

Infants and toddlers leverage their understanding of action goals to evaluate agents who help others

Brandon M. Woo^{1,2}  | Elizabeth S. Spelke^{1,2}

¹Department of Psychology, Harvard University, Cambridge, Massachusetts, USA

²The Center for Brains, Minds, and Machines, Cambridge, Massachusetts, USA

Correspondence

Brandon M. Woo, Department of Psychology, Harvard University, Cambridge, MA 02138, USA.
Email: brandon.matt.woo@gmail.com

Funding information

National Science Foundation, Grant/Award Number: CCF-1231216; Social Sciences and Humanities Research Council of Canada, Grant/Award Number: 752-2020-0474

Abstract

Why do infants and toddlers prefer helpers? Four experiments (conducted from 2019–2022; $n = 136$, 66% White, 15% Asian, 4% Black, 2% Hispanic/Latino, 13% multiracial, majority USA) investigated whether infants and toddlers favor agents whose actions allow others to achieve their goals. In the key experiment, 8-month-old infants and 15-month-old toddlers viewed a protagonist who tried and failed to open a box that contained a toy while two other agents (helpers) observed; then the toys were exchanged and the helpers opened different boxes. Infants and toddlers differently evaluated the two helpers, consistent with their developing means-end understanding. Together, the present four experiments connect infants' and toddlers' evaluations of helping to their understanding of goal-directed behavior.

Understanding others' goals is critical to navigating the social world (Baker et al., 2009; Woodward, 2009). Landmark studies by Hamlin et al. (2007) provided evidence that infants distinguish between agents who help versus hinder others in the pursuit of their goals. In these studies, infants viewed puppet show displays depicting protagonists who attempted to climb a hill (Hamlin et al., 2007, 2010), open a box containing a toy (Hamlin & Wynn, 2011; Steckler et al., 2017), or seek to recover a ball with which they had been playing (Hamlin & Wynn, 2011). The infants in these studies preferentially reached for and looked to agents whose actions brought about the goal states that the protagonist desired, by pushing the agent up a hill that it had struggled to climb, opening the box that it had struggled to open, or returning its ball (Margoni & Surian, 2018; Woo et al., 2022; cf. Schlingloff et al., 2020). Nevertheless, the basis of these evaluations is unclear, because other features distinguished the helping and hindering events, including the protagonist's expressions of positive affect only in helping events (Scarf et al., 2012), and one agent's greater imitation of or synchronization with the protagonist

in the helping events (Benton & Lapan, 2022; Powell & Spelke, 2018b). These differences raise the possibility that infants' preferences depended on the agents' levels of social engagement and interaction with the protagonist, rather than on whether the agents acted to bring about the protagonist's desired goal state. Do early evaluations of agents who help or fail to help an individual achieve a goal reflect an understanding that one agent has acted to foster the individual's goal?

Research provides evidence that adults favor individuals who behave prosocially by imitating others (e.g., Chartrand & Bargh, 1999), synchronizing their actions with others (e.g., Hove & Risen, 2009), or expressing positive affect (e.g., Van Kleef, 2009). Developmental research has revealed that infants and toddlers are also sensitive to these social cues (e.g., Powell & Spelke, 2018b; Scarf et al., 2012; Tunçgenç et al., 2015). To our knowledge, however, no study has clearly shown that infants and toddlers favor those who help others by leveraging their understanding of others' action goals, in situations in which these other cues to prosociality are controlled. In four experiments, we test the hypothesis that infants

Abbreviation: OSF, Open Science Framework

© 2023 The Authors. Child Development © 2023 Society for Research in Child Development.



and toddlers value helpers in situations that require inferences about the goal of the agent in need of help.

Developmental changes in hierarchical action planning and goal understanding

After observing others' actions, adults can infer the goals and action plans driving those actions (Baker et al., 2009, 2017). For example, after observing that a person opens a box and grasps a toy inside, adults can infer that the person sought to grasp the toy: They recognize such actions as means-end actions sequences, in which the intermediate action on one object (the opening of the box) advances the ultimate goal of acting on the second object (grasping the toy). In contrast, a rich literature reveals that this ability is not present in infants until late in the first year. If a desired toy can only be obtained by pulling on the blanket on which it sits (because it is out of reach), infants who are physically capable of pulling the blanket nevertheless fail to do so (Diamond, 1991; Gerson et al., 2015; Piaget, 1952; Sommerville & Woodward, 2005).

Further research has demonstrated a parallel developmental change in infants' inferences about the goals of others who engage in means-end actions. For example, if an actor opens a box that contains a toy, toddlers and older infants infer that the actor's goal is to access the toy, but younger infants do not (Sommerville & Woodward, 2005; Woodward & Sommerville, 2000; see also, Henderson et al., 2013; Henderson & Woodward, 2011). Although infants aged 5 to 10 months can infer the goal of a direct reaching action (in which an agent acts on one object; Woodward, 1998), they are more likely to infer that the goal of a means-end action is the means state (e.g., the opening of the box; Gerson et al., 2015; Sommerville & Woodward, 2005; see also, Gergely et al., 2002; Sommerville et al., 2008). Thus, there are striking limits to young infants' abilities to generate hierarchically structured action plans and to interpret the actions of others who generate such plans. Given this developmental change in means-end understanding, we chose to study evaluations of helping both before and after this change takes place in development. If infants and toddlers leverage their capacities to infer a protagonist's goal in evaluating helpers, then there should be a corresponding developmental change in their evaluations of the very same acts of helping.

The basis of early social evaluation

Although infants' evaluations of helpers are consistent with an understanding of a protagonist's goal (see Tan & Hamlin, 2022), infants may favor helpers for at least two other reasons. First, infants may prefer helpers because their actions elicit positive affect from others. In Hamlin et al.'s (2007, 2010) hill-climbing paradigm, for instance,

the protagonist jumped excitedly upon being helped. If infants see jumping as an expression of excitement or happiness, then they might prefer the helper because it was associated with an expression of positive affect. Consistent with that possibility, one study found that infants could be swayed to prefer a hinderer over a helper when a protagonist jumped after being hindered, but not after being helped (Scarf et al., 2012). Nevertheless, 6- to 11-month-old infants preferred helpers in the hill-climbing paradigm even when no excited jumping occurs (Hamlin, 2015).

Second, infants may prefer helpers over hinderers because the movements of helpers either imitate or synchronize with others' movements. In many studies of infants' evaluations of helping, the helper's actions resemble the actions of the agent whom they help, whereas the hinderer's actions do not. For example, after a protagonist had sought to climb a hill by itself in Hamlin et al.'s (2007, 2010) hill-climbing paradigm, the helper also moved up the hill, together with the protagonist, whereas the hinderer moved down the hill, opposing the protagonist. Likewise, after a protagonist had tried to lift the lid of a box (Hamlin & Wynn, 2011), the helper also lifted the box's lid, together with the protagonist, whereas the hinderer pushed the box's lid down, opposing the protagonist. Finally, after a protagonist had dropped its ball, causing it to move laterally (Hamlin & Wynn, 2011), the helper passed the ball laterally to the protagonist, whereas the hinderer moved away with the ball. Imitation and synchrony are social actions that signal capacities and motivations for social engagement, coordination, and shared intentions (Benton & Lapan, 2022; Powell & Spelke, 2018a, 2018b; Tomasello et al., 2005). Both imitation and synchrony impact social reasoning and behavior in infants, young children, and adults. Notably, infants prefer agents who imitate or synchronize their actions with them (Meltzoff, 1990; Tunçgenç et al., 2015). Infants' and toddlers' selective attention to and reaching for helpers therefore may depend on a preference for agents who coordinate their actions with others, rather than on their inferences about the protagonist's goals.

The present experiments test the role of goal understanding in early social evaluations, by disentangling helping from cues of affect, imitation, and synchrony. If early evaluations of helping reflect an understanding that an agent has fostered someone else's goal, then infants' and toddlers' evaluations may differ as a function of what actions they understand.

Goal understanding and social evaluation

We adapted a paradigm developed by Hamlin and Wynn (2011) in which a protagonist seeks to open a box containing a toy, is helped, and then grasps the toy. Hamlin and Wynn (2011) found that 5- and

9-month-old infants preferred a helper over a hinderer in this paradigm. Their findings do not contradict the literature showing that infants of this age fail to understand means-end actions, however, because infants' favoring of the helper could have occurred for any of three reasons. First, the findings can be explained by preferences for agents who engage in imitation and synchronous behavior, because the helper and the protagonist opened the box together. Second, the findings can be explained by an understanding of the goal of a more direct action: Infants may have viewed the opening of the box as helpful, because they inferred that the protagonist wanted the box to be open (i.e., the intermediate goal), without identifying the protagonist's ultimate goal of obtaining the toy inside it. Finally, young infants may understand the hierarchical structure of a means-end action in helping contexts, even though they fail to do so when acting on their own or when viewing the actions of a single agent. When one agent helps another agent, a single means-end action may be unpacked into its two component parts, because the helper and the protagonist each engage in a distinct, direct action: The helper opens the box, and the protagonist grasps the toy. If helping contexts unpack means-end actions for infants, then infants may better infer the protagonist's ultimate goal of grasping the toy.

Research by Hamlin et al. (2013) at first appears to support the third possibility. Ten-month-old infants first viewed a protagonist who repeatedly grasped one of two toys in a direct action while two agents observed, and then barriers were introduced that obstructed the protagonist's access to the two toys. Infants preferred the observer who removed a barrier to the protagonist's desired toy, allowing the protagonist to grasp the toy, over the observer who removed a barrier to a non-desired toy. These findings cannot be explained by a preference for imitators or synchronizers, as neither agent's action was coordinated with that of the protagonist.

Nevertheless, the experiment by Hamlin et al. (2013) was not designed to test for the infants' means-end understanding, and two of its key features suggest that means-end understanding was not required for success in the study. First, 10-month-old infants are at an age that appears, in non-helping contexts, transitional for their means-end understanding. Studies of younger infants would be stronger tests of the possibility that helping contexts unpack means-end actions. Second, infants had been familiarized with the direct action that reflected the protagonist's ultimate goal at test: During familiarization, no barrier was present, and the protagonist directly grasped its desired toy. Thus, infants may have inferred the protagonist's goal from the protagonist's direct action in familiarization, and consequently favored the helper whose actions led to the protagonist having access to that goal in later

events, without any understanding of the means-end actions.

In summary, the rich literature on early social evaluations has yet to reveal whether infants understand acts of helping in accord with the hierarchically structured goals and action plans of the individuals who are helped. Infants may value helpers of agents who plan means-end actions because infants understand means-end action plans in helping contexts, in which different direct actions (the opening of a box, the grasping of a toy) are performed by different agents. Alternatively, infants may value helpers because their actions precede displays of positive affect, because helpers imitate or act in synchrony with the individuals they help, or because infants can infer a protagonist's goal from a direct action. Indeed, infants may be sensitive to multiple cues to the social value of an individual. Because perception of a well-intentioned person is important for learning, infants may develop multiple mechanisms for evaluating other social beings.

Research overview

In the present experiments, we test whether infants' and toddlers' evaluations of helpful agents are based on infants' and toddlers' understanding of others' actions, in contexts in which displays of positive affect, imitation, and synchrony are not available to guide their evaluations. In Experiment 1, we assessed evaluations of helpers who intervene on means-end action sequences at two ages: 15 and 8 months. We focused on these ages, because children at both ages show preferences for helpers in diverse situations (Margoni & Surian, 2018; Woo et al., 2022), but only the toddlers have been shown to enact and reason about others' means-end actions, including the act of opening a box to retrieve a toy that it contains. Would infants and toddlers leverage their differing understanding of means-end actions to arrive at different evaluations of agents who help others? Or might both infants and toddlers understand hierarchically structured means-end actions in helping contexts, in which each agent performs a different part of the action sequence?

Experiment 1 provided evidence that the toddlers evaluated the helper puppets in accord with a protagonist's ultimate goal (the end state of obtaining a toy), whereas the infants instead evaluated the helpers either in accord with the helper's patterns of imitation or with the protagonist's intermediate goal (the means state of opening a box that the protagonist had tried to open, when it had previously contained the desired toy). Two further experiments with infants distinguished between these two possibilities, and a final experiment, conducted with both infants and toddlers, replicated and extended the evidence for the developmental change found in Experiment 1 while addressing

two further alternative accounts of the evidence for an understanding of helping.

EXPERIMENT 1: EVALUATIONS OF HELPERS AND MEANS-END UNDERSTANDING

In Experiment 1, we familiarized 15-month-old toddlers and 8-month-old infants to a puppet show in which a bear (the protagonist) tried and failed to open one of two boxes, each containing a different toy, as the box appeared alternately on each of the two sides of the stage. The bear then opened the box with the aid of two rabbit puppets, each of whom helped the bear to open the box when it appeared on their side of the stage (the helpers). Once the box was open, the bear grasped the toy inside (Figure 1a). From these familiarization events, the toddlers should infer that the bear seeks the toy inside the box that it attempted to open, whereas the infants should instead infer that the bear seeks to open that box (Figure 1b). After this familiarization, we moved the two toys to the opposite boxes in the presence of the helpers (Figure 1c). Thus, the means and end states of the protagonist's original means-end action were now separated.

The final events followed this switch (Figure 1c): While the protagonist observed, one rabbit (the Means-Helper) opened the original box that the protagonist had tried to open in familiarization, and the other rabbit (the End-Helper) opened the box that contained the toy that the protagonist had grasped before the toys' locations were exchanged. All action paused in the final events after the boxes were opened (i.e., participants never saw the protagonist react to a box being opened or grasp the toy it contained), and participants' looking time to each final event was measured using a participant-controlled procedure. These data enabled us to examine participants' attention to the final events by comparing looking times to the final events involving the Means-Helper and the End-Helper. A long tradition of developmental research

on goal understanding has used infant- and toddler-controlled looking times during familiarization and test (the final events) to probe expectations concerning what an actor will do. As reviewed above, research using such infant-controlled looking times has demonstrated that 12-month-old infants can infer the overarching goal of box-opening sequences like that of the present experiment when the actions in the familiarization and test phases are performed by the same agent (Woodward & Sommerville, 2000). Differential expectations of helpful versus unhelpful actions, however, are typically not found in studies using infant-controlled looking times (e.g., Hamlin et al., 2007; Hamlin & Wynn, 2011). We measured infants' and toddlers' attention to the final events because infants and toddlers may have preferred a helper because they found its actions more interesting, rather than more helpful, in the final events. If so, then the infants and toddlers should also have looked longer at that agent during the final events.

Finally, we tested whether participants selectively reached for the Means-Helper or the End-Helper: a standard measure of social preference used in previous studies (e.g., Hamlin et al., 2007; Hamlin & Wynn, 2011). If looking times in the final events provided no evidence that one helper's actions were more interesting, then reaching preferences during the test events are more likely to be indicators of infants' social evaluations of the two helper puppets.

In these experiments, there were multiple objects, agents, and goals that participants could be tracking, potentially introducing multiple demands on participants' attention and working memory. Our methods and procedures were designed to reduce cognitive load in two ways. First, we facilitated participants' tracking of the helpers' actions by restricting the spatial positions of the two helper agents: For each participant, the Means- and End-Helpers remained in constant regions of the stage throughout the familiarization and final events, and each helper only opened boxes on the side where it was stationed. Although the helpers differed

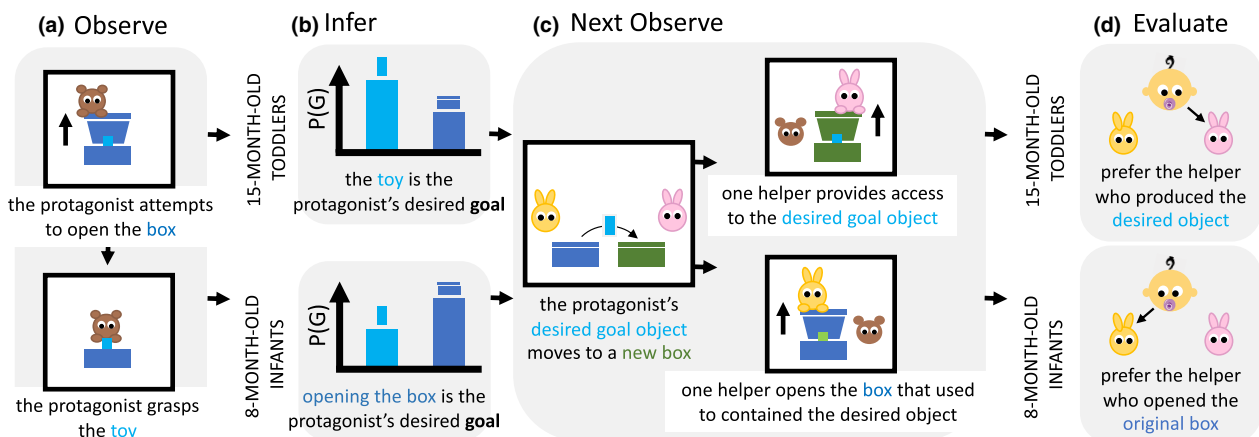


FIGURE 1 A schematic of infants' and toddlers' observations (a, c), inferences (b), and evaluations (d) in Experiment 1.

by color, because each helper always acted on the same side, the sides of each agent allowed the participants to distinguish and track them throughout familiarization, giving them ample time to learn their distinctive colors. Second, the rabbits' physical actions were the same: Both helpers grasped the lid of the box that was closest to them and opened the box. Thus, the participants were not challenged to keep track of two agents who engaged in different physical actions. All the participants needed to track, during the final events, was when each helper performed this action: One did so when the box contained the bear's desired toy and the other did so when the box was the one that did not contain that toy. To form social evaluations, the participants only needed to learn this distinction between the helpers by the end of the last final event.

There are several ways in which our design challenged participants to consider the hierarchical structure of the protagonist's action plan. First, because the means state was separated from the end state of the protagonist's original means-end action (in contrast to the box-opening scenario in Hamlin & Wynn, 2011), our design required participants to determine which state was more relevant to the protagonist's action plan in familiarization, when the two states had been linked. Second, because the protagonist engaged in a means-end action (opening a box, and then grasping the toy inside) in familiarization, our design challenged participants to infer the protagonist's ultimate goal without seeing the protagonist act on the toy directly. Third, because both helpers synchronized their actions with the protagonist during familiarization, and neither helper acted in concert with the protagonist during the final events, synchrony cannot support preferences. Fourth, because the protagonist did not react to the opening of the boxes in the final events, there were no differences in the protagonist's affect that could influence participants' reaching.

Finally, because the box that the protagonist had sought to open in familiarization was not the box that the End-Helper opened in the final events, participants were challenged with considering the consequences of the End-Helper's actions for the protagonist's action plan, rather than the similarity of its actions to those of the protagonist. Indeed, the End-Helper's action was *less* similar to the previous action of the protagonist, who acted on the same box as the Means-Helper and on a different box from the End-Helper. Thus, a preference for the End-Helper could not be explained by a preference for those who imitate others' actions.

The present experiment therefore was designed to investigate how infants' and toddlers' evaluations of helping relate to their ability to recover and reason about others' action plans. If evaluations depend on developing capacities for goal attribution (Figure 1d), then 15-month-old toddlers—who can represent the ultimate goal of others' means-end actions—should prefer the

End-Helper over the Means-Helper, and 8-month-old infants—who are more likely to entertain the possibility that a person's goal in performing a means-end action is the means state—should prefer the Means-Helper over the End-Helper. Further, if helping contexts sufficiently unpack a means-end action into its component parts (as may have been the case with Hamlin et al., 2013), then even 8-month-old infants may prefer the End-Helper over the Means-Helper.

Method

For all our experiments, the methods and analysis plans were preregistered on the Open Science Framework (OSF) at <https://osf.io/k7y5c/>. Stimuli, data, and code are available on the OSF. Data collection occurred from 2019 to 2022.

Participants

Twenty-four full-term 15-month-old toddlers (11 girls; $M_{\text{age}} = 15.05$ months; range = 14;20–15;15), and 24 full-term 8-month-old infants (8 girls; $M_{\text{age}} = 7.86$ months; range = 7;9–8;11) contributed data. An additional nine toddlers were excluded due to fussiness ($n = 5$), inattentiveness ($n = 3$), or failure to choose between puppets in the social preference test ($n = 1$). Likewise, an additional nine infants were excluded due to inattentiveness ($n = 3$), failure to choose between puppets ($n = 3$), caregiver interference ($n = 2$), or fussiness ($n = 1$). In all experiments, experimenters who were naïve about the events seen by participants determined exclusions using preregistered criteria. Lab studies probing infants' evaluations of helping have excluded infants at similar rates (e.g., Hamlin et al., 2013).

In Experiment 1, all participants came from the greater Boston area. About 78% of participants' caregivers completed demographics questionnaires: Approximately 69% of these participants were White, 17% were Asian, 2% were Black, and 11% were multiracial.

For Experiments 1–4, participants were recruited through phone calls or emails to caregivers listed in our lab's database of families who have expressed interest in participating in developmental research (e.g., through responding to mailings or signing up online). Data collection occurred at a university laboratory in Cambridge, Massachusetts. In Experiment 1, caregivers received travel reimbursement and a toy prize for their infants.

Sample size justification

Prior to collecting data, we determined our sample size of 15-month-old toddlers based on power simulations on pilot data ($n = 8$). Similar sample sizes have been used in



prior literature on early social evaluations (e.g., Hamlin et al., 2013).

We collected data on 8-month-old infants after collecting data on 15-month-old toddlers. We determined our sample of infants based on power simulations on the toddler data, and on findings that age does not predict the strength of early evaluations of helpers (Margoni & Surian, 2018). We therefore assumed that 8-month-old infants could form evaluations similar in strength to those of the 15-month-old toddlers, even if that preference were for the Means-Helper.

Displays

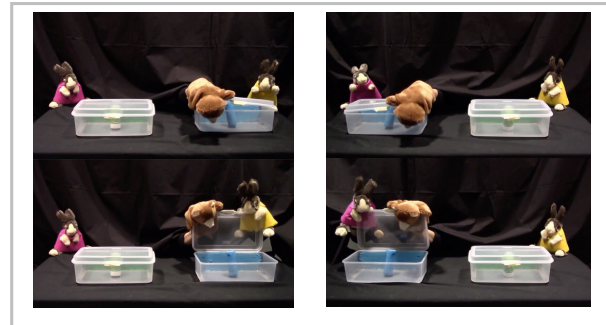
Each participant viewed 6 familiarization events, 1 event in which toys switched positions, and 4 final events, for a total of 11 events. The helping actions were based on Hamlin and Wynn's (2011) box-opening scenario, described above.

All events took place on a stage containing two boxes (one blue, one green), two toys (one blue, one green, matching each box's color) inside the boxes, and two rabbits (one wearing a pink shirt, one wearing a yellow shirt) at the stage's rear corners. In the 6 familiarization events, a protagonist bear puppet consistently tried and failed to open one box alone, as the box appeared alternately on the stage's two sides. Each of the two helpers joined the protagonist on alternating events; together they opened the box, allowing the protagonist to grasp the toy inside.

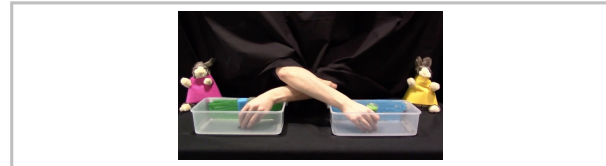
After the familiarization events, participants saw an event in which the toys switched positions, such that each toy now appeared in the box of the contrasting color. We presented participants with videos of two hands that moved the toys, for two reasons. First, if the toys had changed boxes without an event in which they had moved, then infants and toddlers may have focused on the sudden change, rather than on the actions of the agents. Second, presenting infants with hands that moved the objects maximized the likelihood that participants noticed the toys changing positions. This visible movement of the toys was similar to object movements used in studies on early false-belief understanding that require events that make clear that an object has changed locations (e.g., Onishi & Baillargeon, 2005; Rhodes & Brandone, 2014).

In the 4 final events of the show, the protagonist jumped between the two boxes. In alternating events, one helper (the Means-Helper) opened the box that the protagonist had tried to open in familiarization, and the other helper (the End-Helper) opened the other box, which now contained the toy that the protagonist had grasped in familiarization. The action then ceased, and the displays remained present until the infant or toddler looked away from them for 2 s. Events are depicted in Figure 2, and described further in the Supporting Information.

(a) Familiarization



(b) Toys switch



(c) Final events



FIGURE 2 Still images from the video-recorded events in Experiment 1. (a) In the familiarization events, the protagonist repeatedly tried and failed to open the same box, which switched sides on the stage between the events. The puppet on that side of the stage helped the protagonist lift the box's lid, and the protagonist grasped the toy inside. (b) In the toy-switch event, a pair of hands switched the toys' positions in the presence of the helper puppets. (c) At the start of the final events, the protagonist jumped between the boxes. In alternating events, one helper (the End-Helper) opened the box containing the toy that the protagonist had previously grasped, and the other helper (the Means-Helper) opened the box that the protagonist had previously tried to open.

Procedure

Participants sat on their caregiver's lap in a lab environment before an LCD projector screen (102 × 132 cm). (Data collection occurred before the COVID-19 pandemic.) Two speakers located on the screen's sides played all video sounds. Caregivers were instructed to close their eyes, sit quietly, and not influence their children. We required that participants look at the screen while the protagonist struggled to open a box, a helper opened the box, the protagonist grasped the toy inside the box in familiarization, the toys changed positions

in the toy-switch event, and each helper opened a box in the final events. If participants did not see one or more of these critical parts, we repeated the event. (See [Supporting Information](#) for counterbalancing.)

Social preference test

After participants had watched all the events, an experimenter (who was naïve to condition and to the events that participants had seen) presented participants with the two helper puppets in a reaching-based social preference test. Caregivers turned 90° to the left so that they no longer faced the screen, and they closed their eyes. An experimenter who was naïve to the puppets' identities knelt in front of participants and held the Means-Helper and the End-Helper approximately 30 cm apart, initially out of the participants' reach (see [Supporting Information](#) for details). After the participants had looked at both puppets and then back to the experimenter, the puppets were moved within reach, and the experimenter asked, "Who do you like?" The experimenter determined each participant's choice as being the first puppet that the participant touched via a visually guided reach. A second experimenter, also naïve about the puppets' identities, judged which of the two puppets participants looked to and touched first in this test. There was 100% agreement between the two sets of judgments.

Coding and reliability analyses of the events

Looking time data were coded online for all events using Xhab64 (Pinto, 1995) software until participants looked away for 2 consecutive s or until 30 s elapsed. An observer coded participants' looking time (on vs. off the display) through a live video in a separate room, from which this observer could not hear nor see events, and was naïve about counterbalancing. A second, independent observer recorded the final events of a random 25% of participants offline using jHab (Casstevens, 2007) software, following the same criteria while naïve to counterbalancing. The intraclass correlation coefficient between the observers' data was .98 (95% CI [.97, .99]). The coding of looking time followed procedures that are common to studies of familiarization or habituation, followed by detection of novelty or violation of an expectation (for reviews, see Aslin, 2007; Colombo & Mitchell, 2009).

Results

All reported p -values are two-tailed. The 15-month-old toddlers reached more to the End-Helper (21/24 chose the End-Helper, binomial $p < .001$, relative risk = 1.75). In contrast, the 8-month-old infants reached more to the Means-Helper (20/24 chose the Means-Helper, binomial $p < .001$, relative risk = 1.66). Patterns of reaching by the two age groups differed significantly ($\chi^2(1) = 21.37$, $p < .001$, Cohen's $h = 1.49$, odds ratio = 35, 95% CI [6.94, 176.39]; see [Figure 3](#)). At neither age were there significant

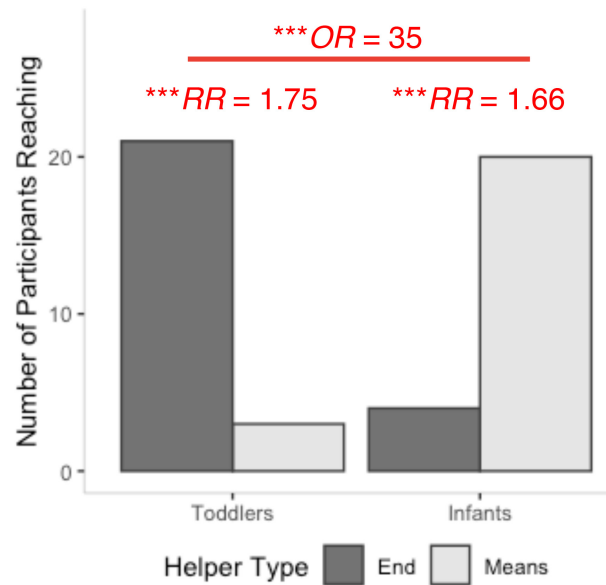


FIGURE 3 The number of 15-month-old toddlers and 8-month-old infants choosing the End-Helper and the Means-Helper in the test of selective reaching in Experiment 1. OR, odds ratio; RR, relative risk (** $p < .001$, two-tailed)

preferences based on the puppets' side of presentation during the show, the order of presentation during the show, the side of presentation during the social preference test, or the color of a puppet (all $ps \geq .541$).

Looking time in the final events, a measure of infants' and toddlers' attention to the helpers' actions did not differ for either age when the Means-Helper and the End-Helper acted ($ps \geq .400$; see [Supporting Information](#) for complete analyses and discussion). These null findings suggest that the infants and toddlers did not expect the helpers to open a specific box, and that their subsequent reaching to the helpers in the preference test was not modulated by differential attention to the helpers' actions in the final events.

Discussion

Fifteen-month-old toddlers selectively reached for the End-Helper, whereas 8-month-old infants selectively reached for the Means-Helper. At neither age can positive affect nor synchrony account for findings, because neither helper was associated with displays of positive affect and because both helpers acted independently of the protagonist in the final events. Moreover, preferences for imitators (Powell & Spelke, 2018b) cannot easily explain the toddlers' preference for the End-Helper, because both helpers engaged in the same action, and because the box that the End-Helper opened in the final events differed from the box that the protagonist had previously tried to open. Finally, infants and toddlers' choices during the social test were not influenced by differing levels of interest in the helping and neutral final events,



because looking times to the two types of events did not differ. By the measure that is standardly used to assess infants' and younger toddlers' social evaluations, therefore, these findings provide evidence that 15-month-old toddlers understood that the protagonist's ultimate goal was to reach a specific toy and evaluated the helpers on the basis of whether they opened the new box containing that toy.

By contrast, 8-month-old infants selectively reached for the Means-Helper, who opened the original box that the protagonist had attempted to open. Infants may have valued the helper who fostered the attainment of the protagonist's intermediate goal (the opening of the original box that had contained the toy) in this means-end action sequence. If so, then these findings suggest that helping does not sufficiently unpack a means-end action into its component parts for 8-month-old infants. It is possible, however, that a preference for imitators accounts for 8-month-old infants' preference for the Means-Helper, because the protagonist's behavior in familiarization was more similar to the Means-Helper's behavior than to the End-Helper's behavior in the final events. In the final events, the Means-Helper opened the same box that the protagonist had tried to open in familiarization, whereas the End-Helper opened the box that the protagonist had ignored in familiarization. Thus, the infants may have preferred the Means-Helper not because it fostered the protagonist's goal but because it imitated the protagonist. The next two experiments aimed to shed light on this possibility and to test infants' evaluations of helping in relation to their understanding of direct actions. In Experiment 2, we tested whether infants inferred the goal of a protagonist who engaged in a direct action: grasping a toy inside a visibly open box. Given the positive findings of that experiment, we probed in Experiment 3 whether 8-month-old infants would prefer an agent who provided a protagonist with access to a desired toy, in a situation in which imitation could not easily support preferences.

EXPERIMENT 2: UNDERSTANDING THE GOAL OF A DIRECT ACTION

In Experiment 2, we familiarized 8-month-old infants to a scenario like that of Experiment 1, except that the box was open, allowing the protagonist to grasp the toy through a direct action. In performing this action, the protagonist never grasped the lid of the box; instead, it directly acted on the toy inside the box and its action ended on contact with that toy. Using looking time measures, here we probed infants' expectations concerning the actions of the protagonist after the toy was removed from the box. Given the findings of Hamlin et al. (2013), we predicted that infants would infer that the protagonist's goal was to get to the toy, rather than to enter the box that the toy had been located in. We preregistered

an analysis testing whether infants looked longer when the protagonist moved to the box than when it moved to the toy.

Method

Participants

Twenty-four 8-month-old infants contributed data to Experiment 2 (9 girls; $M_{\text{age}} = 7.85$ months; range = 7; 9–8;20). An additional two infants began Experiment 2 but were excluded due to inattentiveness ($n = 2$). Participants came from the greater Boston area ($n = 16$), North Carolina ($n = 2$), New Hampshire ($n = 2$), Washington ($n = 1$), Texas ($n = 1$), Puerto Rico ($n = 1$), and Canada ($n = 1$). About 84% of participants' caregivers completed demographics questionnaires: Approximately 54% of these participants were White, 27% were Asian, 4% were Black, and 14% were multiracial.

For Experiments 2–4, caregivers received online gift cards as compensation.

Sample size justification

Prior to collecting data, we determined our sample sizes for Experiment 2 based on power analyses on data from a pilot sample ($n = 8$). Similar sample sizes have been used in prior studies of infants' action understanding (e.g., Sommerville et al., 2005, 2008).

Displays

Each infant viewed 6 familiarization events, 1 event in which a toy was removed from a box, and 6 test events, for a total of 13 events. We included 6 test events here, as in many previous studies of goal attribution (e.g., Sommerville et al., 2005; Woodward, 1998), because our primary interest for Experiment 2 was infants' expectations.

All events took place on a puppet stage. In the familiarization events, the stage contained two open boxes (one blue, one green), each containing a toy (one blue, one green, corresponding to the box's color). In the first 6 events, the familiarization events, a protagonist puppet consistently jumped to grasp the toy inside one of the boxes. After the familiarization events, infants saw an event in which only one box (the one with the protagonist's desired toy) remained on stage, and a hand moved the toy out of the box to approximately the position that the other toy from familiarization would have been in. Then, in 6 test events, the protagonist alternately (i) jumped to and entered the now-empty box and (ii) jumped to and grasped the toy, whereupon all action paused until the infant looked away. Events are depicted in Figure 4, and described in Supporting Information.

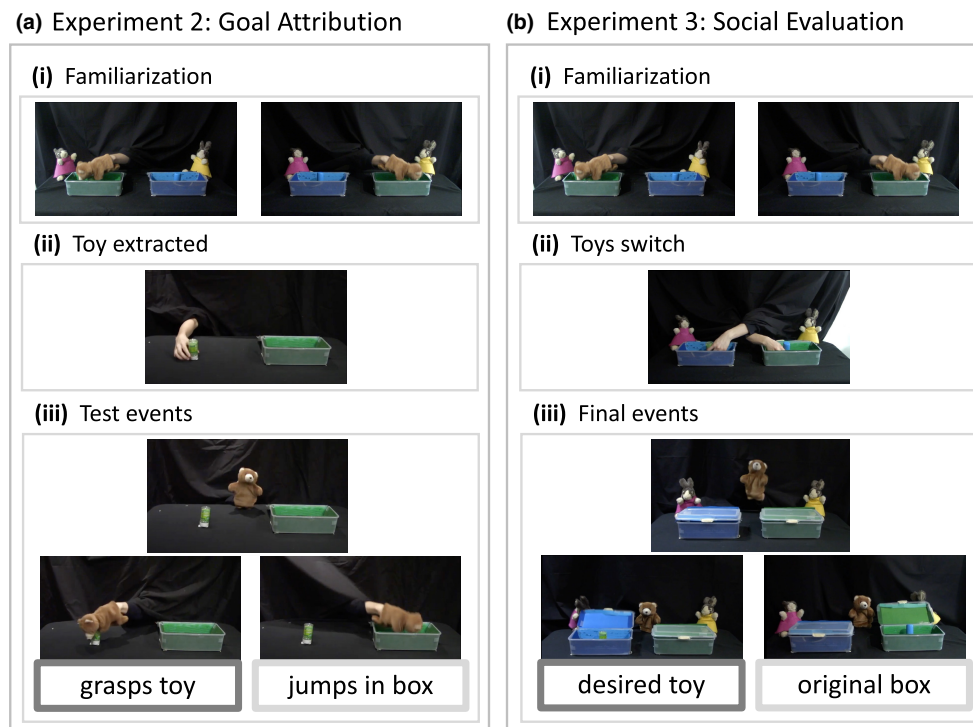


FIGURE 4 Still images from the events in Experiments 2 (a) and 3 (b). (i) In the familiarization events for Experiments 2 and 3, the protagonist repeatedly grasped the toy in the same already-opened box, which switched sides on the stage between the events, while the helper puppets watched without moving. (ii) In Experiment 2's toy-extraction event, a single box appeared onscreen, and a hand removed the toy from the box and placed the toy on the opposite side of the stage. In Experiment 3, the toys switched positions as in Experiment 1. (iii) In Experiment 2, the hand left the stage and the protagonist appeared and jumped between the toy and the now-empty box (top). Then, in alternating events, the protagonist approached and grasped the toy or approached and jumped into the now-empty box (bottom); action froze while infants' looking time was recorded. In Experiment 3, the final events were the same as in Experiment 1.

Procedure

Whereas data collection for Experiment 1 took place in the lab, data collection for Experiments 2–4 took place over Zoom because the COVID-19 pandemic halted in-person research. Caregivers had the option of having their infant seated on their lap or in a highchair. If seated on a caregiver's lap, we asked the caregiver to close their eyes or to look away during the test events. As in Experiment 1, caregivers were instructed to be quiet and to not influence their infants. After the experiment ended, we checked video quality by asking caregivers whether the motion was continuous and smooth in familiarization. None reported issues that would impede infants' understanding of the events. See [Supporting Information](#) for counterbalancing.

In each test event, we required that infants look at the screen when the protagonist grasped a toy or jumped into the box. If infants did not look during this critical time, we repeated the event.

Coding and reliability analyses

Looking time data were coded online using jHAB (Casstevens, 2007) software, following the pausing of actions in test events, until infants looked away for 2 consecutive s or until 30s elapsed. These looking time data served to evaluate participants' expectations of the

protagonist's actions in the test events. A second, independent observer recorded the test events of a random 25% of participants offline using jHAB (Casstevens, 2007) software, following the same criteria while naïve to counterbalancing. The intraclass correlation coefficient between the two observers' data was .96 (95% CI [.94, .98]).

Results

To determine whether the infants looked longer in test events when the protagonist acted on the empty box or the toy, we examined looking times in a mixed-effects model. The dependent variable was looking time. Because a log-normal distribution (log-likelihood = -412.00) fit these data better than did a normal distribution (log-likelihood = -446.87), we log-transformed the data before including it in the model. The fixed effect was event type (Box/Toy), and participant ID and trial pair (1/2/3) were included as random slopes.

The infants looked longer to the test events in which the protagonist moved to and entered the empty box ($M_{\text{box}} = 9.93\text{ s}$) than to the events in which the protagonist moved to and grasped the toy ($M_{\text{toy}} = 7.51\text{ s}$; $\beta = .33$, 95% CI of β [.07, .58], $b = .28$, $t(105) = 2.48$, $p = .014$; see [Figure 5a](#)). Thus, familiarization with the bear directly

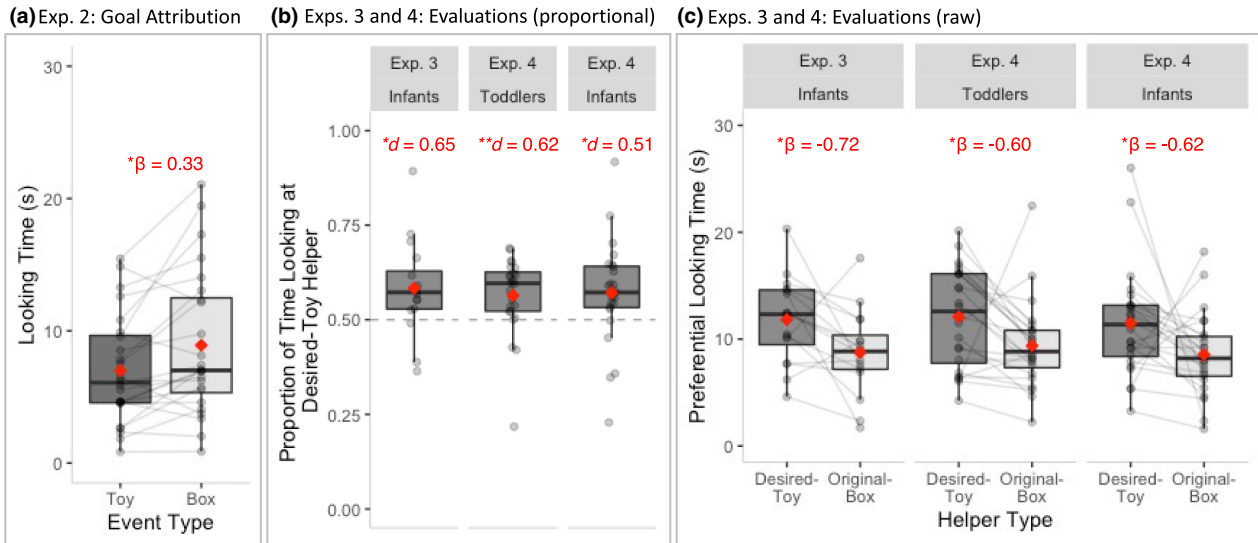


FIGURE 5 Results from Experiments 2–4. Panel a depicts infants' looking time in the final events of Experiment 2, often used as a measure of infants' expectations for behavior. Panels b and c depict the proportion and raw amount of time that the infants and toddlers looked to the Desired-Toy Helper, relative to the Original-Box Helper, in the visual social preference test that followed the events in Experiment 3 and 4, a measure of infants' social evaluations. In all panels, red diamonds indicate means, horizontal lines within boxes indicate medians, boxes indicate interquartile ranges, and pairs of connected dots indicate data from a single infant. Beta coefficients (β) and Cohen's d indicate effect sizes ($*p < .05$, $**p < .001$, two-tailed)

approaching the toy in the box generalized to actions on the toy outside the box, leading to longer looking at actions directed to the box. These findings provide evidence that infants expected the protagonist to continue to act directly on the toy, rather than to enter the box, when the toy and box were separated. Low-level predictors of looking time did not account for these findings (see [Supporting Information](#)).

Discussion

In Experiment 2, infants looked longer when the protagonist jumped into the box that had contained its desired toy than when the protagonist directly grasped the toy. These findings provide evidence that infants represented the protagonist's goal as getting to the toy, not as jumping into the box that the toy had occupied, and they expected the protagonist to approach the toy when it was removed from the box. These findings stand in contrast to studies of younger infants' inferences about the goal of an agent who engages in a means-end action (e.g., Gerson et al., 2015; Sommerville & Woodward, 2005). The critical difference between the present experiment and experiments on means-end action understanding is that the protagonist in the present experiment demonstrated its goal by acting directly on the toy. Because the present experiment did not require the recovery of a hierarchical, means-end action plan, the 8-month-old infants were able to infer the protagonist's goal. Given infants' success in inferring the goal of the protagonist in Experiment 2, the next experiment tested whether

infants prefer an agent who facilitates that goal over an agent who does not, in the absence of confounding social cues.

EXPERIMENT 3: EVALUATIONS OF HELPERS AND UNDERSTANDING DIRECT ACTIONS

In Experiment 3, we tested whether the 8-month-old infants in Experiment 1 preferred the helper who opened the original box because they viewed that helper as aiding the protagonist in reaching its goal (which they took to be the box) or because they viewed that helper as signaling social engagement by imitating the protagonist's action. To distinguish these possibilities in Experiment 2, we familiarized 8-month-old infants to events in which the protagonist engaged in a direct action on the toy without first acting on the box. Next, we moved the toy into a new box as two rabbits observed, and we closed the boxes. One rabbit (the Original-Box Helper) opened the original box that formerly had contained the protagonist's desired toy, whereas the other rabbit (the Desired-Toy Helper) opened the new box that contained the protagonist's desired toy.

Like Experiment 1, Experiment 3 provides a test of the role of goal inference in infants' evaluations of helpers. The events presented to the 8-month-old infants in Experiment 3 were very similar to those presented to the infants in Experiment 1: Only the nature of the protagonist's action on the toy (a means-end action in Experiment 1 and a direct action in Experiment

3) differed. If 8-month-old infants value agents who foster others' goals, and the infants in Experiment 1 preferred the Means-Helper because they inferred that the protagonist's goal was the first object that it acted upon (i.e., the box), then infants should prefer the Desired-Toy Helper in Experiment 3, because the protagonist demonstrates its goal in a direct action in familiarization. Moreover, this finding would conceptually replicate findings that 10-month-old infants prefer an agent who provides access to a protagonist's desired toy over an agent who instead provides access to a different toy (Hamlin et al., 2013), but in a younger age group and with a different action and goal. In contrast, if the infants preferred the Means-Helper in Experiment 1 because its actions imitated the protagonist's actions, then they either should not distinguish between the helpers in Experiment 3, because neither helper performed the same action as the protagonist, or they should prefer the Original-Box Helper, because the Original-Box Helper directed its actions to the same box as the protagonist.

Method

Because data collection occurred during the COVID-19 pandemic, we could not assess the infants' reaching behavior directed at physical puppets (as in Experiment 1). We instead assessed the infants' evaluations through a social preference test focused on preferential looking behavior (as in Geraci et al., 2022; Hamlin et al., 2010; Powell & Spelke, 2018b; Singh, 2020).

Participants

Sixteen 8-month-old infants contributed data to Experiment 3 (8 girls; $M_{\text{age}} = 8.13$ months; range = 7;9–8;25). In contrast to Experiment 1, no infants met the preset exclusion criteria, even though the events in the two experiments were similar in complexity. The lack of attrition may be due to (i) testing occurring online, such that the infants were in their home environments where they felt comfortable, or (ii) events being looped such that infants were more engaged and received greater exposure to the familiarization and final events.

Participants came from the greater Boston area ($n = 14$), Pennsylvania ($n = 1$), and Texas ($n = 1$). About 62% of participants' caregivers completed demographics questionnaires: 70% of these participants were White, 20% were Asian, and 10% were multiracial.

Sample size justification

Prior to collecting data, we determined our sample sizes for Experiment 3 based on power analyses on pilot data ($n = 8$). Additionally, we used the data from the

infants in Experiment 1 for a power analysis, assuming the data reflect infants' capacity to form preferences. Similar sample sizes have been used in research that has measured preferential looking to probe infants' social evaluations (e.g., Hamlin et al., 2010; Powell & Spelke, 2018b).

Displays

Each infant viewed 6 familiarization events, 1 event in which toys switched positions, and 4 final events, for a total of 11 events.

All events took place on a puppet stage containing two boxes and two toys, as in Experiments 1 and 2. The 6 familiarization events were exactly like those of Experiment 2, with a bear protagonist repeatedly jumping to grasp the toy inside a box (see Figure 4; see Supporting Information for additional details). Two rabbits (helpers; one wearing a pink shirt, one wearing a yellow shirt) were present on stage, at the rear corners, for familiarization.

In the toy-switch event, while the helpers were present on stage, a pair of hands took the toys in the boxes, switched them, and closed the boxes, as in Experiment 1.

Each of the 4 final events began with the helpers sitting at the same rear corners of the stage as in familiarization. The boxes were on stage as in familiarization, except that the contents had been switched. In alternating events (as in Experiment 1), one rabbit (the Original-Box Helper) opened the original box that the protagonist had jumped into, even though the toy inside was now different, and the other rabbit opened the box that contained the protagonist's desired toy (the Desired-Toy Helper). These videos were looped 4 times following the opening of the box, and we required that infants see the helper opening the box at least once per trial. This phase was designed to engage infants' attention and to provide evidence of the helpers' behavior so that infants could form evaluations. Because remote testing in the home involves smaller, less compelling displays in an environment with multiple potential distractions, we minimized the duration of the events and maximized the informativeness of each event. These events therefore were not designed to assess infants' expectations about the final events, and relied instead on the findings from past research (e.g., Hamlin et al., 2007; Hamlin & Wynn, 2011) that helpful and unhelpful events are likely to be equally expected (See Supporting Information for counterbalancing details).

Procedure

Caregivers received the same instructions as in Experiment 2 to set up the experiment and optimize data collection.

Social preference test and reliability analyses

As in past research on helping (e.g., Hamlin et al., 2010; Hamlin & Wynn, 2011), the social preference test based on looking time presented infants with the same arrangement of puppets and the same verbal prompt (“Who do you like?”) as did the social preference test based on reaching. After infants had seen all the familiarization and final events, the two rabbits appeared on opposite sides of the screen and moved to an experimenter’s prerecorded voice saying “Hi! Look! Who do you like?” three times, once every 10 s. (Infants and toddlers heard the same speech in Experiment 1, when an experimenter presented real 3D rabbit puppets for participants to reach for.) We coded all looking at the helpers during the social preference test, and we calculated the proportion of this time that infants spent looking at the Desired-Toy Helper.

An experimenter, naïve to condition and to the events that the infants had seen, coded the videos to determine how much time the infants spent looking at each of the two helpers. A second, naïve experimenter coded a randomly selected 25% of infants. The intraclass correlation coefficients between the two observers’ data were .97 (95% CI [.80, .99]) and .99 (95% CI [.91, .99]) for left and right looking, respectively.

Note that the social preference measure differs from the coding of looking in the final events (Experiments 1, 3, and 4) or test events (Experiment 2) of these experiments, in four ways. First, coding for the social preference test occurred after participants had seen all the events, whereas the coding of looking time in the final events occurred *during* the events. Thus, looking time during the social preference test focused on infants’ attention to the individual helper characters who had participated in these events, whereas looking time during the final events focused on infants’ attention to the actions that the characters performed. Second, coding in the social preference test probed whether participants looked to a particular side of the screen (i.e., at an agent), whereas the coding in the final events instead served to examine how long participants attended to any part of the scene during that event. Thus, looking time during the social preference test served as a measure of participants’ differential engagement with each of the characters, whereas looking time during the final events served as a measure of infants’ state of arousal or interest in the events. Third, preferential looking in the social preference test did not stop after participants looked away for 2 s, whereas the coding in the final events did, so as to shed light on how quickly participants lost interest in a given event. Finally, the social preference tests of looking and reaching occurred when the rabbit puppets faced the infants, with accompanying social language (i.e., a voice asking, “Who do you like?”), whereas the looking time measures in final events occurred when the puppets instead attended to and acted on inanimate objects (the boxes),

with no accompanying social language. Thus, the social preference test was more socially directed at the infant than were the final events.

Coding and reliability analyses of the final events

Looking time data for the final events were coded offline for all events using jHab (Casstevens, 2007) software, following the opening of a box in the final events, until participants looked away for 2 consecutive s or until 30 s elapsed. A second, independent observer recorded the final events of a random 25% of participants offline using jHab (Casstevens, 2007) software, following the same criteria while naïve to counterbalancing. The intraclass correlation coefficient between the two observers’ data was .95 (95% CI [.89, .98]). The coding and analysis of looking during the final events were not preregistered or planned; they were undertaken after data collection ended to determine whether differential attention to the final events might be a confounding factor in the social preference test based on looking time.

Results

We ran a one-sample *t*-test to determine whether the proportion of time infants looked at the Desired-Toy Helper over the Original-Box Helper differed from chance (50%) in the social preference test, in which the helpers appeared side by side and infants were asked “Who do you like?” Here, the infants looked longer to the Desired-Toy Helper ($M_{\text{desired-toy}} = 12.32$ s) than to the Original-Box Helper ($M_{\text{original-box}} = 8.73$ s; $M_{\text{desired-toy}}\% = 58.32\%$, $SD = 12.7\%$, 95% CI [0.51, 0.65], one-sample $t(15) = 2.62$, $d = .65$, $p = .019$; see Figure 5b,c). To determine whether the raw amount of time that infants looked at the actors differed in the social preference test, we ran a mixed-effects model (see Supporting Information); the output converged with our analysis on infants’ proportionate looking. There were no significant preferences based on the chosen puppet’s side of presentation, order of presentation, or color (all $ps \geq .177$). See Supporting Information for exploratory analysis comparing Experiments 1 and 3’s patterns of findings in 8-month-old infants.

As in Experiment 1, looking times to the final events involving actions by each of the helpers did not differ ($p = .484$; see Supporting Information for complete analyses). These null findings suggest that infants’ longer looking at the helpful puppet in the social preference test was not produced by differential interest in the puppets.

Discussion

In Experiment 3, infants preferred the helper who directed the protagonist to the new box containing its

desired toy over one who directed the protagonist to the original box that formerly contained its desired toy. This preference was exactly opposite to that of the 8-month-old infants in Experiment 1 (see [Supporting Information](#)), even though the final events depicted the same actions. Eight-month-old infants were better able to infer the protagonist's goal in Experiment 3 than in Experiment 1, because the protagonist acted on its goal in a direct action in Experiment 3, but through a means-end action in Experiment 1.

These findings provide evidence that 8-month-old infants form evaluations based on their analysis of an agent's action plans, in a situation in which goal attribution is not correlated with imitative or synchronized actions. These findings conceptually replicate past findings in 10-month-old infants by Hamlin et al. (2013). Moreover, the findings suggest that 8-month-old infants in Experiment 1 preferred the Means-Helper at least in part because they inferred the protagonist's goal to be the box, not the toy that the box contained. When means-end understanding was not required to attribute the goal of grasping the toy to the protagonist, infants preferred the helper whose actions facilitated that goal in the final events.

In Experiment 4, we tested a further alternative explanation for the findings of Experiments 1 and 3: that toddlers and infants saw the End- and Desired-Toy Helpers, respectively, not as facilitating the protagonist's goal, but either as conforming to the protagonist's goal or as seeking the toy in a box for themselves, because the protagonist's repeated actions towards this toy suggested that it was more valuable. Even though these helpers did not imitate the protagonist's physical actions (instead, they opened a different box than the protagonist had sought to open in familiarization in Experiment 1, and they engaged in box-opening actions where the protagonist had not in Experiment 3), they could be seen as more abstractly imitating or adopting the protagonist's goal, by engaging in a new action that would allow them to pursue the protagonist's original goal.

EXPERIMENT 4: EVALUATIONS OF HELPERS WHO ACT FOR OTHERS' BENEFIT

A final experiment addressed this alternative explanation for toddlers' and infants' preferences for helpers. We presented toddlers and infants with displays like those of Experiments 1 and 3, respectively, except that after opening boxes in the final events, the helpers left the stage. Thus, the helpers' actions were inconsistent both with a desire to signal their similarity to the protagonist by imitating its action, and with a desire to procure the object for themselves because the protagonist's action suggested that it was more valuable. If toddlers' and infants' evaluations were based on whether an agent

fostered the protagonist's goal in Experiments 1 and 3, respectively, then they should continue to prefer the End- and Desired-Toy Helpers in Experiment 4. In contrast, if their social evaluations were based either on imitation at the level of goals, or on rational inferences concerning the greater value of the object that the protagonist had sought, then toddlers and infants should not form preferences between the two helpers when both left the scene without touching either toy.

Method

As in Experiment 3, the COVID-19 pandemic prevented in-person testing. Thus, we instead assessed participants' evaluations through their preferential looking in a social preference test, as in Experiment 2. Although most studies of older infants' and toddlers' evaluations of agents have depended on reaching behavior, there are two reasons that preferential looking would be an effective approach to study toddlers' evaluations. First, studies have successfully used preferential looking to probe older infants' and toddlers' evaluations (Colomer et al., 2020; Geraci et al., 2022). In one study that measured both looking and reaching preferences in the same participants, the findings from these measures converged (Geraci et al., 2022). Second, in studies of infants' and toddlers' social evaluations, older infants' and toddlers' preferential reaching behaviors often align with younger infants' preferential looking behaviors (e.g., Hamlin & Wynn, 2011; Hamlin et al., 2007, 2010; Powell & Spelke, 2018b).

Participants

Both 15-month-old toddlers and 8-month-old infants contributed data to this experiment. There were 24 full-term 15-month-old toddlers (girls; $M_{\text{age}} = 15.01$ months; range = 14;13–15;18), and 24 full-term 8-month-old infants (11 girls; $M_{\text{age}} = 8.12$ months; range = 7;11–8;21). An additional 3 toddlers were excluded due to caregiver interference ($n = 2$) or equipment failure ($n = 1$). An additional 2 infants were excluded due to inattentiveness.

Participants came from the greater Boston area ($n = 37$), North Carolina ($n = 3$), Alabama ($n = 1$), California ($n = 1$), Connecticut ($n = 1$), Nevada ($n = 1$), New York ($n = 1$), Ohio ($n = 1$), Texas ($n = 1$), and Virginia ($n = 1$). About 49% of participants' caregivers completed demographics questionnaires: Approximately 69% of these participants were White, 15% were multiracial, 7% were Black, and 7% were Hispanic or Latino.

Sample size justification

Prior to collecting data, we determined our sample sizes for Experiment 4 based on power analyses on Experiment 3's data and Experiment 1's toddler data.

Displays

Displays for 8-month-old infants were like those of Experiment 3, except that after boxes were opened in the final events, the rabbit who opened the box proceeded to leave the stage. Likewise, displays for 15-month-old toddlers were like those of Experiment 1, except that after boxes were opened in familiarization and in the final events, the rabbit who opened the box proceeded to leave the stage. The final events were therefore the same for 8-month-old infants and 15-month-old toddlers in Experiment 4. For infants, we use the same labels for helpers as in Experiment 3, based on whether a helper opened the original box or provided access to the desired toy in final events. Videos were looped as in Experiment 3 to better engage participants.

Procedure

Caregivers received the same instructions as in Experiment 2 to set up the experiment and optimize data collection.

Social preference test and reliability analyses

The social preference test was the same as in Experiment 3 and was coded in the same way. We again calculated the proportion of time that participants spent looking at the End-Helper versus the Desired-Toy Helper. For toddlers, the intraclass correlation coefficients between two naïve observers' data were .89 (95% CI [.51, .98]) and .95 (95% CI [.75, .99]) for left and right looking, respectively. For infants, the intraclass correlation coefficients between two naïve observers' data were .90 (95% CI [.52, .98]) and .96 (95% CI [.79, .99]) for left and right looking, respectively.

Coding and reliability analyses of the final events

Looking time data were coded online using jHab (Casstevens, 2007) software, following the opening of a box in the final events, until infants and toddlers looked away for 2 s or until 30 s elapsed, whichever came first. These data enabled us to examine participants' attention to the actions during the final events.

A second, independent observer recorded the final events of a random 25% of participants offline using jHab (Casstevens, 2007) software, following the same criteria while naïve to counterbalancing. The intraclass correlation coefficient between the two observers' data was .98 (95% CI [.95, .99]) for the 8-month-old infants and .99 (95% CI [.98, .99]) for the 15-month-old toddlers.

Results

In Experiment 4, the toddlers looked longer to the End-Helper ($M_{\text{end}} = 12.08\text{ s}$) than to the Means-Helper

($M_{\text{means}} = 9.37\text{ s}$; $M_{\text{end}}\% = 56.4\%$, $SD = 10.3\%$, 95% CI [0.52, 0.60], one-sample $t(23) = 3.05$, $p = .005$), as in Experiment 1. As in Experiment 3, the output of a mixed-effects model (see [Supporting Information](#)) on the raw looking times in the social preference test converged with the analysis on toddlers' proportionate looking. There were no significant preferences based on the chosen puppet's side of presentation, order of presentation, or color (all $ps \geq .157$).

Similarly, the infants looked longer to the Desired-Toy Helper ($M_{\text{desired-toy}} = 11.49\text{ s}$) than to the Original-Box Helper ($M_{\text{original-box}} = 8.55\text{ s}$; $M_{\text{desired-toy}}\% = 57.11\%$, $SD = 13.9\%$, 95% CI [0.51, 0.62], one-sample $t(23) = 2.50$, $p = .019$), as in Experiment 3. The output of a mixed-effects model (see [Supporting Information](#)) on the raw looking times in the preference test converged with our analysis on infants' proportionate looking. There were no significant preferences based on the chosen puppet's side of presentation, order of presentation, or color (all $ps \geq .271$). See [Supporting Information](#) for exploratory analysis comparing Experiments 1, 3, and 4's patterns of findings.

As in Experiments 1 and 3, looking time in the final events did not differ when the helpers acted ($ps \geq .151$; see [Supporting Information](#) for complete analyses). Again, these null findings suggest that the infants' and toddlers' preferential looking to the agent who helped the protagonist achieve its goal was not affected by differences in their interest in the two characters.

Discussion

In Experiment 4, both toddlers and infants preferentially looked to the helper who directed the protagonist to the new box containing its desired toy over one who directed the protagonist to the original box that formerly contained its desired toy, as in Experiments 1 and 3, respectively. Experiment 4's findings further show that that evaluations in Experiments 1 and 3 were not based on a preference for others who manifest their sociality by engaging in acts of imitation at the level of goals, or by learning from others' actions about the relative value of the objects on which they act. Neither of these motives was consistent with the helpers' behavior, because the helpers left the scene after opening a box without acting on the object that it contained. Thus, the findings of Experiment 4 (i) replicate those with the toddlers and infants in Experiments 1 and 3, respectively, (ii) provide evidence that social preferences generalize from reaching to looking methods not only in infants but also in toddlers, and (iii) provide evidence that infants robustly value social agents who help other agents to achieve their goals.

GENERAL DISCUSSION

In four experiments, we tested whether infants' and toddlers' evaluations of helpers are modulated by their

understanding of the goals of the protagonist who needs help. In Experiment 1, when a protagonist had demonstrated its goal of getting to a desired toy in a means-end action sequence (opening a box to grasp the toy inside) and then the toy was placed in a new box, 15-month-old toddlers preferred the helper who opened the new box containing the desired toy. By contrast, 8-month-old infants preferred the helper who opened the same box that the protagonist had previously sought to open, even though the desired toy was not there. The infants and toddlers both evaluated the helpers by analyzing the protagonist's goal, but they arrived at opposite evaluations, consistent with their differing understanding of the protagonist's action.

These different patterns of findings accord with the findings of a rich literature on toddlers' and infants' abilities to infer the ultimate goal of a means-end action, with toddlers privileging the end state and younger infants privileging the means state (Gerson et al., 2015; Sommerville & Woodward, 2005; Woodward & Sommerville, 2000; see also, Gergely et al., 2002; Henderson et al., 2013; Henderson & Woodward, 2011; Rakison & Woodward, 2008; Sommerville et al., 2008). Consistent with this interpretation, the 8-month-old infants in Experiments 2 and 3 represented the protagonist's goal as obtaining a toy when the protagonist directly acted on the toy inside an open box. When the toy was moved to a new box in Experiment 3, the infants preferred the helper who opened the new box that now contained the desired toy. In Experiment 4, the toddlers again preferred the helper who opened the new box that now contained the desired toy, and the infants again preferred the helper who opened the new box, even though the helpers left the stage after opening a box, making it clear that they had no interest either in aligning their own goals with those of the protagonist or in obtaining the box's contents for themselves, and instead acted so as to allow the protagonist to achieve its goal.

When children were able to attribute the goal of obtaining the toy to the protagonist in our experiments, both toddlers and infants valued the helper whose action enabled the protagonist to achieve its goal. These findings show that young children can evaluate the socially directed actions of other agents in situations that lack other social cues to a helper's intentions, such as imitation or synchrony. Infants and toddlers do so by drawing on their capacities for generating and recovering the protagonist's action plan, and by determining whether a helper agent's action created conditions under which the protagonist could carry out its plan.

These findings are striking, because the infants and toddlers in Experiments 1 and 3 never saw the protagonist respond to the opening of the boxes in the critical events of these studies, when the protagonist's desired toy was dissociated from the original box that had contained it. To prefer the helper who opened the box that newly contained the desired toy, the toddlers in Experiments 1

and 4 evidently reasoned that the protagonist's goal in familiarization had been to grasp that toy, even though the infants never saw the protagonist perform that action. In contrast, the infants in Experiment 1 either reasoned that the protagonist's goal in familiarization had been to open the box that it had tried to open before, or they responded to the helper's imitative action. The findings of Experiments 2–4 support the first possibility: When the protagonist demonstrated its goal of obtaining a toy directly, infants attributed that goal to the protagonist and preferred the helper who opened the box containing the desired toy, even though the helper's action did not mirror that of the protagonist and the protagonist did not enter the box that the helper had opened. These findings suggest a reanalysis of past studies of 5- and 9-month-old infants' evaluations of helpers and hinderers using a show depicting a protagonist trying to open a box containing a toy (Hamlin & Wynn, 2011): Young infants may have viewed the protagonist in such a study as wanting the box to be open, rather than as wanting the toy that the box contained.

These findings shed light on the bases of infants' and toddlers' evaluations of helpers. Although past research has found that infants prefer agents who help others over agents who hinder others (Hamlin et al., 2007, 2010; Hamlin & Wynn, 2011; Margoni & Surian, 2018), there has been debate as to whether younger infants' preferences for helpers reflect an understanding of acts of helping as social actions that facilitate other agents' goals (Tan & Hamlin, 2022), or an understanding of communicative signals of a helper's desire for social engagement, observed when a social agent expresses positive affect, imitates others, or synchronizes its behavior with others' actions (Powell & Spelke, 2018a, 2018b; Scarf et al., 2012). Although multiple factors can influence infants' evaluations, the present studies show that when infants and toddlers first view a protagonist who engages in means-end or direct actions on objects, their social evaluations depend in part on their understanding of the plans guiding the agent's actions. Because early understanding of means-end action plans changes from 8 to 15 months, participants at these two ages disagreed over which of two actions was more helpful in Experiment 1. When an agent acted directly on a desired goal object in Experiments 3 and 4, infants also preferred the agent who provided access to that object. At both ages, moreover, the participants in Experiment 4 showed this preference even when the agents left the scene immediately after completing their act of helping, with no further action on the objects and no interaction with the protagonist.

While the present experiments address one question concerning how infants and toddlers form evaluations of agents who help other agents, key questions remain. First, why do infants and toddlers reach for and look at agents who have helped other agents? In some contexts, looking and reaching are motivated by a desire for social engagement. Infants and toddlers may choose to look at

and reach for agents who are more helpful because they view them as better social partners for others, including potentially for themselves. In other contexts, looking and reaching are motivated by curiosity, and they foster children's learning about objects and actions (Kidd et al., 2012; Stahl & Feigenson, 2015). Young children may see helpful agents as more competent individuals, who attribute goals to others in line with their own goal understanding and who facilitate others' goals. Infants and toddlers may reach for and look at agents who are more helpful because such agents engage in interesting actions that create new opportunities for other agents to act. We look forward to future research that probes the ways in which early evaluations support young children's exploration, social learning, and affiliative choices.

Second, how do infants and toddlers reason about the action plans of helpers? The present experiments focus on early reasoning about the protagonist's action plan, and therefore were not designed to address this question. Infants and toddlers could have formed preferences in the present experiments without representing the helper's action plan, by focusing on the outcomes of each helper's action for the protagonist and favoring the helper that produced a positive outcome. A growing body of research, however, invites the hypothesis that infants and toddlers understand acts of helping as guided by second-order social goals, such that the protagonist's goal is embedded within the goal of a helper (see Hamlin et al., 2013; Powell, 2022; Ullman et al., 2009; Woo et al., 2017; Woo & Spelke, *in press*). If infants and toddlers in the present experiments indeed represented the helpers' actions as guided by hierarchically structured social goals, then these findings stand in contrast to all the findings with which this paper began, showing that infants below 12 months of age struggle to reason about hierarchically organized nonsocial goals when viewing a solitary agent who engages in a means-end action (e.g., Gerson et al., 2015; Sommerville & Woodward, 2005; Woodward & Sommerville, 2000). Future research could build on the present findings, together with past research, to investigate directly whether infants represent hierarchically organized goals more effectively when the goals of an agent are social (when infants observe acts of helping) rather than nonsocial (when infants observe means-end actions by a single agent) or personal (when infants act on their own).

Third, what is the role of free choice in evaluations of helping? To facilitate tracking of the helpers' actions, helper puppets appeared and acted in constant positions in Experiments 1, 3, and 4: On each trial, each helper chose *whether* to act but not *where* to act. Both infants and toddlers exhibited consistent preferences between the helper puppets under these conditions, providing evidence that they were sensitive to the puppets' choices of when to open or not open a box. Nevertheless, young children may form stronger preferences when helpers vary their locations or actions, providing stronger evidence for their action choices (see Woo et al., 2021).

When adults see an agent's action as helpful to another agent, we draw, in part, on our abilities to infer what the latter agent is trying to do. The present experiments provide evidence that infants and toddlers share this predisposition: Their evaluations of helpers are based on their understanding of a protagonist's action plans. Specifically, infants and toddlers preferred agents who helped a protagonist achieve the goal that they had attributed to the protagonist. These findings build on past evidence for early social evaluations, and they shed light on one process by which infants and toddlers form evaluations of helpers: by recovering a protagonist's action plan, inferring its goal, and determining what actions by other agents will or will not help the protagonist to achieve its goal.

ACKNOWLEDGMENTS

We thank the families who participated in these experiments, Tomer Ullman and the Cambridge Writing Group for feedback on an earlier draft of this paper, the Harvard Laboratory of Developmental Studies for opportunities for discussion, and Bill Pepe, Melyssa Almeida, Lauren Salmans, Will Adams, Kexin Que, and Yuman Li for help with data coding.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data and analytic code necessary to reproduce the findings presented here are publicly accessible on the OSF, as well are the materials necessary to attempt to replicate these findings: <https://osf.io/k7y5cl>. All the analyses were preregistered, except those labeled as exploratory, and the preregistrations are available here too.

ORCID

Brandon M. Woo  <https://orcid.org/0000-0002-8639-2919>

REFERENCES

- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10(1), 48–53. <https://doi.org/10.1111/j.1467-7687.2007.00563.x>
- Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1, 1–10. <https://doi.org/10.1038/s41562-017-0064>
- Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, 113, 329–349. <https://doi.org/10.1016/j.cognition.2009.07.005>
- Benton, D. T., & Lapan, C. (2022). Moral masters or moral apprentices? A connectionist account of sociomoral evaluation in preverbal infants. *Cognitive Development*, 62, 101164. <https://doi.org/10.1016/j.cogdev.2022.101164>
- Casstevens, R. M. (2007). *jHab: Java habituation software (version 1.0.2) [computer software]*. <https://voices.uchicago.edu/woodwardlab/resources/>
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893–910. <https://doi.org/10.1037/0022-3514.76.6.893>

- Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, 92(2), 225–234. <https://doi.org/10.1016/j.nlm.2008.06.002>
- Colomer, M., Bas, J., & Sebastian-Galles, N. (2020). Efficiency as a principle for social preferences in infancy. *Journal of Experimental Child Psychology*, 194, 104823. <https://doi.org/10.1016/j.jecp.2020.104823>
- Diamond, A. (1991). Guidelines for the study of brain–behavior relationships during development. In H. Eisenberg (Ed.), *Frontal lobe function and dysfunction* (pp. 339–378). New York University Press.
- Geraci, A., Simion, F., & Surian, L. (2022). Infants' intention-based evaluations of distributive actions. *Journal of Experimental Child Psychology*, 220, 105429. <https://doi.org/10.1016/j.jecp.2022.105429>
- Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, 415, 755. <https://doi.org/10.1038/415755a>
- Gerson, S. A., Mahajan, N., Sommerville, J. A., Matz, L., & Woodward, A. L. (2015). Shifting goals: Effects of active and observational experience on infants' understanding of higher order goals. *Frontiers in Psychology*, 6, 310. <https://doi.org/10.3389/fpsyg.2015.00310>
- Hamlin, J. K. (2015). The case for social evaluation in preverbal infants: Gazing toward one's goal drives infants' preferences for helpers over hinderers in the hill paradigm. *Frontiers in Psychology*, 5, 1563. <https://doi.org/10.3389/fpsyg.2014.01563>
- Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, 16, 209–226. <https://doi.org/10.1111/desc.12017>
- Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, 26, 30–39. <https://doi.org/10.1016/j.cogdev.2010.09.001>
- Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450, 557–559. <https://doi.org/10.1038/nature06288>
- Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, 13, 923–929. <https://doi.org/10.1111/j.1467-7687.2010.00951.x>
- Henderson, A. M., Wang, Y., Matz, L. E., & Woodward, A. L. (2013). Active experience shapes 10-month-old infants' understanding of collaborative goals. *Infancy*, 18, 10–39. <https://doi.org/10.1111/j.1532-7078.2012.00126.x>
- Henderson, A. M., & Woodward, A. L. (2011). “Let's work together”: What do infants understand about collaborative goals? *Cognition*, 121, 12–21. <https://doi.org/10.1016/j.cognition.2011.05.008>
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27, 949–960. <https://doi.org/10.1521/soco.2009.27.6.949>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One*, 7, e36399. <https://doi.org/10.1371/journal.pone.0036399>
- Margoni, F., & Surian, L. (2018). Infants' evaluation of prosocial and antisocial agents: A meta-analysis. *Developmental Psychology*, 54, 1445–1455. <https://doi.org/10.1037/dev0000538>
- Meltzoff, A. N. (1990). Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling, and self practice in infancy. In D. Cicchetti & M. Beeghly (Eds.), *The self in transition: Infancy to childhood* (pp. 139–164). University of Chicago Press.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. <https://doi.org/10.1126/science.1107621>
- Piaget, J. (1952). *The origins of intelligence in the child*. Routledge & Kegan Paul.
- Pinto, J. (1995). *Xhab64 [computer software]*. Stanford University.
- Powell, L. J. (2022). Adopted utility calculus: Origins of a concept of social affiliation. *Perspectives on Psychological Science*, 17, 1215–1233. <https://doi.org/10.1177/17456916211048487>
- Powell, L. J., & Spelke, E. S. (2018a). Human infants' understanding of social imitation: Inferences of affiliation from third party observations. *Cognition*, 170, 31–48. <https://doi.org/10.1016/j.cognition.2017.09.007>
- Powell, L. J., & Spelke, E. S. (2018b). Third-party preferences for imitators in preverbal infants. *Open Mind*, 2, 61–71. https://doi.org/10.1162/opmi_a_00018
- Rakison, D. H., & Woodward, A. L. (2008). New perspectives on the effects of action on perceptual and cognitive development. *Developmental Psychology*, 44, 1209–1213. <https://doi.org/10.1037/a0012999>
- Rhodes, M., & Brandone, A. C. (2014). Three-year-olds' theories of mind in actions and words. *Frontiers in Psychology*, 5, 263. <https://doi.org/10.3389/fpsyg.2014.00263>
- 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. <https://doi.org/10.1371/journal.pone.0042698>
- Schlingloff, L., Csibra, G., & Tatone, D. (2020). Do 15-month-old infants prefer helpers? A replication of Hamlin et al. (2007). *Royal Society Open Science*, 7, 191795. <https://doi.org/10.1098/rsos.191795>
- Singh, L. (2020). Bilingual infants are more sensitive to morally relevant social behavior than monolingual infants. *Journal of Cognition and Development*, 21, 631–650. <https://doi.org/10.1080/15248372.2020.1807987>
- Sommerville, J. A., Hildebrand, E. A., & Crane, C. C. (2008). Experience matters: The impact of doing versus watching on infants' subsequent perception of tool-use events. *Developmental Psychology*, 44, 1249–1256. <https://doi.org/10.1037/a0012296>
- Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition*, 95, 1–30. <https://doi.org/10.1016/j.cognition.2003.12.004>
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96, B1–B11. <https://doi.org/10.1016/j.cognition.2004.07.004>
- Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, 348, 91–94. <https://doi.org/10.1126/science.aaa3799>
- Steckler, C. M., Woo, B. M., & Hamlin, J. K. (2017). The limits of early social evaluation: 9-month-olds fail to generate social evaluations of individuals who behave inconsistently. *Cognition*, 167, 255–265. <https://doi.org/10.1016/j.cognition.2017.03.018>
- Tan, E., & Hamlin, J. K. (2022). Mechanisms of social evaluation in infancy: A preregistered exploration of infants' eye-movement and pupillary responses to prosocial and antisocial events. *Infancy*, 27, 255–276. <https://doi.org/10.1111/infa.12447>
- 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, 675–691. <https://doi.org/10.1017/S0140525X05000129>
- Tunçgenç, B., Cohen, E., & Fawcett, C. (2015). Rock with me: The role of movement synchrony in infants' social and nonsocial choices. *Child Development*, 86, 976–984. <https://doi.org/10.1111/cdev.12354>
- Ullman, T. D., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D., & Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. *Advances in Neural Information Processing Systems*, 22, 1874–1882.
- Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science*, 18(3), 184–188. <https://doi.org/10.1111/j.1467-8721.2009.01633.x>

- Woo, B. M., Liu, S., & Spelke, E. (2021). *Open-minded, not naïve: Three-month-old infants encode objects as the goals of other people's reaches*. Proceedings of the 43rd annual meeting of the Cognitive Science Society.
- Woo, B. M., & Spelke, E. S. (in press). Toddlers' social evaluations of agents who act on false beliefs. *Developmental Science*. <https://doi.org/10.1111/desc.13314>
- Woo, B. M., Steckler, C. M., Le, D. T., & Hamlin, J. K. (2017). Social evaluation of intentional, truly accidental, and negligently accidental helpers and harmers by 10-month-old infants. *Cognition*, *168*, 154–163. <https://doi.org/10.1016/j.cognition.2017.06.029>
- Woo, B. M., Tan, E., & Hamlin, J. K. (2022). Human morality is based on an early-emerging moral core. *Annual Review of Developmental Psychology*, *4*, 41–61. <https://doi.org/10.1146/annurev-devpsych-121020-023312>
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34. [https://doi.org/10.1016/S0010-0277\(98\)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)
- Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, *18*, 53–57. <https://doi.org/10.1111/j.1467-8721.2009.01605.x>
- Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, *11*, 73–77. <https://doi.org/10.1111/1467-9280.00218>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Woo, B. M., & Spelke, E. S. (2023). Infants and toddlers leverage their understanding of action goals to evaluate agents who help others. *Child Development*, *00*, 1–18. <https://doi.org/10.1111/cdev.13895>