Journal of Experimental Child Psychology 152 (2016) 123–135

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Infants' observation of tool-use events over the first year of life

Klaus Libertus [a,*,1], Marissa L. Greif [d,1], Amy Work Needham [b], Kevin Pelphrey [c]

[a] Department of Psychology and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260, USA
[b] Department of Psychology and Human Development, Vanderbilt University, Nashville, TN 37203, USA
[c] Autism and Neurodevelopmental Disorders Institute, The George Washington University and Children's National Health System, Washington, DC 20037, USA
[d] Florida Atlantic University, USA

ARTICLE INFO

Article history:
Received 14 April 2016
Revised 6 July 2016
Available online 10 August 2016

Keywords:
Event perception
Tool use
Infants
Eye-tracking
Attention
Social cognition

ABSTRACT

How infants observe a goal-directed instrumental action provides a unique window into their understanding of others' behavior. In this study, we investigated eye-gaze patterns while infants observed events in which an actor used a tool on an object. Comparisons among 4-, 7-, 10-, and 12-month-old infants and adults reveal changes in infants' looking patterns with age; following an initial face bias, infants' scan path eventually shows a dynamic integration of both the actor's face and the objects on which they act. This shift may mark a transition in infants' understanding of the critical components of tool-use events and their understanding of others' behavior.

Introduction

Intentional actions begin—by definition—with a goal in mind and are dynamically monitored and adjusted to facilitate goal completion. In contrast, during observation of an action performed by another person, the actor's goal and intentions are not known. Rather, the goal needs to be inferred from the limited information available in the action stream. This leaves room for uncertainty, and

* Corresponding author.
E-mail address: klaus.libertus@gmail.com (K. Libertus).

[1] Klaus Libertus and Marissa L. Greif contributed equally to this article.

http://dx.doi.org/10.1016/j.jecp.2016.07.004
0022-0965/

different interpretations of the same action are possible (e.g., Dretske, 1988). For instance, seeing a person picking up a knife can lead to different interpretations of the intended goal depending on the context (e.g., a cooking show vs. a horror movie). Given enough contextual information, adults quickly form expectations by drawing on their extensive knowledge base acquired over time and are quite good at anticipating the most likely outcomes of observed action sequences (Zacks, 2004; Zacks, Kumar, Abrams, & Mehta, 2009; Zacks & Tversky, 2001). Furthermore, adults show numerous biases linking particular actions with specific types of mental states and dispositions (Iliev, Sachdeva, & Medin, 2012; Malle, 1999; Woolfolk, Doris, & Darley, 2006). Infants, in contrast, are still in the process of acquiring this knowledge and face more ambiguity when interpreting observed actions (e.g., Kuhlmeier, Wynn, & Bloom, 2003). In the absence of extensive experience, how do infants learn about the structure of observed action sequences?

Learning by doing

Infants' own actions play an important role in shaping their interpretations and understanding of observed actions (e.g., Gerson & Woodward, 2010; Hunnius & Bekkering, 2010; van Elk, van Schie, Hunnius, Vesper, & Bekkering, 2008). Correlational evidence for this view comes from studies showing that infants' own grasping skills are related to their ability to predict the goal of an observed action (Kanakogi & Itakura, 2011). Direct evidence comes from training studies where motor skills were selectively manipulated. For example, placing Velcro mittens ("sticky mittens") over the hands of pre-reaching infants allows them to experience successful reaching before they would normally engage in this behavior (Libertus & Needham, 2010; Needham, Barrett, & Peterman, 2002). Experiences of self-produced reaching using sticky mittens have been found to facilitate infants' comprehension of the goal and efficiency of observed actions (Skerry, Carey, & Spelke, 2013; Sommerville, Woodward, & Needham, 2005). Furthermore, reaching training with sticky mittens also fosters infants' perception of complex causal action events (Rakison & Krogh, 2012). Together, these findings suggest that infants learn about actions, their goals, and their structure from their own experiences. A similar position has been put forward in the mirror neuron hypothesis that links action production closely with action perception (Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005; Cannon et al., 2014; Umilta et al., 2001).

Learning by observing

First-hand action experiences are clearly important for learning, but this is not the only way in which infants' understanding of observed actions grows over the first year of life. During a regular day, infants have multiple opportunities to observe others engage in goal-directed activities. Infants learn from such observations, and these experiences influence the development of infants' ability to parse goal-directed events (Esseily, Nadel, & Fagard, 2010). Critically, infants also observe countless actions performed by others that they cannot yet produce themselves. Are infants also able to learn from observation of such never produced actions?

There is evidence that infants can indeed learn from observation alone. For example, infants' observation of a series of visual images containing co-occurrences of certain stimuli and random variation in other stimuli allows them to learn the predictable pairings and respond when these pairings are broken (Fiser & Aslin, 2002). Infants could presumably learn in a similar way about pairings of others' actions and the outcomes of these actions (Buchsbaum, Griffiths, Plunkett, Gopnik, & Baldwin, 2015; Stahl, Romberg, Roseberry, Golinkoff, & Hirsh-Pasek, 2014). Others have shown that infants can learn new skills by watching demonstrations of action sequences (Elsner, 2007; Elsner & Aschersleben, 2003), although the reliable performance of these skills appears somewhat later during the first year of life. For example, Esseily and colleagues (2010) used age-appropriate object retrieval tasks to test infants' ability to learn action sequences from demonstration and found a distinct improvement in sequence reproduction between 10 and 12 months of age.

Importantly, infants also seem to make some sense of observed actions they cannot perform themselves yet. For example, infants can parse adults' actions into meaningful units, including both main action events and transitions that occur between important landmarks in the flow of an action sequence during familiar and completely unfamiliar actions (Baldwin, Baird, Saylor, & Clark, 2001; Hespos, Grossman, & Saylor, 2010; Hespos, Saylor, & Grossman, 2009; Saylor, Baldwin, Baird, & LaBounty, 2007). Thus, infants' ability to parse observed actions into meaningful units does not seem to be exclusively dependent on their own first-hand experiences with the observed actions. By dividing observed actions into meaningful units, infants may also come to better link actions and intended goals of the actor. However, it is not clear what factors facilitate infants' organization of complex sequences, and it is possible that the goal-directed nature and resulting structure of these actions provide critical cues (see Gerson & Woodward, 2013).

From hands to tools

As they grow, infants have multiple opportunities to experience the utility of hands in solving a task (e.g., reaching) and increasingly can hold and manipulate objects that extend their bodies' functionality (McCarty et al., 1999, 2001). However, in many cases, infants have not used the specific tools (e.g., a knife) they observe being used by others. Nonetheless, infants seem to recognize a general event structure in tool-use sequences, which could contribute to emerging tool-using skills in their own behavioral repertoire.

A typical tool-use sequence involves a complex combination of both visible and covert components. For the actor, tool use requires the coordination of means and ends on at least five dimensions: (a) establishing an ultimate goal and breaking it down into sub-goals, (b) matching the physical structures of the tool and target (Wagman & Carello, 2003), (c) planning a motor approach that takes into account the physical constraints of the tool and target, (d) executing and dynamically adapting the action, and (e) knowing when the goal has been successfully attained (for theory and review, see Gerson & Woodward, 2013, and Keen, 2011). To make sense of the action, the infant observer must string these seemingly random actions into a coherent narrative. Indeed, it seems possible that infants may be able to overlay a means–ends structure onto an ongoing tool-use sequence as well as form expectations about congruencies between tool and target to evaluate the utility of the means for achieving the ends. For example, studies using a cloth-pulling paradigm reveal a change in infants' processing of goals between 7 and 12 months of age. In these studies, an out-of-reach toy sits on a cloth, and pulling on the cloth can bring the toy within reach. At 7 months, infants see the cloth as an end in itself and will pick it up to explore it without regard to its connection to the object. However, by 12 months, infants use the cloth to pull in the toy. This suggests that 12-month-olds are able to distinguish between proximal goals (e.g., picking up a tool) and distal goals and become better able to organize this hierarchical structure in goal-related activities performed by others and in their own means–ends problem solving (Sommerville & Woodward, 2005; Willatts, 1999). Furthermore, findings by Gerson and Woodward (2013) suggest that emphasizing the goal is more effective in encouraging infants' own adoption of an effective tool-use strategy than emphasizing the means. Thus, goal completion may be a salient component of tool-use events for infants.

Eye-tracking and looking time studies also provide evidence that infants process sophisticated aspects of object-use events. Specifically, infants expect certain outcomes in goal-directed behaviors performed by others (Gredebäck, Johnson, & von Hofsten, 2010) and by 6 months of age show anticipatory eye movements toward the goal elements of repeated reaching actions (Falck-Ytter, Gredebäck, & von Hofsten, 2006; Kanakogi & Itakura, 2011). Over time, these skills improve; whereas 12-month-olds seem to understand basic functional object–actor behavior (Hunnius & Bekkering, 2010; Kochukhova & Gredebäck, 2010), by 20 months of age infants accurately predict a tool's function based on the way in which an actor initially grasps the tool (Paulus, Hunnius, & Bekkering, 2011). Corroborating these results, Futó, Téglás, Csibra, and Gergely (2010) found that 10-month-olds inferred the presence of two distinct objects behind an occluder after seeing experimenters demonstrate two functions, suggesting that infants expected a one-to-one mapping of functions to objects. Thus, during the first 2 years of life, infants begin to understand the *event structure* of others' manual actions, and this understanding also appears to extend to tool-use events. In addition, young infants might even have some grasp of how functions map onto objects and to their own bodies, which may help them to learn about the link between tool use and goal completion.

The current study

The current study examines infants' spontaneous observation of tool-use sequences using naturalistic stimuli with high levels of complexity. We were interested in whether infants would detect the key components of a tool-use event within an ongoing action stream. Furthermore, we wanted to determine which aspects of tool-use events take priority in infants' distribution of attention. For our purposes, the key components of a tool-use event include visually examining an object before manipulating it, picking up a tool and applying it to the object, and noticing the change in the goal object.

To investigate this issue, we recorded infants' eye gaze during their observation of different tool-use sequences sharing a same general event structure. In these sequences, an actor attempted to manipulate a target object with her hands but failed and then picked up a tool, used the tool on the target object, and finally demonstrated the resolved goal. Infants saw five videos following this sequence depicting different tool-use events. Each tool-use sequence was presented only once. Consequently, infants could not infer tool function through repeated experience with the same tool-use event. In addition, the actor's upper torso was fully visible sitting at a table with a tool and target object in view on the table in order to determine how infants divide their attention between the actor's face and the other relevant aspects of the visual scene (i.e., tool or goal object).

We predicted that younger infants would direct the majority of their looking to faces rather than to other aspects of the scene (e.g., Libertus & Needham, 2014). However, with age and increasing development of selective attention skills, we predicted that infants would eventually direct more attention toward other important elements in the sequence such as the tool used by the actor and the goal object. Based on previous tool-use findings, this shift in looking patterns was expected to occur at around 7 to 12 months of age—the same time as advances in infants' means–ends problem solving emerge (Sommerville & Woodward, 2005).

Method

Participants

Participants in this study were 8 4-month-olds (*M* = 4 months 13 days; 3 girls and 5 boys), 7 7-month-olds (*M* = 7 months 11 days; 4 girls and 3 boys), 8 10-month-olds (*M* = 10 months 19 days; 3 girls and 5 boys), 8 12-month-olds (*M* = 12 months 14 days; 4 girls and 4 boys), and 11 adults (10 women and 1 man). An additional 14 infants did not complete the protocol because of fussiness or distraction during the experiment, and data from 7 infants were discarded because of fussiness before the experiment, resulting in a failure to calibrate. All infants weighed at least 2400 g at birth, and parents did not report any major birth defects or illnesses. Children were recruited via a database of parents who agreed to take part in developmental research. Families were compensated with $5 and an infant T-shirt. Adult participants were recruited through a database of student volunteers and were compensated with $10. Informed consent was obtained from caregivers or adult participants prior to testing. The local institutional review board approved all procedures.

Apparatus

Infants were seated on their parent's lap approximately 60 cm away from a 17-inch computer screen. Participants' gaze position was recorded with a Tobii 1750 remote eye-tracker at a 50-Hz sampling rate (accuracy = 0.5 degree, precision = 1 degree). Prior to the experiment, all participants completed a five-point calibration procedure where an animated character was presented at the center and the four midpoints of each side of the screen.

Materials

Participants watched five videos of an actor using a tool to solve a problem. Videos ranged from 28 to 48 s (mean length = 35.8 s). Problems included opening a box containing a toy (P1), cutting a whole apple (P2), cutting open a cookie package (P3), poking a pom-pom out of a long plastic tube (P4), and cutting a half apple (P5). Videos were presented at the center of the monitor and subtended 21.3 × 16 (H × V) degrees of visual angle.

Each video included three event components (see Fig. 1 for a sample event sequence). First, the actor picked up an object (the "goal") and attempted to act on it with her hands (*goal establishment*). Second, the actor picked up a tool and used the tool to transform the goal object (*tool transformation*). Third, the goal was achieved, and the outcome was clearly presented on the screen (*goal resolution*). The actor's eye gaze was always directed toward the objects, and she never attempted to engage or interact with the viewer. All scenes were filmed against a blue background, and the objects and tools were placed on a black table. The actor was unknown to all participants and was never physically present during the experimental session. Each video was designed to mimic the real-world experience of observing a spontaneous and successful non-pedagogical tool-use attempt.

Procedure

Infants were seated on a parent's lap and presented with five videos, played as a continuous presentation with a 1-s animation of a cartoon character inserted between each video to reorient participants' attention to the center of the screen. Each video was viewed only once, and the procedure ended on completion of the fifth video. The videos were presented in two orders, with each order presented to approximately half of the participants (Order 1: P1, P2, P3, P4, P5; Order 2: P3, P5, P4, P1, P2).

Results

Two complementary approaches to examine infants' processing of the tool-use events were examined. First, focus of attention during the three broad event sequences was examined by calculating participants' fixation durations on key areas of interest (AOIs) in the scenes. Second, gaze patterns were also examined during ambiguous transition phases in the event sequences (i.e., when the actor makes a gaze shift) to determine whether fixations in the first analysis were caused by gaze following behaviors (Moore, 1999) or reflected participants' ability to make higher level conceptual connections between the actor's current activities and possible future components of the event.

Overall fixation duration by event segment

For our first set of analyses, overall gaze durations for AOIs within the visual scene that were placed around the actor's face and left and right hands, the tool, and the goal object were calculated. In addition, two control areas containing only the background were included to test for randomly distributed fixations (see Fig. 2). Only a small number of fixations (<2%) fell into these blank control areas; therefore, the control areas were not analyzed further.

To account for changes in the visual scene as the actor moved the tool or goal object, we split the videos into three sections corresponding to each of the three major event segments: *goal establishment, tool transformation,* and *goal resolution* (see Fig. 2). The AOIs were moved to appropriate portions of the scene within each of these three sections so that target objects were maintained within the AOI during the segments. To adjust for individual differences in total looking time, analyses were performed on the proportion of total fixation time participants spent looking in each AOI during each of the three major events. A separate analysis of variance (ANOVA) was performed on each AOI with age as the between-participants variable.[2] Planned comparisons compared performance in adjacent

[2] Order of presentation was analyzed in an age group *x* order between-subjects ANOVA for all AOIs and video segments. No interaction was ever significant. Additionally, no order main effects were significant for face or tool areas in any video segment. However, there was a main effect of order for goal establishment, $F(1, 32) = 7.90$, $p < .01$, $\eta^2 = .20$ ($M_{Order1} = .25$, $SD_{Order1} = .13$; $M_{Order2} = .31$, $SD_{Order2} = .13$), tool transformation, $F(1, 32) = 6.28$, $p < .02$, $\eta^2 = .16$, ($M_{Order1} = .11$, $SD_{Order1} = .10$; $M_{Order2} = .12$, $SD_{Order2} = 08$), and goal resolution, $F(1, 32) = 5.73$, $p < .03$, $\eta^2 = .15$, ($M_{Order1} = .14$, $SD_{Order1} = .11$; $M_{Order2} = .19$, $SD_{Order2} = .14$). It is possible that watching one order of transformations primed attention in a different way than the other. However, this hypothesis is beyond the scope of the present data and may be answered more systematically in future studies.

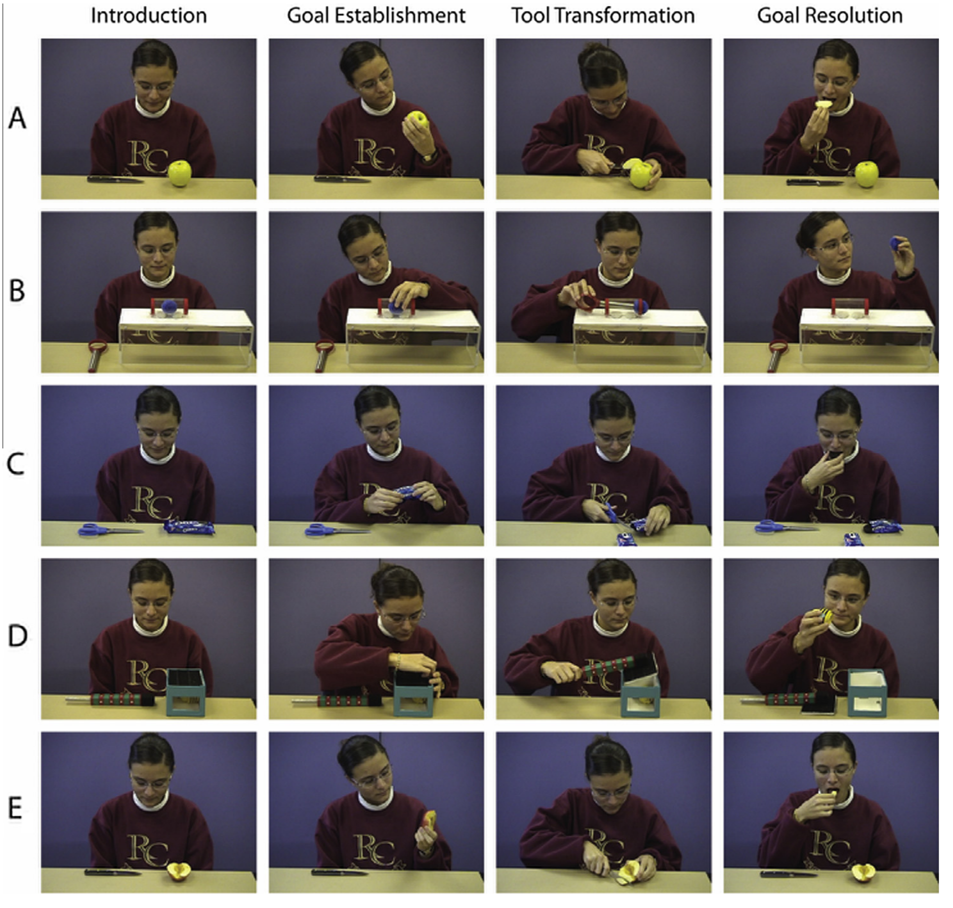

Fig. 1. Examples of the video stimuli used in the experiment. Each row shows stills of the key components of each event (introduction, goal establishment, tool transformation, and goal resolution).

groups in order to assess developmental change. Adjusted *t*- and *p*-values are reported where Levene's tests for equality of variances were significant. Means and standard errors are presented in Table 1; results of statistical comparisons between adjacent age groups are displayed in Fig. 2.

Goal establishment

During goal establishment, the ANOVA revealed significant differences in attention toward the goal object among the five age groups, $F(4, 37) = 5.88$, $p < .01$, $\eta^2 = .39$, with planned comparisons between adjacent age groups indicating an increase in attention toward the goal object between 4 and 7 months of age, $t(13) = 2.34$, $p < .04$, $\eta^2 = .30$, and between 7 and 10 months, $t(13) = 2.12$, $p = .054$, $\eta^2 = .26$. Attention toward the face also differed among age groups, $F(4, 37) = 4.82$, $p < .01$, $\eta^2 = .34$, with planned comparisons indicating a decrease in face attention between 7 and 10 month of age, $t(13) = 2.99$, $p < .02$, $\eta^2 = .41$. Finally, attention toward the tool object also differed among age groups, $F(4, 37) = 17.93$, $p < .01$, $\eta^2 = .66$, with planned comparisons indicating an increase in attention toward the tool between 12-month-olds and adults, $t(17) = 4.16$, $p < .01$, $\eta^2 = .50$.

Tool transformation

While the actor applied the tool to the goal object, ANOVA revealed a significant difference among age groups in their attention toward the goal object, $F(4, 37) = 15.56$, $p < .001$, $\eta^2 = .63$, caused by an

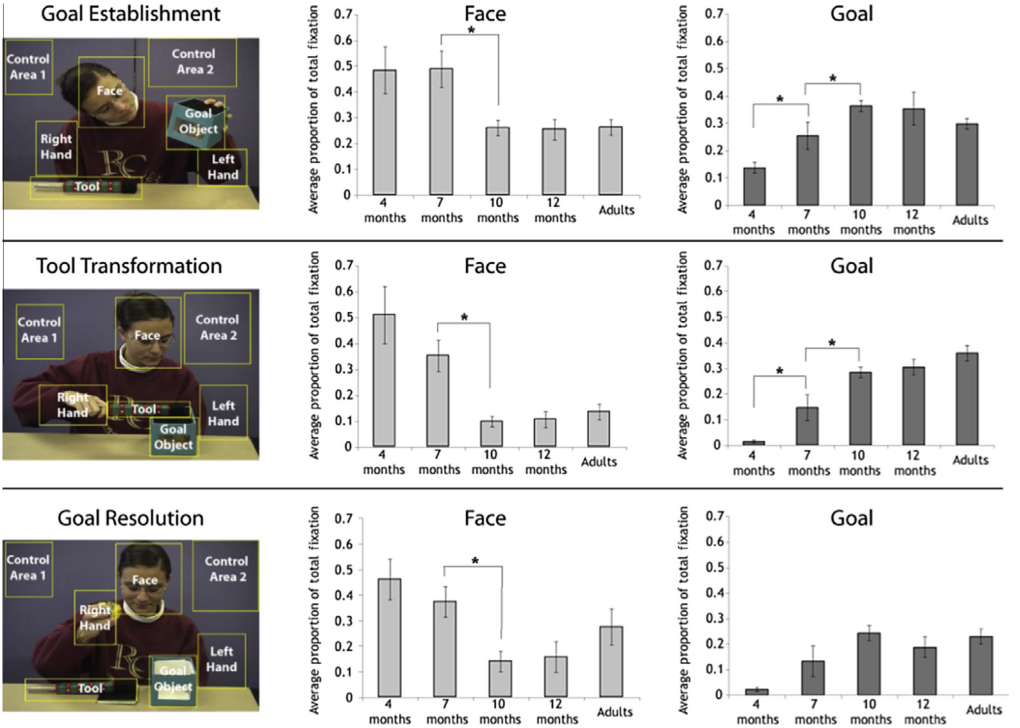

Fig. 2. Proportions of fixation duration on the face and goal by stimulus segment/phase. Representative scenes and the areas of interest from one of the sequences are shown in the left-hand column. [*]$p < .05$.

Table 1
Mean proportions (± standard errors) of fixation during observations of a tool-use sequence.

	Age group				
Area of interest	4 months	7 months	10 months	12 months	Adults
Goal establishment					
Face[*]	.48 (.09)	**.49 (.07)**	**.26 (.03)**	.25 (.04)	.26 (.03)
Goal[*]	**.14 (.02)**	**.25 (.05)**	**.36 (.02)**	.35 (.06)	.29 (.02)
Tool[*]	.00 (.00)	.02 (.007)	.02 (.004)	**.02 (.007)**	**.12 (.02)**
Right hand	.08 (.04)	.02 (.005)	.03 (.01)	.01 (.005)	.03 (.02)
Left hand	.05 (.02)	.04 (.02)	.06 (.02)	.08 (.03)	.04 (.008)
Tool transformation					
Face[*]	.51 (.11)	**.35 (.06)**	**.10 (.02)**	.10 (.03)	.14 (.03)
Goal[*]	**.01 (.005)**	**.15 (.05)**	**.28 (.03)**	.30 (.04)	.36 (.04)
Tool[*]	.04 (.02)	**.09 (.03)**	**.17 (.02)**	.12 (.03)	.14 (.03)
Right hand	.08 (.02)	.11 (.03)	.11 (.03)	.14 (.03)	.06 (.01)
Left hand	.01 (.006)	.01 (.006)	.01 (.004)	.02 (.02)	.01 (.004)
Goal resolution					
Face[*]	.46 (.08)	**.37 (.06)**	**.14 (.04)**	.16 (.06)	.27 (.07)
Goal[*]	.02 (.008)	.13 (.06)	.24 (.03)	.19 (.04)	.23 (.03)
Tool	.03 (.02)	.04 (.02)	.09 (.02)	.09 (.04)	.01 (.002)
Right hand	.10 (.03)	.10 (.03)	.13 (.03)	.13 (.04)	.13 (.02)
Left hand	.03 (.02)	.03 (.01)	.02 (.007)	.04 (.02)	.02 (.004)

Note. Numbers in boldface denote significant adjacent age comparisons.
[*] $p < .05$ for the effect of age in the AOI.

increase in attention toward the goal object between 4 and 7 months of age, $t(13) = 3.00$, $p < .02$, $\eta^2 = .41$, and between 7 and 10 months, $t(13) = 2.48$, $p < .03$, $\eta^2 = .32$. Similarly, significant differences among age groups were evident for participants' attention to the actor's face, $F(4, 37) = 10.46$, $p < .001$, $\eta^2 = .53$, caused by a decrease in face attention between the 7- and 10-month-olds, $t(13) = 4.25$, $p < .01$, $\eta^2 = .58$. Finally, attention toward the tool also differed among age groups, $F(4, 37) = 3.41$, $p < .02$, $\eta^2 = .27$, with interest in the tool increasing between 7 and 10 months of age, $t(13) = 2.28$, $p < .05$, $\eta^2 = .29$.

Goal resolution

During the goal resolution phase, ANOVA revealed a significant difference in attention toward the goal object, $F(4, 37) = 6.21$, $p < .02$, $\eta^2 = .40$, but planned comparisons revealed only a trend toward increased goal attention between 4 and 7 months of age, $t(13) = 1.99$, $p = .068$, $\eta^2 = .23$. Similarly, attention toward the actor's face also differed among age groups, $F(4, 37) = 4.59$, $p < .01$, $\eta^2 = .33$, caused by a significant decrease between 7 and 10 months of age, $t(13) = 3.19$, $p < .01$, $\eta^2 = .46$. However, ANOVA revealed no differences in attention toward the tool.

Visual tracking of scene elements during the actor's gaze shifts

For the second set of analyses, gaze patterns were examined while the actor shifted her gaze between the tools and goal objects. This analysis focused on looking times during two segments in the beginning of each video. In the first segment, the actor held her gaze on the goal object and touched it with her hand. Then, in the second segment, the actor held her gaze on the tool and held it in her hand. If infants were strictly following the actor's gaze, we would expect that they would look at the target of the actor's gaze at the expense of the other object in the scene. Only face, tool, and goal AOIs were included in these calculations and fixation times for each AOI were collapsed across all five videos. Proportions of total looking time were calculated for each AOI and analyzed in separate ANOVAs (see Table 2 for means and standard errors).

Actor's gaze on goal object

When the actor's gaze was focused on the goal object, there was a significant effect of age on proportion of fixation for all three AOIs [face: $F(4, 37) = 5.37$, $p < .01$, $\eta^2 = .37$; goal object: $F(4, 37) = 4.78$, $p < .01$, $\eta^2 = .34$; tool: $F(4, 37) = 5.24$, $p < .01$, $\eta^2 = .362$]. Looking at adjacent age groups, there was a significant decrease in looking at the face AOI, $t(13) = 2.67$, $p < .02$, $\eta^2 = .35$, and a significant increase in looking at the goal object AOI, $t(13) = 2.72$, $p < .02$, $\eta^2 = .36$, between 7 and 10 months of age. The tool area attracted no attention at 4 months of age and only little attention at 7, 10, and 12 months. In contrast, adults spent significantly more time looking at the tool than did 12-month-olds, $t(17) = 2.175$, $p < .05$, $\eta^2 = .22$, although the means are very low and could simply reflect a floor effect.[3]

Actor's gaze on tool

When the actor's gaze was focused on the tool, there was a significant effect of age on proportion of fixation for the tool AOI, $F(4, 37) = 16.62$, $p < .001$, $\eta^2 = .642$, and the face AOI, $F(4, 37) = 4.014$, $p < .01$, $\eta^2 = .303$, but there was no effect for the goal object AOI, $F(4, 37) = 1.763$, $p > .10$, $\eta^2 = .16$. Between 4 and 7 months of age, there was a significant increase in looking at the tool, $t(13) = 2.57$, $p < .05$, $\eta^2 = .34$, which remained unchanged among 7, 10, and 12 months and then increased again between 12 months and adulthood, $t(17) = 3.79$, $p < .01$, $\eta^2 = .46$. Planned comparisons did not reveal any differences in looking at the actor's face between adjacent age groups.

[3] Non-parametric tests showed that 4-month-olds' patterns of looking in the first goal-focused phase were sensitive to the order of video presentation (Face, Mann–Whitney $U = 0$, $p < .05$, $r = .79$; Goal, Mann–Whitney $U = 0$, $p < .05$, $r = .79$). They looked longer at the face in Order 1 (M = .81, SD = .059) than in Order 2 (M = .49, SD = .17). Conversely, they looked longer at the goal in Order 2 (M = .51, SD = .173) than in Order 1 (M = .19, SD = .059).

Table 2
Mean proportions (± standard errors) of fixation during the actor's exclusive gaze on the goal object and tool.

Area of interest	Age group 4 months	7 months	10 months	12 months	Adult
Actor's gaze focused on goal object					
Face*	.61 (.08)	**.57 (.10)**	**.29 (.05)**	.29 (.08)	.35 (.04)
Goal*	.39 (.08)	**.42 (.10)**	**.70 (.05)**	.69 (.07)	.58 (.04)
Tool*	.00 (.00)	.008 (.007)	.008 (.004)	**.01 (.005)**	**.07 (.02)**
Actor's gaze focused on tool					
Face*	.83 (.12)	.64 (.09)	.52 (.07)	.46 (.07)	.45 (.06)
Goal	.15 (.12)	.20 (.07)	.28 (.06)	.33 (.06)	.11 (.03)
Tool*	**.02 (.01)**	**.16 (.06)**	.19 (.04)	**.22 (.04)**	**.45 (.04)**

Note. Numbers in boldface denote significant adjacent age comparisons.
* $p < .05$ for the effect of age in the AOI.

Discussion

Our results suggest that over the first year of life, infants begin to organize their visual assessment of complex tool-use events by increasingly dividing their attention between actors' faces and target objects. At 4 months of age, infants primarily fixate on the actor's face, but between 7 and 10 months a developmental shift occurs. Like the 4-month-olds in our sample, 7-month-olds still primarily fixated on the actor's face but also showed an increased interest in the goal—suggesting a nascent appreciation of some aspects of the relationship between actor and object. By 10 and 12 months of age, infants' general parsing of a tool-use event resembles that of adults—attending to faces while the actor is planning her action (during the goal establishment phase) but attending to the transformation of the goal object as the action is performed (tool transformation phase).

Motor mediation, goal understanding, and selective attention

Three possible mechanisms could support the observed changes in infants' processing of a tool-use event. First, understanding actions might be mediated by a motor mechanism that is facilitated by children's own motor skills (e.g., Sommerville & Woodward, 2005; Stapel, Hunnius, van Elk, & Bekkering, 2010; van Elk et al., 2008). However, such action-mediated understanding does not need to rely on a fully fleshed-out motor solution for specific activities—especially activities that an infant cannot enact yet (see also Southgate & Begus, 2013). For instance, it is possible that infants' own experiences with very simple actions can help them to understand more complex ones. These may include sequences such as touching one object to another, matching object properties to surface characteristics (Bourgeois, Khawar, Neal, & Lockman, 2005), and using analogies between their own bodies (e.g., hands or arms) and tools that extend their bodies' functionality (e.g., a mechanical claw) (Gerson & Woodward, 2012). Even trial-and-error attempts to manipulate objects may underlie comprehension of others' actions (Lockman, 2000).

Second, it is also possible that the shift found in looking after 7 months of age involves a growing understanding of infants' own goal-directed actions. There is considerable evidence that infants increasingly appreciate the parameters of goal-directed actions during the months leading up to their first birthday. For instance, by 12 months of age, infants are well aware that the action of pulling a cloth can be undertaken in an instrumental fashion—to obtain the out-of-reach toy placed on the cloth—whereas 10-month-olds are only beginning to understand this relation (Sommerville & Woodward, 2005; Willatts, 1999). The transition between 7 and 12 months of age reported here is consistent with these observations.

Third, the development of selective attention skills during mid-infancy may allow infants to process tool-use events more efficiently. For infants, faces are highly salient and attract attention (Frank, Amso, & Johnson, 2014; Frank, Vul, & Johnson, 2009; Libertus & Needham, 2014). Once attention has been focused on the face, continued attention at this location can be either facilitated or inhibited. The ability to inhibit previously attended locations seems to increase at around 4 to

6 months of age and may facilitate learning from observation by allowing infants to process all relevant aspects of a naturalistic tool-use event (Amso & Johnson, 2008; Markant & Amso, 2016). Thus, the development of general attention mechanisms may be one agent of developmental change in observational learning. However, selective attention skills seem to develop earlier (4–6 months of age) than the developmental shift observed in the current study (7–12 months). Furthermore, it is unclear whether our results are due to a change in the direction of "inclusion" (infants including more facets of the event to piece together a visual narrative) or to a change in the direction of "exclusion" (older infants using knowledge to focus attention on the more important components). Our data cannot definitively distinguish between these two possibilities, and future research is needed to examine this issue.

Event segmentation

Our results show that infants' ability to understand an event does not depend *only* on their ability to perform the actions themselves. Rather, our findings are aligned with those of Saylor and colleagues (2007), showing that young infants can detect regularities among components of both familiar and unfamiliar event sequences. This is an important skill because events are made up of smaller components or sub-goals and adult observers are able to construe action events as composed of more basic segments or action units (Newtson & Engquist, 1976; Zacks & Tversky, 2001; Zacks, Tversky, & Iyer, 2001; Zacks et al., 2001). Although each event in our study was presented only once and the specific objects in each event varied, the event structure was the same across events; that is, the actor attended to the goal object and tool and then applied the tool to the goal object in order to transform the goal object in some way. This similarity in sequence may have facilitated infants' apprehension and parsing of the event structure. Therefore, although it is likely that knowing how to produce an action facilitates infants' perception of the event, the current findings suggest that being able to perform an action might not be necessary for accurate action perception and goal prediction (see also Southgate & Begus, 2013; Stahl et al., 2014).

The fact that infants can parse specific events that they are not physically capable of fully producing on their own suggests that there could be different levels of processing or understanding for actions. Understanding the relations among actions in an event (including its basic causal structure) could be mediated by a cognitive mechanism such as statistical learning (Saffran, Aslin, & Newport, 1996). For example, once infants have observed a number of similarly structured events during everyday situations, they may come to expect that when an actor initiates an action on an object, the actor will continue to act on it until some sort of resolution is achieved, which itself might be indicated by some regularly occurring behavior. For instance, a mother might open a bag of crackers, take one out of the bag, and eat it. The specific content of the pattern might not be understood; rather, the frequency of pairings of typical behaviors might be averaged over multiple and varied goal-directed events. Over time, infants may become increasingly sensitive to these typical patterns. Future research is needed to investigate this possibility.

This consideration also brings to light a curious finding in our data, namely that infants rarely looked at the tool. It is possible that tools were not as salient to infants as the goal objects. It is also possible that the goals and the tool–target physical connection may have been somewhat opaque for some of our demonstrations. That is, because infants could not clearly understand the relationships between objects, the tool became less relevant in their assessment of the narrative being presented on the screen. The fact that the tools simply occupied less surface area on the screen compared with either the face or goal also could have influenced our results. This issue should be explored in the future by equalizing the size of the tools with those of the goal objects.

Attention to gaze

A final interpretation of our findings is that infants' looking patterns reflect a growing sensitivity to the targets of the actor's gaze and action. Deliberate gaze following seems to emerge at around 10 months of age when infants become sensitive to joint attention cues (Mundy & Newell, 2007; Scaife & Bruner, 1975; Tomasello, 1995). By this account, performance of our 10- and 12-month-olds may result from an emerging gaze-following heuristic to "look where the actor is looking," which may then help infants to learn the contingencies between interesting targets and a person's gaze (MacPherson & Moore, 2007). The shift toward scrutinizing both the goal object and the actor's face at around 7 months of age may reflect this learning process.

Conclusions

Our results suggest that at around 10 months of age, infants' visual scrutinizing of naturalistic tool-use events starts to become more adult-like. This appears to correspond to a similar shift in infants' ability to parse important general landmarks in event structure and their capacity to use objects in goal-directed activities. A complete understanding of this developmental trajectory will be possible only by considering the ways in which infants' understandings of goals, tools, and events come together with their ability to manipulate and control objects.

Acknowledgments

Portions of this work were presented at the 2007 biennial meeting of the Society for Research in Child Development, Boston, MA. This study was supported with funds to Amy Needham and Kevin Pelphrey from Duke University and an NIMH (National Institute of Mental Health) Career Development Award (MH071284) and John Merck Fund fellowship to Kevin Pelphrey.

References

Amso, D., & Johnson, S. P. (2008). Development of visual selection in 3- to 9-month-olds: Evidence from saccades to previously ignored locations. *Infancy, 13*, 675–686.
Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development, 72*, 708–717.
Bourgeois, K. S., Khawar, A. W., Neal, S. A., & Lockman, J. J. (2005). Infant manual exploration of objects, surfaces, and their interrelations. *Infancy, 8*, 233–252.
Buchsbaum, D., Griffiths, T. L., Plunkett, D., Gopnik, A., & Baldwin, D. (2015). Inferring action structure and causal relationships in continuous sequences of human action. *Cognitive Psychology, 76*, 30–77.
Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249.
Cannon, E. N., Yoo, K. H., Vanderwert, R. E., Ferrari, P. F., Woodward, A. L., & Fox, N. A. (2014). Action experience, more than observation, influences mu rhythm desynchronization. *PLoS ONE, 9*(3), e92002.
Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. Cambridge, MA: MIT Press.
Elsner, B. (2007). Infants' imitation of goal-directed actions: The role of movements and action effects. *Acta Psychologica, 124*, 44–59.
Elsner, B., & Aschersleben, G. (2003). Do I get what you get? Learning about the effects of self-performed and observed actions in infancy. *Consciousness and Cognition, 12*, 732–751.
Esseily, R., Nadel, J., & Fagard, J. (2010). Object retrieval through observational learning in 8- to 18-month-old infants. *Infant Behavior and Development, 33*, 695–699.
Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*, 878–879.
Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 15822–15826.
Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology, 118*, 13–26.
Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition, 110*, 160–170.
Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition, 117*, 1–8.
Gerson, S. A., & Woodward, A. L. (2012). A claw is like my hand: Comparison supports goal analysis in infants. *Cognition, 122*, 181–192.
Gerson, S. A., & Woodward, A. L. (2013). The goal trumps the means: Highlighting goals is more beneficial than highlighting means in means–end training. *Infancy, 18*, 289–302.
Gerson, S. A., & Woodward, A. L. (2010). Building intentional action knowledge with one's hands. In S. Johnson (Ed.), *Neoconstructivism: The new science of cognitive development* (pp. 295–313). New York: Oxford University Press.
Gredebäck, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology, 35*, 1–19.
Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous actions: Further evidence. *Neural Networks, 23*, 1026–1032.
Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants' ability to parse continuous actions. *Developmental Psychology, 45*, 575–585.

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: A study of infants' visual anticipations during action observation. *Developmental Psychology, 46*, 446–454.
Iliev, R. I., Sachdeva, S., & Medin, D. L. (2012). Moral kinematics: The role of physical factors in moral judgments. *Memory & Cognition, 40*, 1387–1401.
Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications, 2*, 341.
Keen, R. (2011). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology, 62*, 1–21.
Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development, 81*, 1729–1738.
Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science, 14*, 402–408.
Libertus, K., & Needham, A. (2010). Teach to reach: The effects of active vs. passive reaching experiences on action and perception. *Vision Research, 50*, 2750–2757.
Libertus, K., & Needham, A. (2014). Face preference in infancy and its relation to motor activity. *International Journal of Behavioral Development, 38*, 529–538.
Lockman, J. J. (2000). A perception–action perspective on tool use development. *Child Development, 71*, 137–144.
MacPherson, A. C., & Moore, C. (2007). Attentional control by gaze cues in infancy. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and significance* (pp. 53–75). Mahwah, NJ: Lawrence Erlbaum.
Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3*, 23–48.
Markant, J., & Amso, D. (2016). The development of selective attention orienting is an agent of change in learning and memory efficacy. *Infancy, 21*, 154–176.
McCarty, M. E., Clifton, R. K., & Collard, R. R. (1999). Problem solving in infancy: The emergence of an action plan. *Developmental Psychology, 35*(4), 1091–1101.
McCarty, M. E., Clifton, R. K., & Collard, R. R. (2001). The beginnings of tool use by infants and toddlers. *Infancy, 2*(2), 233–256.
Moore, C. (1999). Gaze-following and the control of attention. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life*. Mahwah, NJ: Lawrence Erlbaum.
Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science, 16*, 269–274.
Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using "sticky mittens" enhances young infants' object exploration skills. *Infant Behavior and Development, 25*, 279–295.
Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology, 12*, 436–450.
Paulus, M., Hunnius, S., & Bekkering, H. (2011). Can 14- to 20-month-old children learn that a tool serves multiple purposes? A developmental study on children's action goal prediction. *Vision Research, 51*, 955–960.
Rakison, D. H., & Krogh, L. (2012). Does causal action facilitate causal perception in infants younger than 6 months of age? *Developmental Science, 15*, 43–53.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.
Saylor, M. M., Baldwin, D. A., Baird, J. A., & LaBounty, J. (2007). Infants' on-line segmentation of dynamic human action. *Journal of Cognition and Development, 8*, 113–128.
Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265–266.
Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 18728–18733.
Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition, 95*, 1–30.
Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*, B1–B11.
Southgate, V., & Begus, K. (2013). Motor activation during the prediction of nonexecutable actions in infants. *Psychological Science, 24*, 828–835.
Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development, 85*, 1821–1826.
Stapel, J. C., Hunnius, S., van Elk, M., & Bekkering, H. (2010). Motor activation during observation of unusual versus ordinary actions in infancy. *Social Neuroscience, 5*, 451–460.
Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum.
Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al (2001). I know what you are doing. A neurophysiological study. *Neuron, 31*, 155–165.
van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *NeuroImage, 43*, 808–814.
Wagman, J. B., & Carello, C. (2003). Haptically creating affordances: The user–tool interface. *Journal of Experimental Psychology: Applied, 9*, 175–186.
Willatts, P. (1999). Development of means–end behavior in young infants: Pulling a support to retrieve a distant object. *Developmental Psychology, 35*, 651–667.
Woolfolk, R. L., Doris, J. M., & Darley, J. M. (2006). Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility. *Cognition, 100*, 283–301.
Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*, 979–1008.
Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., ... Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*, 651–655.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition, 112*, 201–216.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*, 3–21.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130*, 29–58.