Mothers' infant-directed gaze during object demonstration highlights action boundaries and goals

Rebecca J. Brand, *Villanova University*, Emily Hollenbeck, *Northwestern University*, and Jonathan F. Kominsky, *Yale University*

***Abstract*— When demonstrating objects to young children, parents use specialized action features, called "motionese," which elicit attention and facilitate imitation. We hypothesized that the timing of mothers' infant-directed eye gaze in such interactions may provide systematic cues to the structure of action. We asked 35 mothers to demonstrate a series of tasks on objects to their 7- and 12-month-old infants, with three objects affording enabling sequences leading to a salient goal, and three objects affording arbitrary sequences with no goal. We found that mothers' infant-directed gaze was more aligned with action boundary points than expected by chance, and was particularly tightly aligned with the final actions of enabling sequences. For 7- but not 12-month-olds, mothers spent more time with arbitrary than enabling-sequence objects, and provided especially tight alignment for action initiations relative to completions. These findings suggest that infants may be privy to patterns of information in mothers' gaze which signal action boundaries and particularly highlight action goals, and that these patterns shift based on the age or knowledge state of the learner.**

***Index Terms*— eye gaze, infant-directed action, motionese, statistical learning**

I. Introduction

One of the important tasks of childhood is learning to use the myriad artifacts in the human environment. Research has indicated that infants can learn from very subtle patterns in the input [1],[2], making use of cues to object function and actors' intentions [3],[4], among other things. In addition, adults modify the behavior they enact for infants, perhaps highlighting relevant aspects of the input [5],[6]. One such behavior that adults modify and infants are sensitive to is eye gaze [7]-[10]. In the current work, we explore the timing of mothers' infant-directed eye gaze with respect to their actions and the potential role of this cue in highlighting the intentional structure of action.

In order for infants to learn about and ultimately re-enact the actions of others, they have to parse the stream of movement into meaningful units for analysis. Accurate parsing of the action stream is also crucial for verb learning [11]. In addition, in order to engage with tools and artifacts effectively, infants need to distinguish the goal of an action from accidental movements, mere means to the goal, and the like. One clue to assist with parsing and analyzing action comes from the statistical patterns of co-occurrences. Boundaries between action units are characterized by physical regularities to which adults are sensitive [12]-[14]. Such regularities might include sharp changes in arm trajectory, contact and release of objects, changes in direction of bodily movement, and so on. Given the multidimensional and redundant nature of these changes, we might expect infants to be sensitive to them as well [15]. Indeed, infants' skills suggest that they are sensitive to action boundary points by 7-10 months [11],[16]-[17], and action goals by 5-12 months [18]-[19], although the cues they use to make these distinctions are not yet clear. In sum, any pattern that helps distinguish action-boundary from action-midst, and goal-achieving action from other types of action, is likely to help infants in parsing and analyzing action units. Likewise, to the extent that these cues prove useful, researchers should consider implementing them in artificial learning systems.

A. *Infant-directed Action, or "Motionese"*

When parents and other adults interact with infants, they modify numerous aspects of their behavior, including their speech (so-called infant-directed speech, e.g., [20],[21]), as well as their gestures [22], facial expressions [23], and speech-action synchrony [24]-[27]. Csibra and Gergely [28] propose that teaching behaviors such as these are in fact a human adaptation for "natural pedagogy," and that such behavior provided an adaptive advantage as human culture and tool use became increasingly complex. These modifications help direct infants' attention to information as it is presented to them [29]-[33], and seem to support infants' learning of generalizable knowledge [28].

One suite of infant-directed modifications has been documented in adults' demonstration of novel objects to infants and has been dubbed "motionese," or "infant-directed action" [5],[7],[32]. When demonstrating objects for infants, adults use larger, more "square" or indirect movements, as well as more repetition, enthusiasm, turn-taking, and eye gaze. Recent work has indicated that motionese is effective in supporting imitation of novel actions in 2-year-olds [34]. Specifically, children's performance of the actions was

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TAMD.2013.2273057

enhanced in an infant-directed demonstration condition relative to baseline (no demonstration) as well as relative to an adult-directed demonstration.

The benefits of motionese have been hypothesized to work in at least three ways [5]: by enhancing infants' overall attention, by marking action boundaries, and by highlighting the goal-directed nature of actions on objects. Such parental input may work at a largely bottom-up level, by enhancing the regularities that already occur in action (e.g., by making changes in reaching trajectories larger and more salient; [5],[33]), as well as by providing additional regularities (e.g., by using repetition and behavioral cues to mark key points in the action stream). Research with infants suggests that they are particularly attentive when multiple cues occur with the same timing, e.g., lips moving in time with heard speech [15]. If motionese indeed provides such bottom-up regularities, benefits to processing could easily extend to artificial learning systems designed to detect these regularities (e.g., see [27]).

One goal of recent work on infant-directed action has been to explore whether motionese contains features and regularities that make it well-suited to its hypothesized functions of enhancing attention, marking action boundaries, and highlighting goals. Simultaneously, we have been exploring whether infants' responses (i.e., their attention and imitation) when presented with such actions show that these mechanisms are at work [29],[34],[35]. For instance, one study along these lines indicated that motionese is in fact effective in gaining and holding infants' attention ([29]).

A study focusing on repetition found evidence that motionese modifications appear well-suited to serve the other two hypothesized functions: boundary-marking and goal-highlighting [36]. Interestingly, this research suggested that these two functions might interact: mothers appeared to adjust their relative emphasis on boundary-marking depending on whether or not actions were building toward a salient end goal. In one condition, objects afforded an enabling sequence of three actions which led to an interesting goal. For instance, mothers might demonstrate that to open a key safe (here referred to as a "lock box") to find a key, one must press a combination, slide a latch, and pull open the lid. In the other condition, mothers were similarly instructed to perform three actions on each object, but actions were arbitrarily sequenced with respect to one another. For instance, mothers might shake a puzzle toy, twist its adjoining tubes, and tilt it to show off the colorful marbles inside.

Although the distinction between enabling-sequence and arbitrary-sequence objects was not mentioned to mothers, nevertheless they treated the two types of objects quite differently. During arbitrary sequences (e.g., using the puzzle), mothers tended to repeat individual action units several times before proceeding to another unit – presumably assisting infants with parsing. However, for enabling sequences (e.g., using the lock box), mothers tended to cycle through the whole sequence without repeating individual actions – presumably highlighting the salient goal and the necessity of the entire sequence in order. Thus, mothers spontaneously adjusted their repetition patterns in a way that highlighted the *units* within an arbitrary-sequence series but highlighted the *sequence* of units in a series that enabled a salient goal.

Repetition may not be the only aspect of infant-directed communication to highlight boundaries and achievement of goals; several other features seem similarly suited to these functions. Mothers' utterances, for one, appear to be especially well-aligned in time with their movements when the movements make up key steps of achieving a goal [24]. Parents also seem to emphasize the end of an action with strong falling or rising speech intonation and longer silences between actions [32]. Mothers' timing of their turn-taking (i.e., offering the object to the infant) may also be a particularly salient way of marking action boundaries [7].

B. Infant-directed Eye Gaze as an Action Boundary Marker

Eye gaze – specifically, eye gaze directed at the infant – may also function to mark action boundaries and to highlight goal achievement. Eye gaze is important to human communication, particularly in comparison with closely related species [37]. Infants are sensitive to the direction of others' gazes. They prefer when gaze is directed at them rather than averted [9], and show a rudimentary form of gaze following even from birth [38]. By 12 months or earlier, infants understand that adults' gaze is intentionally directed at a target [39]-[41] and they expect (non-self) social partners to look at one another [42]. Given infants' early responsiveness to eye gaze, and in particular, their attraction to gaze directed at them [9], it may be a particularly salient stimulus with which parents can punctuate their demonstrations. In support of this possibility, eye contact during action demonstration is made more frequently and for a longer duration with infant partners compared to adults [7].

We note that infants are likely using gaze in a number of ways simultaneously; certainly, their ability to *follow* the caregivers' gaze will likely play a role in their ability to learn about actions and objects [43],[44]. However, gaze-following is not what is at issue in the current paper. Here we examine infant-directed gaze as a potential low-level cue which may draw infants' attention to action boundaries. Specifically, the goal of the current study was to investigate whether the timing of mothers' gaze at their infants during object demonstrations coincides systematically with action boundaries, and whether this might be particularly true when the action completes a salient goal.

To test this possibility, we measured the timing of mothers' infant-directed gaze onsets, relative to the initiation and completion points of action units. Mothers demonstrated both enabling and arbitrary sequences on objects to their infants [45],[36]. We hypothesized that mothers would use infant-directed gaze onsets to mark action boundaries, especially boundaries around the important event of reaching a goal. Specifically, using an alignment analysis modeled after [46], we first determined the frequency of actions and gazes across the demonstration. Next, we determined how close in time a gaze onset would be to an action boundary if gazes and actions were randomly assorted across the timeframe. Then we

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TAMD.2013.2273057

compared the actual alignment of these events to what the chance alignment would predict. We hypothesized that the closest gaze onset to each action initiation or action completion would be more tightly aligned with that point in time than chance would predict. Finally, we compared the temporal alignment between gaze onset and action boundary for different types of actions and boundaries. We predicted that the alignment of gaze would be even more pronounced when the action unit represents the achievement of a salient goal.

In addition, we hypothesized that gaze would not appear randomly within the temporal gap between action units. Rather, we expected that mothers would be more likely to look *just before* an action initiation, but *just after* the completion of an action, thus providing a "package" of action with the gaze bouts. To test this hypothesis, we measured how often the closest gaze to an action boundary occurred before or after that boundary. We predicted that for action initiations, a higher proportion of closest looks would be preceding the boundary, whereas for action completions, a higher proportion of closest looks would be following the boundary. As with the alignment measure, we expected this packaging effect to be particularly strong for enabling-final actions.

From the middle to the end of the first year, there are striking changes in infants' gaze following, attentional control, and object play, among many others [44],[47],[48]. Further, while previous research indicated many similarities in motionese to infants at 6- to 8- months versus 11- to 13-months [5], there were distinct differences in the number and length of eye gaze bouts [7]. Thus, in the current study, eye gaze was investigated in demonstrations to infants at both 7 months and 12 months of age. We expected that mothers' gaze alignment might change across the first year of life.

II. Method

A. Participants

The sample included 35 mothers who demonstrated each of six objects (or sets of objects) to their infants. Infants were either 7 months old (*n* = 15, 8 males and 7 females, *M* = 7 months, 0 days, *SD* = 26 days) or 12 months old (*n* = 20, 9 males and 11 females, *M* =12 months, 6 days, *SD* = 33 days). Participants were recruited from a commercially-available mailing list and through local libraries in a large city on the West Coast of the United States. Data from 9 additional mothers were excluded due to failure of participant to follow design protocol (1), video or data coding problems (3) and experimenter error (5). Data from this sample were reported in [36], derived from the same task but coded for different variables.

B. Materials

Mothers were asked to demonstrate actions on six objects to their infants. ("Objects" could be a single entity or a related set of entities, but for simplicity each set is referred to here as a single object. See Fig. 1 for two sample objects and Table 1 for the full list with instructions.)

Each object was accompanied by a description of three tasks to complete with it (Table 1). For three of the objects, the three tasks naturally comprised a sequence, each task contingent on the last, to produce a result or goal action.

Fig. 1. Two of the objects provided for mothers. The key safe or "lock box" is an enabling-sequence object, with a series of actions leading to the goal of opening the safe and revealing the key. The "sci-fi puzzle" was an arbitrary - sequence object, with a list of actions that could be done in any order.

Table 1.
Object List and Instructions as Provided for Mothers

Object (set)	Instructions
Enabling-sequence	
Lock box	Your goal is to open the lock box by:
	- pressing in the buttons 1 and 6 and letting go of the buttons
	- pressing down on the upper black square button
	- and lifting the lid
Xylophone chute	Your goal is to make the xylophone chute work by:
	- lifting the blue fabric flap placing the black tube in the hole so that it fits neatly into the inner cardboard tube
	- tossing the yellow cone quickly down the black tube to make it come out the bottom hole
Rock chute	Your goal is to make the rock chute work by:
	- turning the bottom handle on the spiral chute to release the ball from the holding chamber
	- placing the ball in the bottom of the lift
	- turning the blue handle to crank the ball to top -- allowing it to drop onto the spiral chute
Arbitrary-sequence	
Ball of whacks	Your goal is to make the ball of whacks work by:
	- putting the pieces together to form a shape or ball
	- rolling the ball or shape around the table
	- squeezing the ball tightly and firmly to make the pieces fall apart
Colored tubes	Your goal is to make these tubes work by:
	- expanding and contracting the tubes to make an interesting sound
	- twisting them into different shapes
	- snapping them together to form hoops
Sci-fi puzzle	Your goal is to make the sci-fi balls work by:
	- shaking the object firmly back and forth to make a loud sound
	- making the colored balls move from side-to-side
	- and twisting the two sides to allow the balls to move into different tubes

Note: Objects were not labeled as "enabling-sequence" or "arbitrary-sequence" for mothers.

For example, one enabling-sequence object consisted of a key safe or "lock box." Mothers were asked to open the lock box by completing three tasks: (1) pushing two numbered buttons on a keypad, (2) pressing down on a lever to free the lock, and (3) removing the lid to find a key inside. For the other three objects, the three tasks did not comprise a specific sequence and instead might logically be produced in any order. These arbitrary-sequence objects afforded separate actions with no salient goal outcome.

Infants were seated in a high chair facing their mother at the end of a table approximately 82 inches long. Thus, mothers and infants had an unobstructed view of one another, but were too far apart to touch or exchange objects. Mothers' demonstrations were recorded at 30 frames/sec from a digital video camera positioned just above the infants' head.

C. Procedure

Mothers were informed that the goal of the study was to investigate how we communicate information about novel objects to one another. Mothers were led to a small room adjacent to the main laboratory room where the six objects and the instructions for each object were displayed. For each object, a small note card with the three tasks was provided. Each mother was asked to practice the actions while out of sight of her infant.

The mother and her infant were then moved to the main experiment room in order to begin the demonstrations. Orders for demonstration were pre-determined such that enabling-sequence and arbitrary-sequence objects were blocked but objects within a block appeared in a randomly assigned order. Both first block and order of objects within a block were counterbalanced across dyads. Mothers were never made aware that there were two blocks or categories of objects; objects were simply laid out in order and labeled 1-6. Objects were on a small table within the reach of the mother but out of the view of the infant.

D. Coding and Data Analysis

Separate raters, both naïve to the research questions, coded the actions and gaze. Each viewed the demonstration video with the sound off and used SuperCoder [49], a computer program with frame-by-frame replay functions, to mark events within the ongoing flow. For the rater coding the actions, the mother's face was obscured, and for the rater coding the eye gaze, the mother's arms and body were obscured, thus minimizing the extent to which the timing of one feature could inadvertently influence judgments of the timing of the other feature.

The rater coding gaze was told that the demonstrator's audience was seated just below the camera. The infant was not visible and instructions avoided reference to the fact that the onlooker was an infant. The rater was instructed to mark the onset and offset of every gaze directed at that person. Other than the object itself, and in rare instances, the camera, mothers directed their looks to little else in the room, and it was possible to code infant-directed gaze reliably (see below for reliability information).

The rater coding action was provided a list of commonly-seen actions in these demonstrations, as determined by [31]. See Table 2. This included the scripted actions (e.g., pressing, sliding, and opening the lock box) but also other non-scripted actions that mothers regularly displayed. For instance, mothers often shook the key inside the lock box. In addition, we originally envisioned mothers removing the lid to immediately reveal the key (i.e., by tilting the box toward the infant) as the final action in the sequence, but in fact, many mothers used a separate action after lifting the lid to lift and show the key. Therefore, this action was split into two separate codes. Uncommon actions were also identified and labeled "other."

For enabling-sequence objects, all actions were considered *enabling-middle* if these actions occurred in the action sequence but did not directly produce the intended goal. Actions that produced the intended goal in the enabling-sequence stream were labeled *enabling-final*. Thus, for instance, opening the lock box and producing the key were considered enabling-final actions, while all other actions on the lock box were considered enabling-middle, whether they were among the suggested actions (e.g., pressing buttons) or mothers' spontaneous actions (e.g., shaking the box). All scripted and non-scripted actions produced with the arbitrary-sequence objects were considered *arbitrary*.

For the purpose of this project, the coder identified the start and stop point for each of these coded actions. The initiation of an action was defined as when the mother's hand made contact with the object to begin the motion of an action (e.g., touching the buttons on the lock box), or when the trajectory changed markedly from one direction to another between actions (e.g., when the movement on the tubes changed from contraction to expansion). Similarly, the action completion was coded as the moment the object was released or the trajectory changed.

Based on these data, it was possible to determine not only start and stop times for each action and gaze but also the duration of actions and gazes, as well as the absolute number of actions and gazes. Regarding number of actions, each token was counted separately, so if a mother cycled through the three scripted actions two times and did no other actions, this would be a total of six actions during her demonstration.

One quarter of the data were randomly selected for reliability coding by two additional naïve raters (one for gaze and one for action). In this case, the action coder was not provided with a top-down definition of actions, but was given one example (as in the preceding paragraph) and was asked to code all action onsets and offsets using similar criteria.

Because of a glitch in the coding program, absolute start and stop times were not comparable across coding passes, but length between start and stop time (in seconds) could still be computed accurately for each pass and compared. For computing reliability, four summary variables were tallied for each object: number of actions; average length of actions (in seconds); number of gaze bouts; and average length of gaze. Reliability was good for number of actions (Chronbach's alpha = 0.87) and length of actions (0.94) and was excellent for number of gazes (1.00) and length of gaze (0.98).

Table 2.
Scripted and Unscripted Actions Coded for Each Object

Object	Actions Coded
Enabling-sequence	
Lock box	
	- Push buttons
	- Slide latch
	- *Lid off*
	- *Key out*
	- Key in
	- Lid on
	- Shake/rattle
Xylophone chute	
	- Lift flap
	- Tube in opening
	- *Cone in tube*
	- Remove cone
	- Tube out
	- Look through tube
Rock chute	
	- Dump holding chamber
	- Ball onto lift
	- Turn lift handle
	- *Ball drops down chute*
	- Roll ball on table
Arbitrary-sequence	
Ball of whacks	
	- Put pieces together
	- Take pieces apart
	- Roll ball on table
	- Squeeze to explode ball
Colored tubes	
	- Twist
	- Snap together
	- Unsnap
	- Contract
	- Expand
	- Look through
	- Put on body (e.g., wrist, head)
Sci-fi puzzle	
	- Shake/rattle
	- Tilt
	- Twist

Actions in italics were considered "enabling-final" actions; all other enabling-sequence actions were considered "enabling-middle." Actions coded but not labeled here were referred to as "other" and were assigned to the enabling-middle or arbitrary categories, depending on object type.

We modeled our primary analyses after [46], in which the authors computed the chance and actual alignment of two action boundaries within a scene. We adapted this to determine chance and actual alignment between gaze onsets and action boundaries. We reasoned that gaze onset was likely more salient than gaze offset; for alignment analyses, only gaze onsets were analyzed. For each action boundary (initiation and completion), we then selected the closest gaze onset, whether this came just before, just after, or simultaneously with the action boundary. For instance, imagine the following scenario: a mother looks up to check whether her child is attending (gaze onset), turns her gaze toward the object, places her fingers on the lock box buttons (initiation – in this case, of an enabling-middle action), and then removes them a few seconds later (completion of the enabling-middle action). Then she quickly looks up with a "ta-da!" expression (a new gaze onset). We might determine that the closest gaze onset to the action initiation was the preceding gaze and the closest gaze onset to the action completion was the gaze that followed the action. In this way, every action boundary was paired with a gaze onset – whichever one was closest in time. The temporal distance (absolute value in seconds) between the gaze onset and the action boundary was then computed for that pair.

III. Results

A. Descriptives and Preliminary Analyses

Subjects with outliers on any measure (any score more than 3 *SDs* from the mean) were removed from analyses. This resulted in a final sample of 30 subjects (13 mothers of 7-month-olds and 17 mothers of 12-month-olds). (Conducting analyses without removing the outliers did not affect the key findings and the overall pattern of findings was largely the same.) Preliminary analyses revealed no main effect of sex, so data were collapsed across this variable.

We first explored whether mothers treated enabling-sequence objects differently from arbitrary-sequence objects in terms of general time spent and engagement with each object. To do so, we used a 2 (type of object: enabling-sequence or arbitrary-sequence) x 2 (age group: 7 month vs. 12 month) MANOVA with overall time spent demonstrating, the number of gazes directed at infants, the average duration of infant-directed gazes, and the number of actions used for each object as dependent variables. Because of the possibility of differences in the length of the demonstration, number of gazes and actions were measured proportionally (e.g., gazes per minute). Note, however, that using absolute (rather than proportional) values here did not change the findings. The MANOVA revealed a main effect of object type, $F(3,26) = 13.52$, $p < .001$, $\eta_p^2 = 0.61$. There was no significant effect of age group; however, age group and object type did interact, $F(3,26) = 4.36$, $p = 0.01$, $\eta_p^2 = 0.34$.

In order to understand precisely how mothers' demonstrations differed on each dependent variable, we next examined univariate ANOVAs for each measure. For overall demonstration time, there was no main effect of object type. There was a marginal effect of age, with mothers spending slightly more time demonstrating to 12-month-olds (*M* = 138 sec, *SD* = 69) than to 7-month-olds (*M* = 95 sec, *SD* = 63), *F* (1, 28) = 3.43, *p* = 0.07, $\eta_p^2 = 0.11$.

There was also an interaction between age and object, *F*(1,28) = 9.43, *p* = .005, $\eta_p^2 = 0.25$. For 7-month-olds, mothers spent significantly more time demonstrating arbitrary-sequence objects (*M* = 108 sec, *SD* = 73) than enabling-sequence objects (*M* = 81, *SD* = 54), paired samples *t*(12) = 2.49, *p* = 0.028, Cohen's *d* = 0.42. On the other hand, for 12-month-olds, mothers spent marginally more time demonstrating enabling-sequence objects (*M* = 147, *SD* = 73) than arbitrary-sequence objects (*M* = 130, *SD* = 66), *t*(16) = 1.80, *p* = 0.091, *d* = 0.24. See Fig. 2.

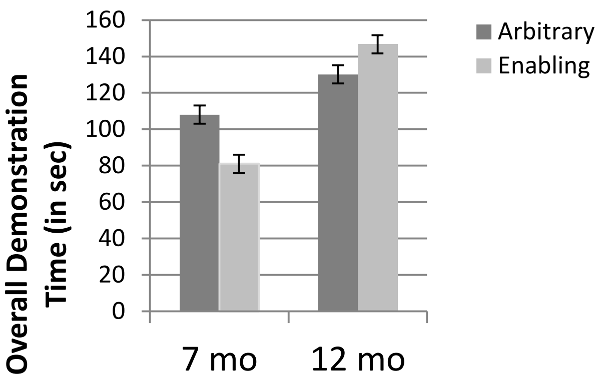


Fig. 2. Mean number of seconds mothers spent demonstrating objects, by age and object type. Error bars in all figures represent standard error.

There was a main effect of object type for both number of gazes and number of actions performed within the demonstration. Mothers gazed at their children more often while demonstrating arbitrary-sequence objects (*M* = 15 gazes/min, *SD* = 6) than when demonstrating enabling-sequence objects (*M* = 9 gaze/min, *SD* = 2), *F* (1,28) = 30.14, $p < 0.001$, $\eta_p^2 = 0.52$. Likewise, they also demonstrated proportionally more actions on arbitrary-sequence (*M* = 13 actions/min, *SD* = 5) compared to enabling-sequence objects (*M* = 9 actions/min, *SD* = 2), *F* = (1,28) = 34.25, $p < 0.001$, $\eta_p^2 = 0.55$. There were no interactions involving number of gaze and number of actions. There were also no main effects or interactions for duration of gaze.

B. *Gaze-Action Alignment Analyses*

1) *Comparison against chance*

In order to determine whether infant-directed eye gaze and action boundaries were systematically aligned, we compared the average temporal distance between gaze onsets and the closest action boundaries with the distance one would expect just due to chance. Conceptually, chance was based on this question: given the number of actions and gazes in a demonstration, how close would they be if they were randomly distributed? For instance, based on the above data, we see that during enabling sequences, mothers performed 9 actions/min and 9 gazes/min. Thus, just by chance, one can expect one action completion and one gaze onset every 6.67 seconds. If gazes were as far as possible from action boundaries (not aligned), we would expect a gaze onset to occur an average of about 3.3 seconds from every action completion. In reality, the average time between gaze onset and action boundaries is much shorter than this, indicating tighter alignment.

In order to quantify this more precisely, we modified a formula used in [46] for estimating the chance temporal distance between two action boundaries. This chance distance, calculated for each demonstration, represents the average distance between action boundaries and gaze boundaries, assuming gaze is randomly distributed with respect to action. In this formula, x = the timestamp beginning the demonstration; y = the timestamp of a given event (an action initiation or completion, which were analyzed separately); z = the timestamp ending the last action in the demonstration; and n = the number of events of that type in that demonstration.

$$Chance\ distance = \{\frac{x^2}{2} + \sum_{i=1}^{n-1}[\left(\frac{y_{i+1}-y_i}{2}\right)^2]\}/z \qquad (1)$$

Chance distance for action initiations and completions were calculated separately for each object for each participant. Average distances (absolute value in seconds) predicted by chance, as well as actual average distances for each action type, are shown in Fig. 3. We then compared the actual average distance between each action boundary and the closest gaze to the chance distance see if that distance was significantly shorter (i.e., more tightly aligned) than chance would predict. For every action type (arbitrary, enabling-middle, enabling-final), and for both types of action boundaries (initiations and completions), gaze was more aligned with action boundaries than expected due to chance, paired-sample *t*s (29) = 3.60 – 7.60, *p*s ≤ 0.001, *d* = 0.91 – 1.74.

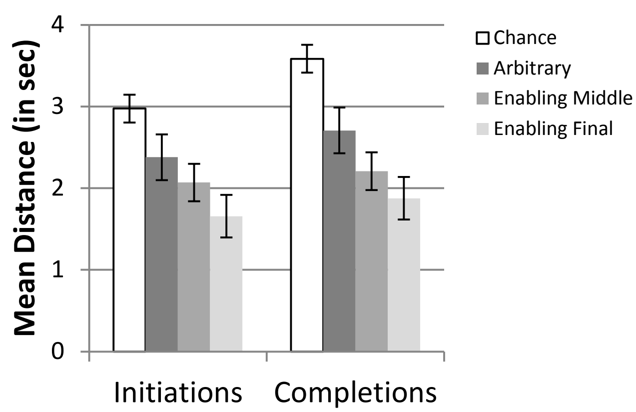


Fig. 3. Mean temporal distance (absolute value, in seconds) between each action boundary and the nearest gaze onset, as a function of action boundary (initiation or completion). Here, "chance" is averaged across all initiations and all completions for illustration purposes. For analyses, chance was computed for each action type separately (arbitrary, enabling-middle, and enabling-final), and all types of action were more aligned than chance would predict.

2) Relative alignment across action types

In order to investigate whether some action types were more likely to be closely aligned with gaze onset than others, we conducted a 3 (action type: arbitrary, enabling-middle, or enabling-final) x 2 (action boundary: initiation vs. completion) x 2 (age group: 7 vs. 12 months) ANOVA with repeated measures on the first two factors. The dependent variable was the absolute temporal distance between action boundary and gaze onset in seconds, allowing us to combine all paired gazes, whether they came just before or just after the action boundary. The dependent variable was log transformed before analysis to reduce skew. See Fig. 4 for mean temporal distances before transformation.

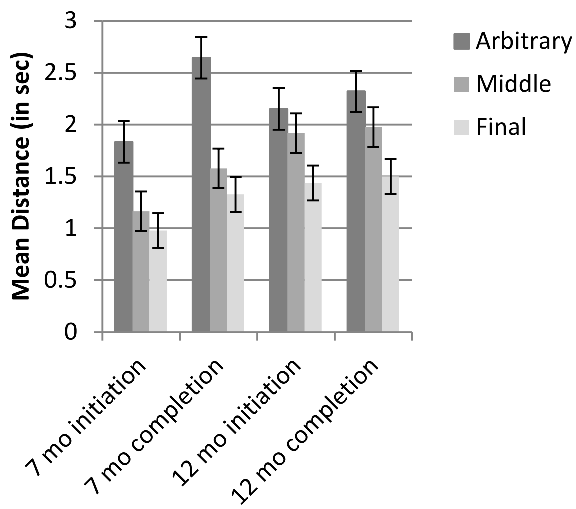


Fig. 4. Mean temporal distance (absolute value, in seconds) between each action boundary and the nearest gaze onset, as a function of age, boundary type, and action type (arbitrary, enabling-middle, and enabling-final). This Fig. shows the consistent pattern of enabling-final actions being the most tightly aligned with gaze.

There was a main effect of action type on gaze-action alignment, F (2,27) = 20.46, $p < 0.001$, $\eta_p^2 = 0.60$. Contrasts indicate that, in line with our hypothesis, the final actions in an enabling sequence (M = 1.31 sec, SD = 0.75) were more tightly aligned with gaze onset than enabling-middle (M = 1.65 sec, SD =0.89) or arbitrary actions (M = 2.24 sec, SD = 0.86), F (1,28) = 42.02, $p < 0.001$, $\eta_p^2 = 0.60$. Enabling-middle actions were also more tightly aligned than arbitrary actions, F (1,28) = 16.08, $p < 0.001$, $\eta_p^2 = 0.37$. Action type did not interact with any of the other variables, indicating that this pattern held across both initiations and completions and across age groups.

There was also a main effect of action boundary, such that action initiations (M = 1.58 sec, SD = 0.74) were on average more tightly aligned with gaze onset than action completions (M = 1.89 sec, SD = 0.93), F (1,28) = 9.38, p = 0.005, $\eta_p^2 =$ 0.25; however, this interacted with age, F (1,28) = 4.27, p = 0.048, $\eta_p^2 = 0.13$. Specifically, for 7-month-olds, initiations (M = 1.33 sec, SD = 0.48) were significantly more tightly aligned than completions (M = 1.85 sec, SD = .99), F (1,12) = 7.14, p = 0.02, $\eta_p^2 = .373$, but for 12-month-olds, the temporal distance between gaze and action boundary for initiations (M = 1.83 sec, SD = 0.87) and completions (M = 1.93 sec, SD = 0.82) did not differ, F (1,16) = 1.08, *ns*.

3) When does gaze precede vs. follow action boundaries?

Finally, we investigated whether the closest gaze onset was more likely to come just before (preceding) or just after (following) the action boundary, and whether that differed across boundaries or action types. A 2 (action boundary: initiation vs. completion) by 3 (action type: arbitrary, enabling-middle, enabling-final) by 2 (age group: 7 months vs. 12 months) ANOVA on the proportion of action boundaries for which the closest gaze was following rather than preceding showed two main effects and an interaction. See Fig. 5.

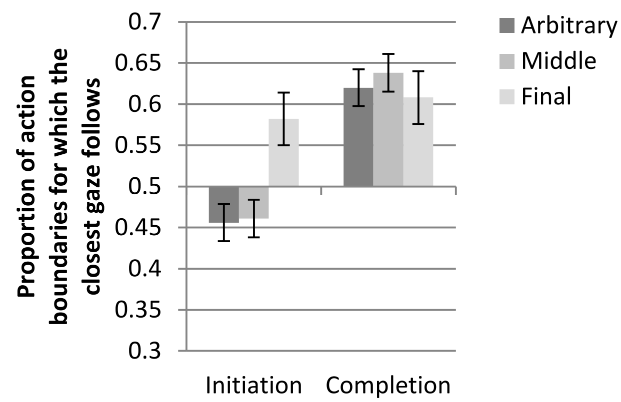


Fig. 5. Proportion of action boundaries for which the closest gaze came immediately after the boundary. Scores below 0.50 indicate a tendency for the closest gaze to come before the boundary.

There was a main effect of action boundary, F (1, 28) = 18.10, $p < .001$, $\eta_p^2 = 0.39$, such that action completions were more likely to be paired with following gazes (M = 0.62, SD = .16) than were action initiations (M = 0.47, SD = 0.06). In addition to comparing initiations to completions, we also asked whether either type was significantly more likely than chance to be paired with a preceding or following gaze. We found that action completions were significantly more likely than chance (.50) to be paired with a following gaze, $t(29)$ = 2.38, p = 0.02, d = 0.75, whereas action initiations were more likely than chance to be paired with a preceding gaze, $t(29)$ = 4.17, $p < 0.001$, d = 0.67. In other words, as predicted, mothers tended to look up just before starting and just after finishing an action.

There was also a main effect of action type, $F(2,27)$ = 3.78, p = .035, $\eta_p^2 = 0.22$. Contrasts indicated that enabling-final actions (M = 0.60, SD = 0.12) were more likely to be paired with a following gaze than either arbitrary (M = 0.54, SD = 0.12) or enabling-middle actions (M = 0.55, SD = 0.13), F(1, 28) = 7.45, p = .011, $\eta_p^2 = 0.21$. Arbitrary and enabling-middle actions did not differ from one another on this measure, F (1, 28) < 1, *ns*.

This main effect of action type appears best explained in the context of the interaction between action type and action boundary, F (2, 27) = 5.15, p = 0.013, η_p^2= 0.28. Post-hoc t-tests show that action completions in all types of actions were equally likely to be paired with a following gaze, *t*s(29) = 0.38-1.15, *ns*, see Fig. 5, right side. However, the pattern was more varied for action initiations (see Fig. 5, left side). Action initiations for enabling-final actions were more likely to be paired with a following gaze (M = 0.59, SD = 0.17) than either arbitrary-sequence (M = 0.46, SD = 0.08) or enabling-middle actions (M = 0.46, SD = 0.09). In fact, comparing each against chance, action initiations for arbitrary-sequence and enabling-middle actions were significantly more likely than chance to be paired with a preceding gaze, *t*s (29) = 2.23-3.07, *p*s < 0.035, while action initiations for enabling final actions were more likely than chance to be paired with following gaze t (29) = 2.72, p = 0.01. In short, the pattern that emerged is that mothers tended to look up at infants soon after completing actions of any sort, and just before most action initiations, but in the case of enabling-final action, mothers were most likely to look up just *after* beginning rather than just *before* beginning.

IV. Discussion

The current study examined the role of infant-directed eye gaze in mothers' demonstrations of objects to their infants. We asked mothers to demonstrate two types of objects: *enabling-sequence*, for which three actions represented a meaningful sequence, with each step contingent on the last, and which produced a salient outcome; and *arbitrary-sequence*, for which three possible actions could be performed in any order and led to no overarching goal. We measured the temporal distance between each action boundary (initiation or completion) and the closest onset of gaze toward the infant. As predicted, we found that gaze onset was systematically aligned with action boundaries, particularly for the final action in an enabling sequence.

Specifically, we found that across all conditions, gaze was more aligned to actions than would be expected by chance. This finding demonstrates that mothers' bouts of eye gaze toward the infant were not randomly distributed throughout their demonstration; rather, they offered systematic information with regard to boundary points in the ongoing stream of action. Whatever other regularities occur at action boundaries[12],[14] the current study indicates that when mothers are specifically demonstrating objects for infants at 7 and 12 months, mothers' own eye gaze toward infants is also a reliable marker of an action boundary.

Gaze-action alignment was prevalent for all action boundaries, but gaze was particularly tightly aligned with the boundaries of actions which resulted in a salient outcome (called enabling-final actions). This supports our hypothesis, that mothers would use eye gaze especially to mark actions which completed a multi-step intention. In fact, we also found that the earlier steps in an enabling-sequence sequence (called enabling-middle actions) contained tighter gaze-action alignment than actions in the arbitrary sequence (although still less than enabling-final actions). This tighter alignment for all actions on objects with a meaningful sequence might indicate that mothers stressed each and every step as important, since each was necessary to reach the goal. The fact that mothers moved more slowly (i.e., provided fewer actions and gazes per second) when demonstrating the enabling-sequence versus arbitrary-sequence objects also supports the interpretation that they were particularly emphasizing the steps in the enabling sequence.

As predicted, the beginning points of action (initiations) tended to be preceded immediately by gaze, whereas the end points of action (completions) tended to be followed immediately by gaze. Thus, moments of gaze onset typically "book-end" the action unit and thus provide a reliable package of information. To the extent that maternal gaze draws attention, not only to the face but to the on-going behaviors (e.g., [31]), infants are provided a lens to focus on key moments of intentional change in the stream of motion. Thus, gaze may serve as a marker of action boundaries or may draw attention to other salient markers of action boundaries, such as trajectory changes. From the perspective of a bottom-up learning system, gaze alone, or gaze in combination with other markers, could potentially provide probabilistic information for parsing. In other words, a system that segmented an action stream at moments of learner-directed gaze would end up with units that correspond fairly closely to those representing intentional actions.

As with the especially tight alignment, enabling-final actions differ from other types when it comes to precisely how gaze is used to package action. Across all action types, the completion point of action units is likely to be immediately followed by gaze, and across most action types, the initiation of actions tends to be preceded immediately by gaze. However, enabling-final actions show a different pattern for action initiations. Packaging is similar to that of other actions in terms of their completions, but the closest gaze to the initiation of the enabling-final actions tends to be immediately after. In other words, the typical course of events is that mothers begin a final action, such as placing their fingers to take the lid off the lock box, and then immediately look at the infant. Thus, rather than using gaze as the standard before-and-after package, mothers typically interrupt that final action to gaze at their infants. Mothers appear to be checking – or eliciting – infant attention to the crucial final action, ensuring that infants see the necessary placement of hands and fingers, and that they are watching when the salient outcome occurs. Along these lines, it appears that to a bottom-up system, gaze might be more useful in drawing attention to other action markers, rather than being a reliable marker itself, as it sometimes comes before and sometimes comes after the action starts.

As in some prior motionese research [7], we found evidence that mothers tailored their interactions with respect to infant age. Specifically, for 7-month-olds, but not for 12-month-olds, mothers spent more time demonstrating arbitrary-sequence than enabling-sequence objects. One possible explanation for this finding is that mothers spent more time repeating the simpler, individual actions of the arbitrary-sequence objects for the younger babies, reasoning that babies could more easily learn the individual actions rather than complex sequences. Indeed, evidence shows marked development in infants' ability to learn sequences across the first few years of life [50]. Future research which includes infant response to similar demonstrations could confirm whether perhaps mothers' time spent on action sequences changes as a function of their infants' attention to and learning of such sequences.

In a second age-related finding, we showed that for 7-month-olds, but not for 12-month-olds, mothers' gaze onset was more tightly aligned to action initiations than completions. It may be that 12-month-olds are more aware of action boundaries (having ample experience with their own object-directed actions by this time) and thus mothers intuit that they do not need such a tight alignment to signal the beginning of an action. Overall, these age differences align with findings in other domains that maternal interaction changes systematically with development [51],[21], but future research on the feedback loop between tutor and learner will likely be necessary to establish the specific function of these changes.

There are a number of potential explanations of our finding that mothers' eye gaze is aligned with their action. One possibility is that mothers intuitively know that action initiations and completions are important points for their children to see, and they may look up to check whether their child is attending. In addition to merely checking attention, mothers' gaze likely elicits attention even from infants whose focus has drifted, so gazes before initiating and after completing actions may call attention to the action. This explanation aligns with Csibra and Gergely's [24] suggestion that eye gaze at the beginning of a demonstration is naturally elicited by any adult in any teaching event. Perhaps making eye contact at key points throughout the demonstration is a similarly spontaneous aspect of teaching behavior that would emerge for any learner. However, based on prior research showing that adults make substantially more eye contact when teaching infants than other adults [7], looking throughout the demonstration likely reflects some sensitivity to the specific attention abilities of the learner. This also accords with research showing that even a robot garners similar gazing behavior in a teaching scenario as long as it displays infant-like attention characteristics [52],[53].

Another, more mechanistic, explanation is that mothers look at their children as much as possible for purely social (rather than pedagogical) reasons, but must turn their attention to the object itself when they are manipulating it. Note that this mechanism would provide the same result: that mothers' looks to their infants tend to happen at the boundary points of the action. Whether the timing of mothers' gaze is intended to communicate information about action or not, it nevertheless does communicate such information. Even assuming no inherent teaching motivation or sensitivity to the needs of the partner, such a pattern would nevertheless be available for infants to provide clues to parsing the action. However, because actions and gazes are particularly tightly aligned for the final actions in an enabling sequence, even relative to other actions on that object, a purely mechanistic explanation seems unlikely. This suggests that it is not the objects per se (and inherent difficulty manipulating them) which elicited differences in eye gaze alignment, but the special status of the action step which achieves a goal.

Another possibility is that mothers' gaze in the current study is elicited by some behavior on the part of the infant, rather than being initiated by mothers themselves. Given the findings of differential timing of gaze, this would indicate that it is the 7- to 12-month-old infants who already discriminate not only enabling from arbitrary sequences, but also the goal action from the enabling means actions. As the infants have no prior familiarity with these toys, this seems unlikely; however, only future research which measures both infant and maternal behaviors simultaneously can answer this question definitively.

If the timing of infant-directed gaze with respect to action boundaries stems at least in part from maternal sensitivity to infant attention and developmental stage, we might expect variability in this input. Specifically, research with mothers of children with autism, or children who are blind, would be informative about the degree to which mothers are responding to features of the infant. Also, mothers who are lower in sensitivity or mind-mindedness [54], or those with depressive disorders [55], might fail to provide the tight alignment seen here in this community sample. Future research should explore whether such variability exists and whether any such variability impacts infants' attention or parsing.

In sum, when demonstrating novel objects to their infants, mothers systematically aligned their infant-directed gaze onsets with their actions. Mothers were particularly apt to align gaze with the boundary points of actions which represented the culmination of an enabling, goal-directed sequence. The typical pattern involved mothers looking up just before beginning and just after ending an action, although for the final actions of an enabling sequence, mothers were likely to look up just after beginning as well as just after completing the action. Thus, gaze may function as a probabilistic marker of action boundaries or as an attentional spotlight on other markers which occur at boundary points. Identification of these gaze-action patterns opens the door for future research, in which we can explore whether such patterns support learning in infants or artificial systems.

ACKNOWLEDGMENTS

The authors would like to thank the Reed College Department of Psychology; Emily Blumenthal, Meredith Meyer, and three anonymous reviewers for their insightful feedback; Bridgette Martin Hard for advice on alignment calculation; and the Cognitive Development Project members at Reed College and Villanova University for their assistance. Portions of these data were presented as the Reed College Honor's thesis of the second author, and at the Society for Research in Child Development, 2010.

References

[1] N. Z. Kirkham, J. A. Slemmer, and S. P. Johnson, "Visual statistical learning in infancy: Evidence of a domain general learning mechanism," *Cognition*, vol. 83, no. 2, pp. B35-B42, Mar. 2002.

[2] J. Saffran, R. Aslin, and E. Newport, "Statistical learning by 8-month-old infants," *Sci*., vol. *274,* no. 5294, pp.1926-1928, Dec. 1996.

[3] D. Buchsbaum, A. Gopnik, T. L. Griffiths, and P. Shafto, "Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence" *Cognition*, vol. 120, no. 3, pp. 331-340, Sep. 2011.

[4] A. N. Meltzoff, "Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children," *Dev. Psych*. vol. 31, no. 5, pp. 838-850. Sep. 1995.

[5] R. Brand, D. Baldwin, and L. Ashburn, "Evidence for 'motionese': modifications in mothers' infant-directed action," *Dev. Sci*., vol. *5 no. 1,* pp. 72-83, Mar. 2002.

[6] G. Gergely and G. Csibra, "The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy," *Interaction Stud.*, vol. 6, no. 3, pp. 463-481, 2005.

[7] R. Brand, W. Shallcross, M. Sabatos, and K. Massie, "Fine-grained analysis of motionese: Eye gaze, object exchanges and action units in infant-versus adult- directed action," *Infancy*, vol. 11, no. 2, pp. 203-214, Mar. 2007.

[8] G. Gergely, K. Egyed, and I. Király, "On pedagogy," *Dev. Sci*., vol. 10, no. 1, pp. 139-146, Jan. 2007.

[9] S. M. J. Hains and D. W. Muir, "Infant sensitivity to adult eye direction," *Child Dev*., vol. 67, no. 5, pp. 1940–1951, Oct. 1996.

[10] A. Senju, G. Csibra, and M. H. Johnson, "Understanding the referential nature of looking: Infants' preference for object-directed gaze," *Cognition*. vol. 108, no. 2, pp.303-319, Aug. 2008.

[11] S. Roseberry, R. Richie, K. Hirsh-Pasek, R. Golinkoff, and T. Shipley, "Babies Catch a Break: 7-to 9-Month-Olds Track Statistical Probabilities in Continuous Dynamic Events." *Psych. Sci.*, vol.22, no. 11 pp. 1422-1424, Nov. 2011.

[12] D. Baldwin and J. Baird, "Action analysis: A gateway to intentional inference," in *Early social cognition*, P. Rochat, Ed. Hillsdale: Lawrence Erlbaum Associates, 1999, pp. 215-240.

[13] B. M. Hard, B. Tversky, and D. Lang, "Making sense of abstract events: Building event schemas," *Memory and Cognition*. vol 34, no. 6, pp. 1221-1235, Sep. 2006.

[14] J. M. Zacks, "Using movement and intentions to understand simple events," *Cognitive Sci.*, vol. 28, no. 6, pp. 979-1008, Nov. 2004.

[15] L. E. Bahrick, R. Lickliter, and R. Flom, "Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy," *Cur. Dir. in Psych. Sci.*, vol. 13, no. 3, pp. 99-102, Jun. 2004.

[16] D. A. Baldwin, J. A. Baird, M. M. Saylor, and M. A. Clark, "Infants parse dynamic action," *Child Dev*., vol. 72, no. 3, pp. 708–717, May 2001.

[17] M. M. Saylor, D. A. Baldwin, J. A. Baird, and J. LaBounty, "Infants' on-line segmentation of dynamic human action," *J. of Cog. and Dev*., vol. 8, no. 1, pp. 113-128, Dec. 2007.

[18] G. Csibra, G, Gergely, S. Bíró, O. Koos, and M. Brockbank, "Goal attribution without agency cues: the perception of 'pure reason' in infancy." *Cognition*, vol. 72, no. 3, pp. 237-267, Oct.1999.

[19] L. Lakusta, L. Wagner, K. O'Hearn, and B. Landau, "Conceptual foundations of spatial language: Evidence for a goal bias in infants," *Lang. Learning and Dev*., vol. 3, no.3, pp. 179-197, Dec. 2007.

[20] A. Fernald and T. Simon, "Expanded intonation contours in mothers' speech to newborns," *Dev.Psych.*, vol. 20, no. 1, p. 104, 1984.

[21] A. Fernald, T. Taeschner, J. Dunn, M. Papousek, B. Boysson-Bardies, and I. Fukui, "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants," *J. of Child Language*. vol. 16, no. 3, pp. 477-501, Oct. 1989.

[22] J. M. Iverson, O. Capirci, E. Longobardi, and M. C. Caselli, "Gesturing in mother-child interactions," *Cog. Dev.*, vol. 14, no. 1, pp. 57-75, Jan. 1999.

[23] S. Chong, J. F. Werker, J. A. Russell, and J. M. Carroll, "Three facial expressions mothers direct to their infants," *Infant and Child Dev*., vol. 12, no. 3, pp. 211–232. Mar. 2003.

[24] L. J. Gogate, L. E. Bahrick, and J. D. Watson, "A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures," *Child Dev*., vol. 71, no. 4, pp. 878-894, Jul. 2000.

[25] M. Meyer, B. Hard, R. Brand, M. McGarvey, and D. A. Baldwin, "Acoustic packaging: Maternal speech and action synchrony," *IEEE Trans. Autonomous Mental Dev.*, vol. 3, no. 2, pp. 154-162, Jun. 2011.

[26] L. Schillingmann, B. Wrede, and K. Rohlfing, "A computational model of acoustic packaging." Autonomous Mental Development, IEEE Transactions on 1.4 (2009): 226-237.

[27] M. Rolf, M. Matthias, M. Hanheide, and K. J. Rohlfing. "Attention via synchrony: Making use of multimodal cues in social learning." *IEEE Trans. Autonomous Mental Dev. vol. 1, no. 1*, pp. 55-67, 2009.

[28] G. Csibra and G. Gergely, "Natural pedagogy as evolutionary adaptation," *Phil. Trans. of the Royal Soc. B*, vol. 366, no. 1567, pp. 1149-1157, Apr. 2011.

[29] R. J. Brand and W. Shallcross, "Infants prefer motionese to adult-directed action," *Dev. Sci.*, vol. 11, no. 6, pp. 853-861, Nov. 2008.

[30] A. Senju and G. Csibra, "Gaze following in human infants depends on communicative signals," *Current Biology*, vol. 18, no. 9, pp.668-671, May 2008.

[31] R. Wu, A. Gopnik, D. C. Richardson, and N. K. Kirkham, "Infants learn about objects from statistics and people," *Dev. Psych*, vol. 47, no. 5, pp. 1220-1229, Sep. 2011.

[32] K. J. Rohlfing, J. B. Fritsch, Wrede, and T. Jungmann, "How can multimodal cues from child-directed interaction reduce learning complexity in robots?" *Advanced Robotics*, vol. 20, no. 10, pp. 1183–1199, 2006.

[33] Y. Nagai and K. J. Rohlfing. "Computational analysis of motionese toward scaffolding robot action learning." *IEEE Trans. Autonomous Mental Dev*., vol. 1, no. 1, pp. 44-54, 2009.

[34] R. Williamson and R. J. Brand, "Infant-directed action promotes 2-year-olds' imitation," in press, *Journal of Experimental Child Psych*.

[35] E. A. Koterba and J. M. Iverson, "Investigating motionese: The effect of infant-directed action on infants' attention and object exploration," *Infant Beh. and Dev.*, vol.32, pp. 437–444, Nov. 2009.

[36] R. J. Brand, A. McGee, J. F. Kominsky, K. Briggs, A. Gruneisen, and T. Orbach, "Repetition in infant-directed action depends on the goal structure of the object: Evidence for statistical regularities," *Gesture*, vol. 9, no. 3, pp. 337-353, Dec. 2009.

[37] M. Tomasello, " Human culture in evolutionary perspective," in *Advances in Culture and Psych*, M. Gelfand, Ed. New York: Oxford University Press, 2011, pp.5-51.

[38] T. Farroni, G. Csibra, F. Simion, and M. Johnson, "Eye contact detection in humans from birth," *Proc. of the Academy of Natural Sci.*, vol. 99, no. 14, pp. 9602-9605, Jul. 2002.

[39] D. A. Baldwin, "Infants' ability to consult the speaker for clues to word reference," *J. of Child Lang.*, vol. 20, no. 2, pp. 395-418, Jun. 1993.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TAMD.2013.2273057

[40] R. Brooks and A. N. Meltzoff, "The importance of eyes: How infants interpret adult looking behavior," *Dev. Psych*. vol. 38, no. 6, pp. 958-966, Nov. 2001.

[41] A. L. Woodward, "Infants' Developing understanding of the link between looker and object," *Dev. Sci*.,. *v*ol. 6, no. 3, pp. 297-311, Jun. 2003.

[42] J. S. Beier and E. S. Spelke, "Infants' Developinging understanding of social gaze," *Child Dev.*, vol. 83, no. 2, pp. 486-496, Jan. 2012.

[43] A. Belardinelli, F. Pirri and A. Carbone, "Bottom-up gaze shifts and fixations learning by imitation." *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 37, no. 2, pp. 256-271, 2007.

[44] Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. "Social cognition, joint attention, and communicative competence from 9 to 15 months of age." Monographs of the society for research in child development, i-174, 1998.

[45] P. J. Bauer, "Holding it all together: How enabling relations facilitate young children's event recall," *Cog. Dev*., vol. 7, no. 1, pp. 1-28, Jan. 1992.

[46] J. M. Zacks, B. Tversky, and G. Iyer, "Perceiving, remembering, and communicating structure in events," *J. of Experimental Psych: General*, vol. 130, no. 1, pp. 29-58, Mar. 2001.

[47] C. Trevarthen and P. Hubley, "Secondary intersubjectivity: Confidence, confiding, and acts of meaning in the first year," in *Action, Gesture, Symbol: The emergence of language*, A. Lock, Ed London: Academic Press, pp. 183-229, 1978.

[48] H. A. Ruff and M. K. Rothbart. "Attention in Early Development: Themes and Variations." Oxford University Press, 2001.

[49] G. Hollich. (2003). SuperCoder (Version 1.5). [Online]. Available: http://hincapie.psych.purdue.edu/Splitscreen/SuperCoder.dmg

[50] P. J. Bauer, L. A. Hertsgaard, P. Dropik, and B. P. Daly, "When even arbitrary order becomes important: Developments in reliable temporal sequencing of arbitrarily ordered events," *Memory*, vol. 6, no. 2, pp. 165-198, 1998.

[51] A. Fernald, "Four-month-old infants prefer to listen to motherese," *Infant Behavior and Dev*., vol. 8, no. 2, pp. 181-195, Apr. 1985.

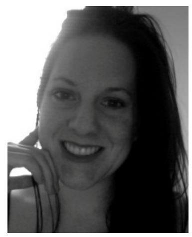
[52] A. Vollmer, K. S. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, K. J. Rohlfing, and B. Wrede, "People Modify Their Tutoring Behavior in Robot-Directed Interaction for Action Learning," in Proc. 8th IEEE Int. Con. on Dev. and Learning, Shanghai, China, 2009, pp. 1-6.

[53] Y. Nagai, A. Nakatani, and M. Asada, "How a robot's attention shapes the way people teach," in Proc. of the 10th Int. Con. on Epigenetic Robotics, Örenäs Slott, Sweden, 2010, pp. 81-88.

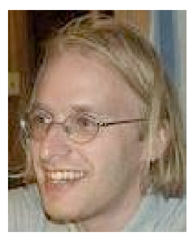
[54] E. Meins and C. Fernyhough, "Linguistic acquisitional style and mentalizing development: the role of maternal mind-mindedness," *Cog. Dev.*, vol. 14, no. 3, pp. 363-380, Sep. 1999.

[55] E.M. Cummings and P.T. Davies, "Maternal depression and child development," *J. of Child Psych. & Psych.*, vol. 35, no. 1, pp. 73-122, Jan. 1994.

Rebecca J. Brand received a B.A. (1996) degree in cognitive science from Vassar College, Poughkeepsie, NY, and an M.S. (1998) and Ph.D. (2002) degrees in psychology from the University of Oregon, Eugene. She is currently an Associate Professor of Psychology at Villanova University, Villanova, PA. Her research focuses on infants' social–cognitive skills and the role of parental input in the development of these skills. Her current work investigates infant-directed action (motionese), as well as speech-action alignment (acoustic packaging).



Emily Hollenbeck graduated from Reed College in 2008 with a B.A. in psychology. She is currently a Ph.D. student in cognitive psychology at Northwestern University. Her research interests include infant social cognition and infants' understanding of basic physical principles. Her current work explores infants' attention to causal information and infants' expectations about novel substances.



Jonathan F. Kominsky received his B.A. in Psychology from Reed College in 2009. He is currently working on a Ph.D. in Psychology at Yale University. His research focuses on how children and adults perceive and understand the multitudes of causal systems they encounter in day-to-day life.