussion of prosocial actions.

17. In a laudable attempt to make helping and comforting tasks more similar, Svetlova et al. (2010) tested infants with three out-of-reach helping scenarios modeled after those of Warneken and Tomasello (2006) and three novel out-of-reach comforting scenarios. At 18 months, 44% of the infants tested spontaneously produced appropriate responses in one or more of the helping scenarios, but only 13% did so in the comforting scenarios, perhaps due in part to difficulties in understanding the behavior of the distressed experimenter. In one scenario, for example, the experimenter first placed her teddy bear near the infant while explaining that it was her special toy and made her happy. She then sat on the floor some distance from the infant, was approached by an assistant who whispered something in her ear, and immediately began to sigh and sob. In order to succeed at the task, the infants had to infer that the experimenter was told something distressing, and that the way to comfort her was to bring her the teddy bear. Neither inference might have been obvious for the 18-month-olds, who might well have been perplexed by the

experimenter's sobbing.