



Full length article

Let's get it together: Infants generate visual predictions based on collaborative goals



Sheila Krogh-Jespersen^{a,*}, Annette M.E. Henderson^b, Amanda L. Woodward^c

^a Department of Medical Social Science, Feinberg School of Medicine, Northwestern University, Chicago, IL 60614, United States

^b School of Psychology, The University of Auckland, New Zealand

^c Department of Psychology, University of Chicago, United States

ARTICLE INFO

Keywords:

Infancy
Collaboration
Goal understanding
Eye-tracking
Social cognition

ABSTRACT

Infants engage in social interactions that include multiple partners from very early in development. A growing body of research shows that infants visually predict the outcomes of an individual's intentional actions, such as a person reaching towards an object (e.g., Krogh-Jespersen & Woodward, 2014), and even show sophistication in their predictions regarding failed actions (e.g., Brandone, Horwitz, Aslin, & Wellman, 2014). Less is known about infants' understanding of actions involving more than one individual (e.g., collaborative actions), which require representing each partners' actions in light of the shared goal. Using eye-tracking, Study 1 examined whether 14-month-old infants visually predict the actions of an individual based on her previously shared goal. Infants viewed videos of two women engaged in either a collaborative or noncollaborative interaction. At test, only one woman was present and infants' visual predictions regarding her future actions were measured. Fourteen-month-olds anticipated an individual's future actions based on her past collaborative behavior. Study 2 revealed that 11-month-old infants only visually predict higher-order shared goals after engaging in a collaborative intervention. Together, our results indicate that by the second year after birth, infants perceive others' collaborative actions as structured by shared goals and that active engagement in collaboration strengthens this understanding in young infants.

1. Let's get it together: Infants generate rapid visual predictions based on higher-order collaborative goals

Infants anticipate the outcomes of individual actions, such as reaching for an object, early in development (Cannon & Woodward, 2012; Gredebäck, Stasiewicz, Falck-Ytter, Rosander, & von Hofsten, 2009; Kanakogi & Itakura, 2010; Krogh-Jespersen & Woodward, 2014). Recent work has also shown that infants' experience performing actions themselves relates to their perception of the same actions being performed by others (e.g., Ambrosini, Reddy, de Looper, Costantini, & Lopez, 2013; Brune & Woodward, 2007; Cannon, Woodward, Gredebäck, Von Hofsten, & Turek, 2012; Kanakogi & Itakura, 2011; Loucks & Sommerville, 2012; Sommerville & Woodward, 2005). While this mapping of action and perception sheds light on infants' understanding of the intentions and goals of individuals, it leaves open the question of whether infants infer intentionality in more complicated social interactions, specifically interactions that involve multiple individuals. Many actions in infants' social worlds have goal structures that are less concrete and involve multiple individuals, such as collaborative interactions when actors may not perform an immediate action on their goal. Collaborative activities are activities in which the actions of multiple individuals are complementary and critical to attaining a shared

* Corresponding author.

E-mail address: sheilakj@northwestern.edu (S. Krogh-Jespersen).

<https://doi.org/10.1016/j.infbeh.2020.101446>

Received 6 September 2019; Received in revised form 23 March 2020; Accepted 24 March 2020

Available online 20 April 2020

0163-6383/ © 2020 Elsevier Inc. All rights reserved.

goal and pervade everyday human life (see also [Bratman, 1992](#)). Understanding higher-order collaborative goal structures is a critical marker in development that is demonstrated by infants early in the second year after their birth ([Henderson & Woodward, 2011](#)). Whether infants engage this type of goal understanding when predicting future actions is an open question addressed by the current research.

Initial studies utilizing visual habituation paradigms, which measure infants' retrospective intentional understanding, demonstrate that infants structure meaningful actions in terms of the relation between individuals and their goals; that is, they expect a person to continue to perform actions, such as reaching and grasping, to reach the same goal (e.g., [Biro & Leslie, 2007](#); [Woodward, 1998, 1999](#)). Recent developments in the field of eye-tracking have built upon and extended the pioneer work on infants' understanding of goal-directed actions by demonstrating that infants show prospective understanding by visually anticipating the outcomes of simple action events, such as reaching for or displacing an object ([Brandone, Horwitz, Aslin, & Wellman, 2014](#); [Gredebäck et al., 2009](#); [Henrichs, Elsner, Elsner, Wilkinson, & Gredebäck, 2014](#)). These studies have demonstrated that infants can anticipate simple action events in the first year of their lives. However, several of these studies utilized a measure that confounds goal understanding with lower level processes, such as the movement/trajectory of the agent's action. This confound in study design leaves open the possibility that infants rely on repeated, predictable visual patterns to generate visual anticipations, similar to how they predict the trajectory of an inanimate object rolling behind an occluder ([von Hofsten, Kochukhova, & Rosander, 2007](#); [Wentworth & Haith, 1998](#)).

To address this limitation, [Cannon and Woodward \(2012\)](#) developed a novel eye-tracking paradigm to measure infants' visual anticipation of goal-directed actions. In this paradigm, infants watched three familiarization events in which an agent, either a human hand or a claw, performed the goal-directed action of reaching for and grasping one of two objects. The objects' locations were then switched, and the agent reappeared but stopped midway between the objects (i.e., did not contact either object). Cannon and Woodward then measured 11-month-olds' visual predictive fixations from the agent to an object to determine whether they engaged in an on-line goal analysis of the agent's actions. Indeed, infants generated visual predictions to the prior goal object when viewing a human hand; in contrast, infants generated visual predictions to the prior location object when viewing a claw. By removing the confound of movement and trajectory cues at test, this paradigm revealed that 11-month-old infants generate visual predictions for human agents based on the goal structure of an event. Further, using an adaptation of this paradigm, [Krogh-Jespersen and Woodward \(2014\)](#) found that 15-month-old infants required more time to generate goal-based visual predictions than simpler, movement-based predictions, which suggests that there may be a cognitive burden associated with generating goal predictions.

Given the existing evidence that infants visually predict the goals of single individuals, one question of interest is whether infants can recruit their knowledge to predict the goals of multiple individuals engaged in more complex social interactions, such as collaborations. On the one hand, collaborative activities may be more difficult for infants to encode given the structure of the social interaction, as collaborations involve multiple individuals who produce actions that may or may not be the same, but are jointly necessary for attaining a shared goal. Moreover, collaborative partners' actions often reflect abstract higher-order goals and thus, they often do not reflect immediate goal satisfaction. For example, when working together to bake a cake, one person may be mixing the batter while another turns the stove on. In this example, the actions of each individual are quite different on the surface, however, both are causally instrumental to attaining the shared goal of producing a delicious dessert. On the other hand, evidence from habituation studies indicates that infants represent the actions of individuals engaging in a collaborative activity in terms of the shared goal by 14 months of age ([Henderson & Woodward, 2011](#)), and this understanding is evident in 10-month-old infants following a brief collaborative training intervention ([Henderson, Wang, Matz, & Woodward, 2013](#)). It is yet unknown whether infants can use their understanding of collaboration to appropriately predict the actions of collaborative partners towards the shared goal.

A first step toward insight comes from [Fawcett and Gredebäck \(2013\)](#) who used eye-tracking to investigate whether 18-month-old infants rely on social engagement cues as an indicator of collaboration. Infants watched two actors engage in a sequence of actions in which each actor moved a block from one location to another. Half of the infants saw the actors engage in a friendly interaction (e.g., smiling at one another) before moving the block, whereas the other half saw the same block-moving events without the social interaction. At test, only one of the collaborators was present and infants' visual anticipations regarding her next action were measured. Infants were more likely to predict that the actor would continue to act toward the shared goal, even though she was alone, when they had previously watched her engage in a friendly interaction compared to when she had not. Fawcett and Gredebäck's study sheds light on infants' understanding of the importance of social cues to collaborative interactions, however, two issues are evident: 1) this study does not clarify whether infants structured their understanding based on the individuals' actions, as only the presence or absence of social engagement cues was manipulated; and 2) this study includes the aforementioned confound in that movement and trajectory information was evident during the test trials. In a follow-up study, [Fawcett and Gredebäck \(2015\)](#) controlled for the second issue by inserting an occluder into similar test trial events, yet they used a different measure of infants' action understanding. They examined 18- and 14-month-olds' increase in overall attention rather than visual anticipations. Here, they found that 18-month-old infants attended more toward the location where the shared goal would be completed following a socially engaged interaction than a nonsocial interaction. This increase in attention due to the social engagement of two actors was not evident for the 14-month-old infants. As such, older infants rely on clues toward social engagement when representing a shared goal, but the question remains regarding whether infants understand higher-order collaborative goals based on the manual actions of multiple individuals.

The current study examines 14-month-old infants' ability to predict an actor's goal after viewing the actor engaging in either a collaborative or noncollaborative (i.e., on-looker) interaction with another actor. The study design was based on [Henderson and Woodward's \(2011\)](#) habituation paradigm, which examined infants' retrospective understanding of collaborative events. The use of eye-tracking methodology in the current study allowed us to ask whether infants could rapidly recruit their understanding of higher-

order shared goals to predict an actor's future behavior in real-time. It has been suggested that habituation and eye-tracking measures reflect differences in infants' understanding of social events and their ability to use that understanding to generate predictions regarding future scenarios (see [Woodward & Cannon, 2013](#) for a more detailed discussion of prospective vs. retrospective abilities). During infant-controlled habituation paradigms, infants have a generous time window to encode events that they have just watched, whereas when utilizing eye-tracking anticipation measures, infants must rapidly recruit and deploy their knowledge of an event to generate a visual prediction. If infants at 14 months possess a robust understanding of the shared goal structure of collaboration (as argued by [Henderson & Woodward, 2011](#)), they should predict that the actor will behave in a manner consistent with attaining the shared goal, even in the absence of the social partner. However, it is also possible that infants' ability to demonstrate an understanding of collaborative goals in a habituation paradigm might emerge earlier than their ability to use their understanding to anticipate the actions of collaborative partners. If this were the case and 14-month-old infants' understanding of collaboration is not as robust as it would be in older infants (as the collaborative behavioral differences between 14- and 18-month-olds demonstrated by [Warneken & Tomasello, 2007](#) suggest), infants may predict that the actor will continue to act in a manner consistent with her immediate (i.e., instrumental) goal, but not the shared goal.

2. Study 1

2.1. Participants

Thirty-two 13- to 15-month-old infants participated in this study ($M = 14;14$ months, range: 13;16–15;21 months). Half of the infants were randomly assigned to the Collaborative condition (8 males, 8 females; $M = 14;10$ months) and half to the On-looker condition (8 males, 8 females; $M = 14;17$ months). Infants' ages were not significantly different between conditions, $t(30) = 1.13$, $p = .27$. All infants were full term (i.e., minimum 37 weeks gestation). Participants were recruited from an urban population, and were 41 % White, 28 % African American, 16 % Hispanic, 9 % Asian and 6 % multiracial. An additional 14 infants were tested and excluded from further analysis due to distress (3) or failure to generate a visual prediction during either test trial (6 in the Collaborative condition and 5 in the On-looker condition). This high number of exclusions reflects the strict criteria for the definition of a visual prediction in the current study (see Test Trials section for details) and it is important to note that each condition experienced similarly high exclusion numbers.

2.2. Procedure

All procedures performed in studies involving human participants were in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Infants viewed videos presented on a 24-inch monitor equipped with a Tobii T60XL corneal reflection eye-tracking system (accuracy 0.5°, sampling rate 60 Hz). Infants were seated in their parents' laps approximately 65-cm from the monitor. Calibration was performed with a 9-point procedure using the standard infant calibration setting. When necessary, the calibration process was repeated to improve accuracy. The data were collected and analyzed using Tobii Studio (Tobii Technology, Sweden). The order of the video presentation for both conditions was the following: one welcome trial, three familiarization trials, one orientation trial, and two test trials (for a schematic, see [Fig. 1](#)). The videos were filmed to match the interactions presented in [Henderson and Woodward \(2011\)](#).

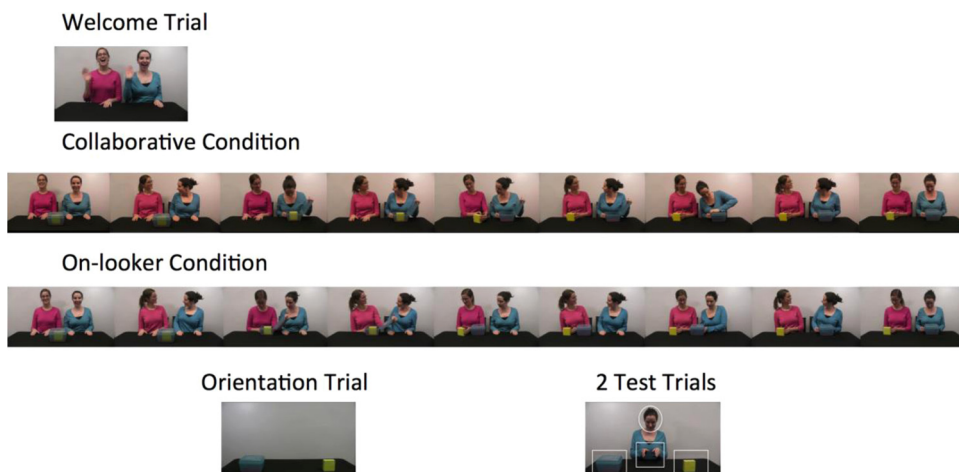


Fig. 1. Key frames from the Collaborative and On-looker conditions. All infants saw one welcome trial, three familiarization trials, one orientation trial, and two identical test trials. The Areas of Interest are presented in white for the test trials. Videos of the events are presented in supplementary materials.

2.2.1. Welcome trial

This single 7-second video introduced two women, who sat next to each other with their gaze down. They each took turns looking up, smiling, and waving while saying “hi”, with both directing their attention to the camera and not each other. They directed their attention straight-ahead and not to each other. The sides that the women sat during the interactions were held constant within infants yet counterbalanced across infants. No objects were present during this trial.

2.2.2. Familiarization trials

Infants then viewed three 25-second familiarization videos. In the Collaborative condition, two women engaged in a collaborative interaction with the shared goal of getting a toy block out of a box to play with it. The video started with the two women saying “hi” individually toward the camera and then smiling at each other. One woman (the box-opener) then reached for a transparent blue-tinted box that had a yellow toy block visible inside of it. She opened the box and then turned to smile at the other woman. The other woman (the toy-getter) reached into the box and removed the toy block from the box. The toy-getter then played with the toy block while the box-opener watched. The women then looked at each other and smiled. The video concluded with the box-opener closing the box, the women smiling at each other, and then looking down at their hands holding their respective objects.

In the On-looker condition, the two women engaged in an interaction that did not have a shared goal, as one woman observed the other perform the necessary actions of getting the toy block out of the box to play with it. To start, each woman said “hi” individually and then smiled at each other. Next, one woman (the toy-getter) opened the box and then smiled at the other woman (the on-looker). Following this, the toy-getter removed the toy from the box and played with it. She then smiled at the on-looker. After this, the on-looker reached for and grasped the box and brought it to rest in front of her. The women smiled at each other again and then looked down at their hand holding their respective objects.

Across conditions, infants viewed the same number of social cues to suggest that the women were engaging with each other; however, the individual actions differed in their contribution to attaining the shared goal of getting the toy block out of the box. Counterbalanced factors included the side of the toy-getter, the location of the box with the block inside of it at the start of the trial, and which woman acted as the toy-getter across infants.

2.2.3. Orientation trial

Following the familiarization period, infants viewed a single 5-second switch trial in which the toy block and box were presented on opposite sides of the table. The sides of the objects during the switch trial were always the opposite of where they appeared during the familiarization trials (e.g., the final location of the box in familiarization was on the left; therefore, it appeared on the right for the switch and test trials). The side that the objects appeared upon was counterbalanced across infants. No actor was visible for this trial. To keep infants’ visual attention, the sound of a ringing bell was played.

2.2.4. Test trials

Infants in both conditions then viewed the same 7-second test videos, with each infant viewing two identical test trials. During the test videos, the woman who acted as the box-opener/on-looker was seated alone at the table. The toy block and the box were located on the sides of the table that matched their locations during the orientation trial. The woman said “hi” toward the camera and then began to reach for an object, but did not complete the reaching behavior; rather, her hands, which were ambiguously positioned to avoid revealing her intended goal, paused centered between the objects (see Fig. 1). The woman did not contact or look at either object during the test trials.

Our primary unit of measurement was infants’ predictive fixations to the objects. As the actor never completes the reaching action during the test trials, this measurement reflects infants’ ability to predict which object the actor will reach toward in the future. Participants could generate a visual prediction at any time point from the start of the video during these trials, yet the eye movements had to meet the following criterion: A predictive fixation is defined as a visual fixation to the actor’s hand Area of Interest (AOI) followed by a visual fixation to either the toy block AOI (i.e., the shared goal) or the box AOI (i.e., the individual goal). Therefore, for each trial, infants’ visual predictions were coded as either to the shared goal object or the individual goal object. Subsequent visual fixations following this visual anticipation were not coded. If no predictive fixations occurred, the trial was coded as “No Prediction”.

2.2.5. Data reduction

Fixation data were extracted from Tobii Studio to calculate where and when infants fixated during the familiarization, orientation, and test trials using the data tools available in the program, which include calculating total fixation durations to the Areas of Interest (AOIs) and the order in which infants fixated to the relevant AOIs. During the familiarization trials, the following four AOIs were generated: two AOIs to encompass each woman’s face, one AOI to encompass the area in which the actions on the objects occurred, and one AOI to encompass the whole screen.

During the test trials, five static AOIs were created to encompass the woman’s face, her hands, the box, the toy block, and the entire viewing screen (see Fig. 1). Although the objects had differing heights and widths, the AOIs were drawn to be identical shapes and sizes to allow for comparable data collection. The AOIs for the objects were located equally distant from the hand AOI during the test trials. Given that the test trial stimuli were identical across conditions, the AOI sizes could not account for differential predictions.

The Tobii fixation filter was used to define fixations, which is the default fixation algorithm for Tobii Studio. A fixation was defined as a stable gaze (within 0.75 visual degrees) for a minimum of 200-ms. Saccades through an AOI without a fixation within the AOI were not coded as visual predictions (as defined above).

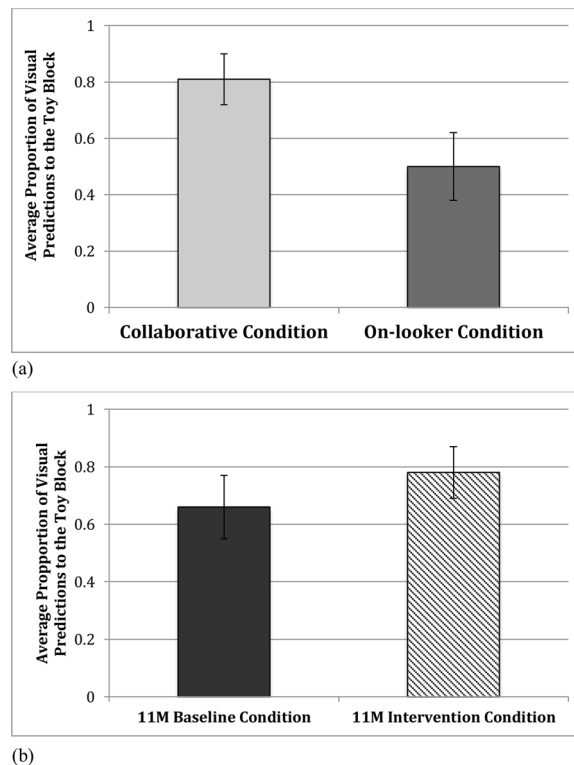


Fig. 2. Means and standard errors for visual predictions to the toy block averaged across the two test trials by condition for (a) the 14-month-old infants in Study 1 and (b) the 11-month-old infants in Study 2.

3. Results

The average percentage of fixation data collected, as a measure of total visual attention to the events as they were presented on screen, did not differ between the Collaborative condition ($M = 83.7\%$, $SD = 13.55$) and the On-Looker condition ($M = 83.1\%$, $SD = 19.19$), $t(30) = .10$, $p = .92$. To examine whether infants perceived the events in terms of shared goals, their predictive fixations were scored per trial as to the shared goal (i.e., the toy block), the individual goal (i.e., the box), or as no prediction (i.e., infants did not generate a fixation from the agent's hand to an object). A mixed-effect regression with condition (Collaborative vs. On-looker) and test trial (test trial 1 vs. 2) as fixed effects, and participant as a random effect revealed condition significantly predicted infants' visual prediction responses (Shared Goal vs. Individual Goal vs. No Prediction) (condition: $B = .40$; $t(61) = 2.13$, $p = .037$; test trial: $B = -.09$; $t(61) = -.09$, $p = .63$). Importantly, all 32 infants in the final sample generated a visual prediction on at least one trial, yet the meaning of a lack of visual prediction is unclear. Therefore, further analyses exclude these trials, which resulted in 9 individual trials excluded for the Collaborative condition and 14 individual trials excluded for the On-looker condition (see Krogh-Jespersen & Woodward, 2014 for similar coding criterion for visual prediction data). Predictive fixations were scored binomially with a score of 1 reflecting a visual prediction to the toy block (i.e., the shared goal) and a score of 0 assigned to the box (i.e., the individual goal). Infants' predictive fixations were averaged across the two test trials, revealing that infants in the Collaborative condition were more likely to generate visual predictions to the toy block than those in the On-looker condition ($t(30) = 2.08$, $p = .047$, $d = .73$; Means and standard errors are presented in Fig. 2a). Tests against chance (test value = .5) revealed that infants in the Collaborative condition generated visual predictions to the shared goal at above chance levels ($t(15) = 3.48$, $p = .003$), whereas those in the On-looker condition were at chance ($t(15) < .001$, $p > .999$). These patterns support the conclusion that infants in the Collaborative condition attributed the individual actions of the box-opener as being directed toward the attainment of the shared goal, whereas infants in the On-looker condition may have been unsure of the on-looker's individual goal. One possibility is that infants could respond based on instrumental goals, supporting an associative account of our current findings. In this account, the box-opener/on-looker should reach for the box during the test trials as this was the object that she acted upon during the familiarization events. However, our results indicate that this pattern of responding was not evident, as infants did not consistently generate visual predictions to the box in either condition.

One alternative possibility for the divergent visual prediction patterns on the test trials across the conditions is that infants' attention was entrained differently during the familiarization events. To examine this possibility, infants' attention to three AOIs, one for each of the women's faces and one for the area in which the actions on the objects occurred, was calculated as a proportion of their total attention to the whole screen AOI and averaged across the three familiarization trials (see Fig. 3a for means and standard errors). A repeated measures ANOVA revealed a significant main effect of AOI ($F(2,60) = 93.36$, $p < .001$), with higher levels of

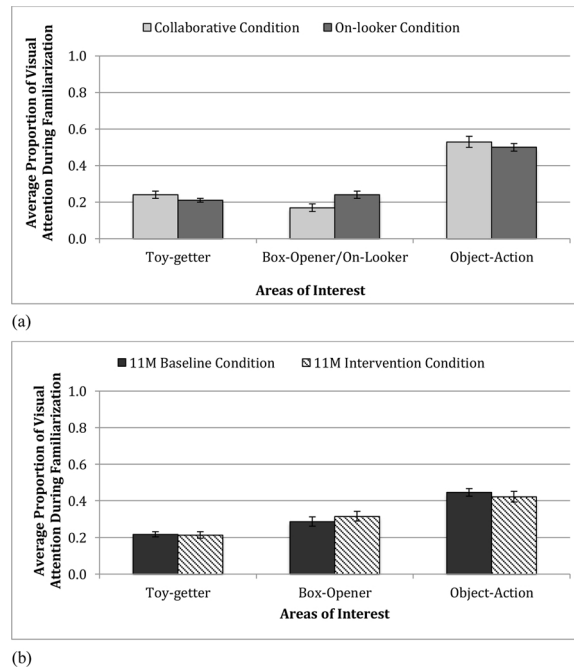


Fig. 3. a and b. Means and standard errors for infants' attention to the three AOIs averaged across the three familiarization trials by condition for (a) the 14-month-old infants in Study 1 and (b) the 11-month-old infants in Study 2.

attention paid to the object-action AOI than to each of the women's faces in both conditions. Given that the actions that the women performed were evident in the object-action AOI, this result is consistent with infants attending to their actions during the familiarization event. Neither the main effect of condition nor the interaction were significant, $F(1,30) = 1.09$, $p = .31$ and $F(2,60) = 2.36$, $p = .10$, respectively.

Another possibility is that infants may have shown a visual preference for the toy block, as they may have considered this to be more interesting than the box. To examine this, infants' attention to the toy block and box during the orientation trial was calculated as a proportion of their total attention. A repeated measures ANOVA revealed no main effect of AOI ($F(1,30) = .30$, $p = .59$) or condition ($F(1,30) = .002$, $p = .96$) and no significant interaction ($F(1,30) = .30$, $p = .59$). Moreover, infants' attention to either object during the orientation trial did not correlate with their visual predictions during the test trials (all $ps > .18$). Therefore, infants' attention to the objects does not appear to have influenced their generation of higher-order goal predictions.

A final analysis examined infants' attention during the test trials to determine whether they viewed these events differently across conditions. Four AOIs were created, one for the woman's face, one for her hand, and one for each object (i.e., the toy block and box). Infants' fixation durations to these AOIs prior to generating a visual prediction were averaged across the test trials. A repeated measures ANOVA indicated a main effect of AOI ($F(3,90) = 146.92$, $p < .001$), revealing higher levels of attention paid to the person than to the objects (see Fig. 4a for means and standard errors). The main effect of condition was not significant ($F(1,30) = 1.07$, $p = .31$) and no interaction was evident ($F(3,90) = 1.11$, $p = .35$). Thus, differences in attention prior to generating visual predictions were not evident across the two conditions.

4. Discussion

The results of Study 1 provide the first evidence that 14-month-old infants are able to recruit their representation of shared goals rapidly to predict an individual's future responses based on past collaborative behavior. After viewing two individuals engaged in a collaborative activity, 14-month-old infants were able to predict that an individual actor would reach toward the object that was the shared goal of the collaboration. When viewing two individuals engaged in a non-collaborative (i.e., on-looker) activity, infants did not generate systematic predictions regarding the future behavior of an individual actor. These findings indicate that early in the second year after their birth, infants can perceive others' collaborative actions as structured by shared, abstract goals.

Additionally, the current findings address limitations within the existing body of research examining collaboration using eye-tracking measures. Infants in the current study could not rely on movement or trajectory information during the test trials to generate their visual predictions. In fact, these trials presented a novel situation as the infants had never seen an individual perform a reaching action in the context in which they were tested. Therefore, infants were required to implement their knowledge of the shared goal to generate a prediction regarding the most likely outcome of this new situation. Although this may have contributed to greater difficulty in generating visual predictions at this age, it sheds light on how infants perceive two individuals engaged in a collaborative social interaction, as 14-month-old infants were able to generate expectations about how one individual would act when alone.

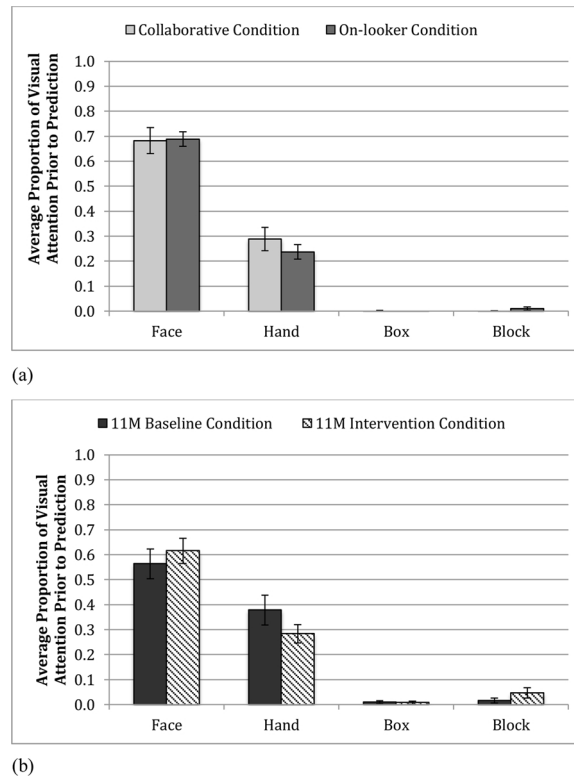


Fig. 4. a and b. Means and standard errors for infants' attention to the four AOIs prior to generating visual predictions averaged across test trials by condition for (a) the 14-month-old infants in Study 1 and (b) the 11-month-old infants in Study 2.

An open question is whether this ability is present by infants' first birthday. Behavioral studies show that, by age 1, infants engage in collaborative games, such as peek-a-boo (e.g., Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Bruner, 1983; Duncan & Farley, 1990; Hubley & Trevarthen, 1979; Ross & Lollis, 1987), and their collaborative interactions become increasingly complex with age (Bakeman & Adamson, 1984; Brenner & Mueller, 1982; Brownell & Carriger, 1990, 1991; Brownell, Ramani, & Zerwas, 2006; Eckerman, Davis, & Didow, 1989; Eckerman & Didow, 1989; Warneken & Tomasello, 2007; Warneken, Chen, & Tomasello, 2006). Interestingly, Ross and Lollis (1987) demonstrated that 9-month-old infants showed frustration when social partners stopped engaging in a collaborative activity, providing evidence that just prior to the end of the first year after their birth, infants perceive that the actions performed by their social partners are necessary to attain the shared collaborative goal (see Warneken et al., 2006 for similar results with older infants). Yet, utilizing a visual habituation paradigm, 10-month-old infants did not successfully represent a collaborative interaction unless they engaged in a collaborative interaction first (Henderson et al., 2013). A growing number of studies show a link between infants' competence in motor performance and their success when viewing third-person action events during visual habituation studies (Gerson & Woodward, 2014; Sommerville, Hildebrand, & Crane, 2008). These findings raise an interesting open question: do infants in their first year demonstrate the ability to predict collaborative goals following an intervention designed to engage their own understanding of collaboration?

Study 2 examines whether 11-month-old infants are able to generate visual predictions based on a shared goal. Here, infants viewed the same video sequence as did infants in the Collaborative Condition in Study 1. Infants in the Intervention Condition engaged in a collaborative intervention before viewing the video sequence, whereas infants in the Baseline Condition only participated in the eye-tracking paradigm.

5. Study 2

5.1. Participants

Thirty-two 11-month-old infants participated in the current study ($M = 11;01$ months, range: 10;10–11;19 months). Half of the infants were randomly assigned to the Baseline condition (8 males, 8 females; $M = 11;00$ months) and half to the Intervention condition (8 males, 8 females; $M = 11;02$ months). Infants' ages were not significantly different between conditions, $t(30) = .31, p = .76$. All infants were full term (i.e., minimum 37 weeks gestation). Participants were recruited from an urban population, and were 56 % White, 16 % African American, 6% Hispanic, 9% Asian and 13 % multiracial. An additional 10 infants were tested and excluded from further analysis due to distress (3) or failure to generate a visual prediction during either test trial in the Baseline condition (4)

and in the Intervention condition (3). An additional 5 infants were excluded because they did not engage with the experimenter and objects during the Intervention condition.

5.2. Procedure

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Review Board and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. After consent processes, 11-month-old infants in the Baseline condition participated in the same eye-tracking procedure as did the 14-month-old infants in the Collaboration Condition in Study 1. Infants in the Intervention Condition participated in a brief collaborative interaction based on Henderson et al.'s (2013) study. During this intervention, infants were seated in the parents' laps in front of a table with an experimenter who sat to their left side. The experimenter picked up an opaque box containing a small bath toy (e.g., a dolphin) that was visible to the infants and placed the box on the table out of reach of the infants. The experimenter directed her attention and gaze toward the box for 20-seconds and then looked at the infants and said, "Should we get it? Let's get it". The experimenter then reached toward the box with two hands, opened the lid, and asked the infants to get the toy out of the box ("Can you get it?"). Once the infants successfully removed the toy from the box, the experimenter praised the infants and then moved to the next box. There were a total of 4 boxes, each with a different bath toy presented in a randomized order. Presentation of the boxes was identical and if infants did not attempt to reach for the toy, the experimenter would place the bath toy in the infants' hands. After the fourth round, infants in the Intervention condition viewed the video series from the Collaborative condition in Study 1. Data coding and analyses were identical to Study 1.

6. Results

The average percentage of fixation data collected did not differ between the Baseline condition ($M = 72.1\%$, $SD = 17.98$) and the Intervention condition ($M = 72.4\%$, $SD = 15.85$), $t(30) = .06$, $p = .95$). To give consideration to the three response choices (Shared Goal vs. Individual Goal vs. No Prediction) as we did in Study 1, a mixed-effect regression with condition (Baseline vs. Intervention) and test trial (test trial 1 vs. 2) as fixed effects, and participant as a random effect did not reach significance with regard to infants' visual prediction responses (condition: $B = -.19$; $t(61) = -.93$, $p = .36$.; test trial: $B = .06$; $t(61) = .31$, $p = .76$). Again, trials in which infants did not generate a predictive fixation were excluded from further analysis, which resulted in 9 individual trials excluded for the Baseline condition and 11 individual trials excluded for the Intervention condition. Importantly, as in Study 1, all infants in the final sample of 32 generated at least one visual prediction across the test trials. Infants' responses were scored binomially and averaged across the two test trials, revealing that infants in the Intervention condition did not differ significantly from those in the Baseline condition ($t(30) = .88$, $p = .39$); Means and standard errors are presented in Fig. 2b). However, tests against chance (value of .5) revealed that infants in the Intervention condition generated visual predictions to the shared goal at above chance levels ($t(15) = 3.09$, $p = .007$), whereas those in the Baseline condition did not differ from chance ($t(15) = 1.43$, $p = .17$). These patterns of responding reveal that 11-month-old infants may be on the cusp of recruiting their understanding of shared goals to generate on-line visual predictions of future behaviors, as the intervention condition seemed to boost infants' propensity to form predictive fixations.

One possibility is that the two conditions may have differed in their attention towards the familiarization event. That is, it is possible that infants in the Baseline condition were not attending to relevant aspects of the Collaborative event. To examine this possibility, infants' attention to three AOIs, one for each of the women's faces and one for the area in which the actions on the objects occurred, was calculated as a proportion of their total attention to the whole screen AOI and averaged across the three familiarization trials (see Fig. 3b for means and standard errors). A repeated measures ANOVA revealed a significant main effect of AOI ($F(2,60) = 31.82$, $p < .001$), with higher levels of attention paid to the object-action AOI than to each of the women's faces in both conditions. Given that the actions that the women performed were evident in the object-action AOI, this result is consistent with the infants attending to the relevant actions during the familiarization event. Importantly, neither the main effect of condition nor the interaction were significant, $F(1,30) = .07$, $p = .80$ and $F(2,60) = .60$, $p = .45$, respectively.

To examine whether infants in either condition had a visual preference for the toy block, their attention to the toy block and box during the orientation trial was calculated as a proportion of their total attention. A repeated measures ANOVA revealed no main effect of AOI ($F(1,30) = .52$, $p = .48$) or condition ($F(1,30) = .08$, $p = .78$) and no significant interaction ($F(1,30) = 2.26$, $p = .14$). Moreover, infants' attention to either object during the orientation trial did not correlate with their visual predictions during the test trials (all $ps > .19$).

Finally, infants' averaged attention during the test trials was examined to determine whether they viewed these events differently across conditions prior to generating a predictive fixation. A repeated measures ANOVA indicated a main effect of AOI ($F(3,90) = 75.60$, $p < .001$), revealing higher levels of attention paid to the person than to the objects (see Fig. 4b for means and standard errors). Neither the main effect of condition nor interaction were significant, $F(1,30) = .46$, $p = .50$ and $F(3,90) = 1.02$, $p = .39$, respectively. Together, these final analyses demonstrate that differences in attention prior to generating visual predictions were not evident between the two conditions, which supports our finding that the intervention shaped infants' ability to form predictive fixations in a similar collaborative action.

7. General discussion

The results from Study 1 suggest that infants may understand collaborative interactions in the form of a simple shared goal by soon after their first birthday. These results are consistent with those of [Henderson and Woodward \(2011\)](#), who utilized a visual habituation paradigm, and extend the previous work to demonstrate that 14-month-old infants not only represent higher-order shared goals, they are able to recruit and deploy this understanding to generate visual predictions regarding an individual's next likely behavior. Moreover, Study 2 reveals that this ability may be evident even earlier in development when 11-month-old infants are given the opportunity to first engage in a similar collaborative interaction themselves. Without a collaborative experience prior to the eye-tracking paradigm, 11-month-old infants did not show any evidence of generating visual predictions to either the instrumental goal or the shared goal, as they generated visual predictions to either goal at chance levels. After engaging in a first-person collaborative interaction that was similar in nature to the experimental stimuli, young infants appeared to accurately represent the actions of the individuals in terms of the higher-order shared goal and, importantly, use this understanding to predict the actor's subsequent actions. These findings are consistent with the findings of [Henderson et al. \(2013\)](#), but extend this work by demonstrating that a lab-based intervention shapes young infants' ability to form visual predictions of an individual's actions towards a shared goal.

The present research revealed that infants can make action predictions towards a shared goal at the same age (i.e., 14-months) at which they have previously demonstrated the ability to identify shared goals (i.e., in the visual habituation paradigm recruited by [Henderson & Woodward, 2011](#)). These findings contrast with the research on infants' understanding of actions involving single individuals, which has shown that infants typically show an action understanding at different ages depending on the paradigms and measures being utilized. For example, [Daum, Attig, Gunawan, Prinz, and Gredebäck \(2012\)](#) used a single paradigm to measure goal sensitivity in infants, which included 1) looking time via visual habituation; 2) anticipatory looking when movement/trajectory information was evident, and 3) goal-based predictive looking without movement/trajectory information. They found that 9-month-old infants showed evidence of goal understanding according to the visual habituation and anticipatory looking measures, but not in the goal-based predictive looking measure. Future work could examine whether infants younger than 14 months (but older than 10 months; [Henderson & Woodward, 2011](#)) might be able to identify shared goals in a habituation paradigm, but not show action prediction towards a shared goal.

The current findings also highlight the importance of first-person engagement in the development of infants' ability to represent collaborative goals. Behaviorally, infants in their first year engage in familiar routines, such as rolling a ball back and forth with a partner, and ritualized collaborative interactions, such as peek-a-boo ([Duncan & Farley, 1990](#); [Hubley & Trevarthen, 1979](#); [Ross & Lollis, 1987](#)). In these routines, infants and their partners produce similar actions. By 11 months, infants visually predict the outcomes of a single individual's goal-directed actions in an eye-tracking study ([Cannon & Woodward, 2012](#)). However, at 10–11 months, infants do not form shared-goal interpretations of collaborative actions either retrospectively when shown the action outcome in a visual habituation paradigm ([Henderson et al., 2013](#)) or prospectively when shown an ambiguous beginning to an action in an eye-tracking paradigm (the current study). Thus, 11 months seems to be an age of transition in which infants can represent the goal of a single individual's actions but have difficulty representing shared goals between multiple agents. That is, unless infants have recently engaged in a similar first-person collaborative intervention with another individual. The current and [Henderson et al.'s \(2013\)](#) findings reveal that infants' ability to identify the goals of multiple individuals is bolstered by their own experience engaging in a similar interaction. Importantly, the current study also reveals that infants can actively apply their knowledge of others' shared goals to a novel situation in which the outcome of an individual's actions is unclear. The first-person intervention may have highlighted the actions produced by each partner in order to successfully determine the shared goal, thereby strengthening infants' ability to identify the shared-goal structure underlying the collaborative action sequence.

During the second year of their lives, infants show increasing sophistication in their understanding of both first-person and third-person collaborative interactions. In a visual habituation paradigm, 14-month-old infants no longer require an intervention to represent the actions of two individuals as collaborative in nature ([Henderson & Woodward, 2011](#)) and the current study shows that infants at this age can actively recruit and deploy this knowledge in real-time to generate a prediction about an individuals' next action based on her shared goals. In first-person interactions, 14-month-olds show basic understanding of the rules for collaborative engagement by making efforts to re-engage a distracted social partner ([Warneken & Tomasello, 2007](#)), with these abilities becoming more sophisticated by 18- to 24-months of age ([Warneken et al., 2006](#)). Taken together, the results from behavioral and visual paradigms provide a consistent perspective of the development of infants' understanding of collaborative goals.

The present findings raise questions surrounding the role that infants' understanding of shared goals (as indicated by their performance on looking time and eye-tracking tasks) might play in their engagement in collaborative actions. One possibility is that infants who demonstrate a stronger understanding of collaboration might be better able to collaborate with others. This possibility is consistent with previous research linking eye-tracking performance with interactive behaviors (e.g., [Hunnius, Bekkering, & Cillessen, 2009](#); [Krogh-Jespersen, Liberman, & Woodward, 2015](#); [Meyer et al., 2015](#)).

In sum, the current study demonstrates that infants can interpret third-party collaborative interactions early in their lives. By 14 months, infants structure collaborative interactions in terms of goals that are shared between social partners and engage this understanding to generate predictions regarding individuals' future goal-based behaviors. By 11 months of age, infants rely on their own first-person experience with collaboration to represent the actions of two individuals engaged in a third-party collaborative interaction and use this representation to predict future behavior. The ability to predict the most likely future outcome of an individual's behavior in collaborative contexts involving multiple actors should facilitate effective social interactions during which infants contingently and appropriately respond to their social partner in order to attain shared goals.

Funding information

This research was supported by grants to A.L. Woodward from NICHD (P01 HD064653 and R03 HD079714) and NSF (DLS 0951489) and to A. Henderson from The Royal Society of New Zealand (Marsden Fast Start).

CRediT authorship contribution statement

Sheila Krogh-Jespersen: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Annette M.E. Henderson:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Funding acquisition. **Amanda L. Woodward:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Funding acquisition.

Declaration of Competing Interest

Sheila Krogh-Jespersen declares that she has no conflict of interest. Annette M. E. Henderson declares that she has no conflict of interest. Amanda L. Woodward declares that she has no conflict of interest.

Acknowledgements

We thank the families that participated in this study.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.infbeh.2020.101446>.

References

- Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., et al. (2013). Looking ahead: Anticipatory gaze and motor ability in infancy. *PLoS ONE*, *8*(7), e67916. <https://doi.org/10.1371/journal.pone.0067916>.
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development*, *55*, 1278–1289. <https://doi.org/10.2307/1129997>.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.
- Biro, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: Development through cue-based bootstrapping. *Developmental Science*, *10*, 379–398. <https://doi.org/10.1111/j.1467-7687.2006.00544.x>.
- Brandone, A. C., Horwitz, S. R., Aslin, R. N., & Wellman, H. M. (2014). Infants' goal anticipation during failed and successful reaching actions. *Developmental Science*, *17*, 23–34. <https://doi.org/10.1111/desc.12095>.
- Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*, 327–341. <https://doi.org/10.2307/2185537>.
- Brenner, J., & Mueller, E. (1982). Shared meaning in boy toddlers' peer relations. *Child Development*, *53*, 380–391. <https://doi.org/10.2307/1128980>.
- Brownell, C. A., & Carriger, M. S. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Development*, *61*, 1164–1174. <https://doi.org/10.2307/1130884>.
- Brownell, C. A., & Carriger, M. S. (1991). Collaborations among toddler peers: Individual contributions to social contexts. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 365–383). Washington D.C: American Psychological Association.
- Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006). Becoming a social partner with peers: Cooperation and social understanding in one- and two-year-olds. *Child Development*, *77*, 803–821. <https://doi.org/10.1111/j.1467-8624.2006.00904.x>.
- Brune, C. W., & Woodward, A. L. (2007). Social cognition and social responsiveness in 10-month-old infants. *Journal of Cognition and Development*, *8*, 138–158. <https://doi.org/10.1080/15248370701202331>.
- Bruner, J. (1983). *Child's talk: Learning to use language*. New York: W.W. Norton and Company, Inc.
- Cannon, E., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science*, *15*, 292–298. <https://doi.org/10.1111/j.1467-7687.2011.01127.x>.
- Cannon, E., Woodward, A., Gredebäck, G., Von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science*, *15*, 35–42. <https://doi.org/10.1111/j.1467-7687.2011.01095.x>.
- Daum, M., Attig, M., Gunawan, R., Prinz, W., & Gredebäck, G. (2012). Actions seen through the babies' eyes: A dissociation between looking and predictive gaze. *Frontiers in Psychology*, *3*, 370.
- Duncan, S., & Farley, A. M. (1990). Achieving parent-child coordination through convention: Fixed- and variable-sequence conventions. *Child Development*, *61*, 742–753. <https://doi.org/10.2307/1130959>.
- Eckerman, C. O., & Didow, S. M. (1989). Toddlers' social coordinations: Changing responses to another's invitation to play. *Developmental Psychology*, *25*, 794–804. <https://doi.org/10.1037/0012-1649.25.5.794>.
- Eckerman, C. O., Davis, C. C., & Didow, S. M. (1989). Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, *60*, 440–453. <https://doi.org/10.2307/1130988>.
- Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science*, *16*, 841–849. <https://doi.org/10.1111/desc.12074>.
- Fawcett, C., & Gredebäck, G. (2015). Eighteen-month-olds, but not 14-month-olds, use social context to bind action sequences. *Infancy*, *20*, 115–125. <https://doi.org/10.1111/inf.12062>.
- Gerson, S., & Woodward, A. (2014). The joint role of trained, untrained, and observed actions at the origins of goal recognition. *Infant Behavior and Development*, *37*, 94–104. <https://doi.org/10.1016/j.infbeh.2013.12.013>.
- Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., & von Hofsten, C. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Developmental Psychology*, *45*, 1190–1194 doi: <http://dx.doi.org.proxy.uchicago.edu/10.1037/a0015667>.
- Henderson, A. M. E., & 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>.

- Henderson, A. M. E., 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>.
- Henrichs, I., Elsner, C., Elsner, B., Wilkinson, N., & Gredebäck, G. (2014). Goal certainty modulates infants' goal-directed gaze shifts. *Developmental Psychology*, 50, 100–107. <https://doi.org/10.1037/a0032664>.
- Hubley, P., & Trevarthen, C. (1979). Sharing a task in infancy. *New Directions in Child and Adolescent Development*, 4, 57–80. <https://doi.org/10.1002/cd.23219790406>.
- Hunnus, S., Bekkering, H., & Cillessen, A. H. N. (2009). The association between intention understanding and peer cooperation in toddlers. *European Journal of Developmental Science*, 3, 368–388.
- Kanakogi, Y., & Itakura, S. (2010). The link between perception and action in early infancy: From the viewpoint of the direct-matching hypothesis. *Japanese Psychological Research*, 52, 121–131. <https://doi.org/10.1111/j.1468-5884.2010.00429.x>.
- Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications*, 2, 341. doi: 10.1038/ncomms1342.
- Krogh-Jespersen, S., & Woodward, A. L. (2014). Making smart social judgments takes time: Infants' recruitment of goal information when generating action predictions. *PLoS ONE*, 9(5), e98085. <https://doi.org/10.1371/journal.pone.0098085>.
- Krogh-Jespersen, S., Liberman, Z., & Woodward, A. L. (2015). Think fast! The relationship between goal prediction speed and social competence in infants. *Developmental Science*, 1–9. <https://doi.org/10.1111/desc.12249>.
- Loucks, J., & Sommerville, J. A. (2012). Developmental changes in the discrimination of dynamic human actions in infancy. *Developmental Science*, 15, 123–130. <https://doi.org/10.1111/j.1467-7687.2011.01099.x>.
- Meyer, M., Bekkering, H., Haartsen, R., Stapel, J. C., & Hunnius, S. (2015). The role of action prediction and inhibitory control for joint action coordination in toddlers. *Journal of Experimental Child Psychology*, 139, 203–220. <https://doi.org/10.1016/j.jecp.2015.06.005>.
- Ross, H. S., & Lollis, S. P. (1987). Communication within infant social games. *Developmental Psychology*, 23, 241–248. <https://doi.org/10.1037/0012-1649.23.2.241>.
- Sommerville, J. A., & Woodward, A. L. (2005). Infants' sensitivity to casual features of means-end support relations in action and perception. *Infancy*, 8, 119–145. https://doi.org/10.1207/s15327078in0802_2.
- 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>.
- von Hofsten, C., Kochukhova, O., & Rosander, K. (2007). Predictive tracking over occlusion by 4-month-old infants. *Developmental Science*, 10, 625–640. <https://doi.org/10.1111/j.1467-7687.2007.00604.x>.
- Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11, 271–294. <https://doi.org/10.1111/j.1532-7078.2007.tb00227.x>.
- Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77, 640–663. <https://doi.org/10.1111/j.1467-8624.2006.00895.x>.
- Wentworth, N., & Haith, M. M. (1998). Infants' acquisition of spatiotemporal expectations. *Developmental Psychology*, 34, 247–257. <https://doi.org/10.1037/0012-1649.34.2.247>.
- 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. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, 22, 145–160. [https://doi.org/10.1016/S0163-6383\(99\)00007-7](https://doi.org/10.1016/S0163-6383(99)00007-7).
- Woodward, A. L., & Cannon, E. (2013). Online action analysis: Infants' anticipation of others' intentional actions. In M. Rutherford, & V. Kuhlmeier (Eds.). *Social perception* (pp. 383–404). MA: MIT Press.