Linking Joint Attention with Hand-Eye Coordination – A Sensorimotor Approach to Understanding Child-Parent Social Interaction

Chen Yu and Linda B. Smith (chenyu, smith4@indiana.edu)
Department of Psychological and Brain Sciences, and Cognitive Science Program, Indiana University
Bloomington, IN, 47405 USA

Abstract

An understanding of human collaboration requires a level of analysis that concentrates on sensorimotor behaviors in which the behaviors of social partners continually adjust to and influence each other. A suite of individual differences in partners' ability to both read the social cues of others and to send effective behavioral cues to others create dyad differences in joint attention and joint action. The present paper shows that infant and dyad differences in hand-eye coordination predict dyad differences in joint attention. In the study reported here, 51 toddlers and their parents wore head-mounted eye-trackers as they played together with objects. This method allowed us to track the gaze direction of each participant to determine when they attended to the same object. We found that physically active toddlers align their looking behavior with their parent, and achieve a high proportion of time spent jointly attending to the same object in toy play. However, joint attention bouts in toy play don't depend on gaze following but rather on the coordination of gaze with hand actions on objects. Both infants and parents attend to their partner's object manipulations and in so doing fixate the object visually attended by their partner. Thus, the present results provide evidence for another pathway to joint attention – hand following instead of gaze following. Moreover, dyad differences in joint attention are associated with dyad differences in hand following, and specifically parents' and infants' manual activities on objects and the within- and between-partner coordination of hands and eyes during parent-infant interactions. In particular, infants' manual actions on objects play a critical role in organizing parent-infant joint attention to an object.

Keywords: joint attention, eye tracking, perception and action

Introduction

Everyday human collaborative behavior, from maintaining a conversation to jointly solving a physical problem, seems so effortless that we often notice it only when it goes awry (Shockley, Richardson, & Dale, 2009). Many contemporary theories of social interaction, collaboration, and joint attention concentrate on internal representations and inferences from those representations For example, one common psychological explanation of how we manage to (typically) work so well together is called "mind-reading" (Baron-Cohen, 1997). The idea is that we form models of and make inferences about the internal states of others; for example, along the lines of "He is looking at the object and so must want me to pick that up." However, it is not at all clear that such mental models about the states of others – and inferences from such internal representations – can explain the real-time smooth fluidity of such collaborative behaviors as everyday conversation or joint action depends on coordinated adjustments on the time scales of fractions of seconds. Accordingly, there is growing interest (de Barbaro, Johnson, Forster, & Deak, 2013; Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012; Pereira, Smith, & Yu, 2008; Richardson, Dale, & Tomlinson, 2009; Shockley, Santana, & Fowler, 2003) in more micro-analytic studies of just what happens – in real time – as individual agents interact. These new approaches concentrate on the real-time dynamics of the behaviors of collaborating social partners – rapid shifts of eye movements, head turns, and hand gestures – and how they co-organize across partners in an interaction. Within these newer approaches, an understanding of human collaboration requires a level of analysis that concentrates on sensorimotor behaviors in a complex dynamical system in which the behaviors of social partners continually adjust to and influence each other.

Here we focus on momentary looking behavior in coordination with hand actions. Eye gaze is a micro-behavior tightly tied to one's internal attentional state (Baron-Cohen, 1997; Frischen, Bayliss, & Tipper, 2007) and used by partners in social interactions to establish common ground and smooth social engagements (Argyle, 2007). The importance of gaze following to developmental process is well documented: individual differences in infants' and children's ability to coordinate visual attention with a social partner is correlated with individual differences in language, social, and cognitive development (Brooks & Meltzoff, 2005; Mundy & Gomes, 1998). Considerable research also suggests that both adults and infants are highly attentive to instrumental hand actions of their social partners (Land & Hayhoe, 2001). We know less about how attention to hands may interact with gaze following.

The present study tests the hypothesis that infant's success in visually attending to the same object with their parent depends on hand-eye coordination *within* the infant and *between* the infant and parent. The hypothesis was suggested by the results from a prior study of how parents and 12-month old infants coordinate visual attention (Yu & Smith, 2013). In that study, parents and infants played with multiple toys at a time while both of them wore head-mounted eye tracking systems that measured the momentary gaze of each partner and provided a precise measure of the coordination of visual attention. Consistent with findings from a growing number of studies (Aslin, 2009; Deák, Krasno, Triesch, Lewis, & Sepeta, 2014; Franchak, Kretch, Soska, & Adolph, 2011; Frank, Vul, & Saxe, 2012; Yoshida & Smith, 2008), the fixation data indicated that in this active context, infants rarely looked to their parent's face, a fact that precludes gaze following by the infant as a contributing factor to joint attention. Further, the dynamics of parent and infant attention to objects during the interaction were very different: parents rapidly shifted eye gaze among all the visual targets while infants generated longer and stickier looks on objects (Smith, Yu, & Pereira, 2011; Yu & Smith, 2012). Despite these differences in gaze patterns, parents' and infants' visual attention was often coordinated such that they were fixated on the same object at the same time. Moreover, the two partners appeared to shift attention to an object nearly in unison suggesting between-partner coordination at the sensorimotor level. The gaze pattern from both partners indicated that joint attention moments were often formed by one partner handling an object and the other looking to those hand actions.

In the present paper, we used a method similar to that in the previous dual eye-tracking study of 12 month olds and their parents (Yu &Smith, 2013). However, here we collected a larger sample; the infant participants ranged in age from 11 to 24 months in an effort to capture individual differences in object manipulation and socially-coordinated visual attention. We focused on the second year of life because past research suggests that during this age period, infants become increasingly active and autonomous and individual differences in motor behavior and joint attention become noticeable (Landa, Gross, Stuart, & Faherty, 2013).

Method

Participants. The final sample consisted of 51 (24 male infants) parent-infant dyads with the infants ranging in age from 11 to 24 months (mean = 17.92, SD= 4.15); 14 additional dyads began the study but the infants refused to wear the measuring equipment. The children were recruited from a population of working and middle class families in a Midwestern town.

Stimuli. There were 6 unique novel "toys" constructed in the laboratory and pilot-tested to be interesting and engaging to infants. Each novel toy was a complex object made from multiple and often moveable parts and were of similar size, on average, 288 cm^3. These were organized into two sets of three so that each object in the set had a unique uniform color.

Experimental setup. Parents and toddlers sat across from each other at a small table (61cm × 91cm × 64cm). Parents sat on the floor such that their eyes and heads were at approximately the same distance from the tabletop as those of the toddlers, a posture that parents reported to be natural and comfortable. Both participants wore head-mounted eye trackers (positive science, LLC; also see Franchak et al., 2011). Each eye-tracking system includes an infrared camera – mounted on the head and pointed to the right eye of the participant – that records eye images, and a scene camera that captures the events from the participant's perspective. The scene camera's visual field is 90 degrees, providing a broad view but one less than the full visual field --approximately 170^{o} (Smith, Yu, Yoshida, & Fausey, 2014). Each eye tracking system recorded both the egocentric-view video and eye-in-head position (x and y) in the captured scene at a sampling rate of 30 Hz. Another high-resolution camera (recording rate 30 frames per sec) was mounted above the table and provided a bird's eye view that was independent of participants' movements.

Procedure. One experimenter played with the infant while another experimenter placed the eye-tracking gear low on the forehead at a moment when the infant was engaged with the toy. To collect calibration points for eye tracking, the experimenter directed the infant's attention toward an attractive toy. This procedure was repeated 15 times with the toy placed in various locations on the tabletop to ensure a sufficient number of calibration points. To calibrate the parent's eye tracker, the experimenter asked the parent to look at one of the objects on the table, placed close to the toddler, and then repeated the same procedure to obtain at least 15 calibration points from the parent. Parents were told that the goal of the experiment was to study how parents and toddlers interacted with objects during play and therefore they were asked to engage their toddlers with the toys and to do so as naturally as possible. Each of the two sets of toys was played with twice for 1.5 min, resulting in 6 minutes of play data from each dyad. Order of sets (ABAB or BABA) was counterbalanced across dyads.

Data processing. Four regions-of-interest (ROIs) were defined: the three toy objects and the partner's face. These ROIs were coded manually by a coder who watched the first-person view video with a cross-hair indicating gaze direction, frame-by-frame, and annotated when the cross-hairs overlapped any portion of the four ROIs. Thus, each dyad provided two gaze data streams containing four ROIs (3 objects, partner's face) as shown in Figure 1. The second coder independently coded a randomly selected 10% of the frames with 95% agreement. Also shown in Figure 1 is a measure of sustained joint attention that will be described in detail in the following section.

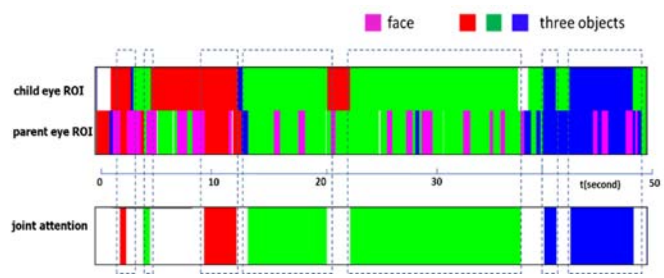

Figure 1. An overview of raw and derived data. Top: two gaze data streams from child and parent include four regions of interest – three toy objects and the partner's face. Bottom: Sustained visual attention (highlighted by dotted lines) is derived based on child's and parent's gaze data.

Results

Individual gaze patterns and joint attention

Figure 1 shows a representative example of the raw gaze data streams for one dyad and Table 1 provides the summary statistics of several measures of infant and parent fixations for the entire sample, including overall looks, looks to faces and looks to objects. For each type of looking behaviors, we report three measures: 1) percentage of total looking time to ROIs, 2) frequency with which these looks were formed (in rate/min), and 3) mean duration of looks (in sec). For all measures, correlations with age were small and not significant, with one exception (proportion of time infants looked at the ROIs versus “off-task”), the lack of an overall age effects is consistent with a period of rapid developmental change in which individuals progress at different rates and thus a period of marked individual differences (see, Smith, Yu & Pereira, 2011; Yu & Smith, 2012, 2013). However infants and their parents differed considerably and reliably on all measures: Infants and parents spent a high proportion of time fixating the ROIs but parents spent more total time overall than infants (M_{parent}=82.58%, M_{infant}=75.76%) and exhibited more attentional switches between objects and faces (M_{parent}=61.29 switches per min, switches per minute) than infants (M_{infant}=25.46). Infants, in contrast, had longer unbroken fixations on the ROIs than did the parents (M_{parent}=806ms, M_{infant}=1825ms), showing the “stickiness” often observed in infant and toddler attention during object play (Kannass, Oakes, & Shaddy, 2006; Yu & Smith, 2013). Finally and consistent with past findings (Yu & Smith, 2013), infants rarely looked to their parents’ face during the play session (M_{parent}=34.03%, M_{infant}=11.61%).

We used the method developed by Yu & Smith (2013) to find joint attention episodes. We first determined –frame by frame – the frames in which parents and infants fixated on the same ROI. Meaningful shared attention should last some amount of time longer than a frame (33msec) but might also include very brief looks elsewhere. Therefore, a joint attention (JA) bout was defined as a continuous alignment of parent’s and toddler’s fixation to the same ROI that lasted longer than 500 msec and included segments of fixations that were to the same object but separated by brief looks away by one partner of no longer than 300 msec. Examples of the so-defined joint attention bouts from one dyad’s gaze streams are shown in Figure 1.

Table 1. Parent and infant differences in fixations on the defined ROIs: the three objects and each other’s face.

	Infant mean SD	correlation with age	Parent mean SD	correlation with age	infant-parent comparison
Fixations to ROIs					
% of time	75.76	0.256*	82.58	0.016	$t(50)=2.77$
	-12.65		-12.23		$p<.001$
frequency (rate/min)	25.46	0.145	61.29	0.032	$t(50)=17.366$
	-8.33		-13.71		$p<.001$
duration (msec)	1825	0.103	806	0.021	$t(50)=8.79$
	-745		-282		$p<.001$
Looks to faces					
% of time	11.61	-0.04	34.03	-0.203	$t(50)=8.038$
	-7.14		-13.87		$p<.001$
frequency (rate/min)	4.74	0.099	21.52	-0.155	$t(50)=15.805$
	-2.28		-7.22		$p<.001$
duration (msec)	1364	0.085	791	-0.01	$t(50)=6.514$
	-530		-338		$p<.001$
Looks to objects					
% of time	64.8	-0.127	55.33	-0.235	$t(50)=5.573$
	-13.71		-11.63		$p<.001$
frequency (rate/min)	20.71	0.123	39.77	0.145	$t(50)=10.13$
	-7.96		-10.8		$p<.001$
duration (msec)	2093	0.139	826	0.049	$t(50)=9.86$
	-885		-243		$p<.001$

Column 1 of Table 2 summarizes a set of statistics on joint attention measures across the whole sample: the percentage of overall time in joint attention to any ROI, the frequency with which joint attention bouts were formed (in rate/min), and the mean duration of these bouts (in sec). These same statistics are provided for the two subcomponents of overall joint attention – mutual gaze and joint attention to an object. Overall, parents and toddlers looked at the same object at the same time over 34.72% of the play session; however, there were substantial dyad differences -- joint attention episodes varied from near 14% to over 63% of the play session.

As a first indicator of the potential importance of hand actions, across dyads, 82.34% of joint attention moments on a visual object, the target object was manually contacted by at least one partner. In contrast, the percentage time in which an object was being manually contacted by one partner in non-joint attention moments was only 43.67% ($t(100)=29.40$, $p<0.001$, $d=6.07$). Moreover, infants were holding the jointly attended object 45.23% of time (SD=5.63%) and parents were holding the jointly attended object 37.72% of time (SD=5.81%). In brief, the target object was always in one of the social partners’ hands at JA moments; and during those moments, infants held it more than parents did ($t(100)=5.85$, $p<0.001$, $d=1.17$).

Low and High JA groups

We partitioned the dyads into those with high and low incidence of JA bouts using a median split of the overall percentage of joint attention time. Columns 3 and 4 of Table 2 provide the statistics for the two defined groups for the measures of percentage of JA time, frequency of JA bouts and duration of JA bouts. Because the two groups were defined by the overall percentage of time in joint attention, the expectation is that they would differ on all the components contributing to this overall measure. As shown in column 5 of Table 2, this is generally true with the exception of measures of mutual gaze, a low frequency behavior in the present study, and one that at least in the context of active toy play may not be linked to the likelihood of joint attention (see also Yu & Smith, 2013). High and Low JA dyads also did not differ in the frequency with which parents looked to infant faces, M_{high}=22.77, M_{low}=20.23, $t(49) < 1.00$, nor in the frequency with which infants looked to parent faces, M_{high}=4.56, M_{low}=4.94, $t(49) < 1.00$. The High JA infants were older than the Low JA infants but the difference was not reliable; the mean age of the High JA infants was 18.78 months (SD = 4.24) and the mean of the Low JA infants was 19.8 months (SD=3.91), $t(49) = 1.61$, p=.113. Thus, neither age nor gaze following seems to be a determining factor of individual differences in joint attention.

The next set of analyses speak directly to the hypothesis that dyad differences in joint attention are related to within- and between-partner hand-eye coordination. By hypothesis, object manipulation matters in establishing joint attention because partners look to their own hand actions on objects and to the hand actions of their partner. To test this hypothesis, we measured the proportion of total play time that gaze was fixated on a hand-held object. Hand-eye coordination was measured by combining gaze and hand streams, frame by frame, to determine whether the visually attended object at a moment was manually handled by a partner. With gaze and hand streams from both partners, this yielded four measures: infant hand-eye coordination, parent hand-eye coordination, infant eye to parent hand, and parent eye to infant hand. Figure 2 shows the main results.

Consider first the case in Figure 2(a) which coincides with a hand-following path to joint attention: the infant was handling an object and the parent was not touching any object. High and Low JA infants differed in the degree to which they looked at the own manual actions on objects, with high JA infants looking at their own object handling more than Low JA infants. High and Low JA parents also differed, with High JA parents being more attentive to the hand actions of their infants than Low JA parents. These conclusions were confirmed via a 2 (JA group) X 2 (Participant) analysis of looking behavior for the cases when the (only) infant was handling an object. The analysis revealed only two main effects – Low versus High JA, $F(1,98) = 12.11$, $p<0.001$, $\eta_p^2 = 0.08$, and Parent versus Infant, $F(1,98)=18.37$, $p<0.001$, $\eta_p^2 = 0.16$. Across both groups, infants looked at their own hand actions more than did their parents, but High JA infants and High JA parents looked more at the object handled by infant self, than did Low JA infants and Low JA parents. These findings provide support for the main hypotheses from the infant side of manual actions: Infants who are more likely to achieve joint attention bouts with their parents not only manually act on objects more, but they also look more to their own hand actions. Moreover, they have parents who visually follow the hand actions from those infants to an object more than the parents of Low JA infants.

Figure 2(b) reports the patterns for another hand-following path to joint attention when the parents were in manual contact with an object (and the infants were not). A 2 (high vs. low JA group) x 2 (parent gaze vs. child gaze) analysis of variance was conducted with no significant main effects or interactions found ($F_{group}(1,98)=3.56, p=0.06$; $F_{agent}(1,98) = 0.17, p=0.67$, n.s.; $F_{interaction}(1,98)=0.11, p=0.73$, n.s.). This indicates that both parents' and toddlers' attention to what the parent held was not related to whether they had more or fewer joint attention bouts. Overall, the results from the first two cases – infant holding and parent holding -- suggest that dyad differences derive from differences in infant hand actions on objects and from parent visual responses to those hand actions from infants. That is, the

Table 2. Measures of Joint Attention Low and High Joint Attention (JA) Dyads.

	whole sample	correlation with age	Low JA	High JA	low-high comparison
overall					
% of time	39.24	0.295*	28.08	48.41	t(49)=10.61
	-12.07		-8.06	-4.85	p<.001
frequency (rate/min)	9.44	0.246	7.82	10.76	t(49) =6.45
	-2.19		-1.55	-1.7	p < .001
duration (msec)	2.38	0.199	2.02	2.68	t(49) = 5.49
	-0.53		-0.43	-0.44	p < .001
mutual gaze					
% of time	5.74	0.093	4.27	4.92	t(49) = 0.72
	-4.56		-4.05	-6.62	n.s.
frequency (rate/min)	1.78	0.06	1.37	1.81	t(49)= 0.93
	-1.28		-0.96	-1.43	n.s.
duration (msec)	2.44	-0.02	1.54	1.65	t(49) = 0.78
	-0.68		-0.69	-0.73	n.s.
JA to object					
% of time	34.72	-0.092	23.82	41.76	t(49) = 8.94
	-4.56		-7.68	-6.9	p<.001
frequency (rate/min)	7.66	0.212	6.45	8.65	t(49) = 5.09
	-1.65		-1.41	-1.07	p<.001
duration (msec)	1.68	0.42	1.99	2.82	t(49) =5.77
	-0.71		-0.46	-0.61	p<.001

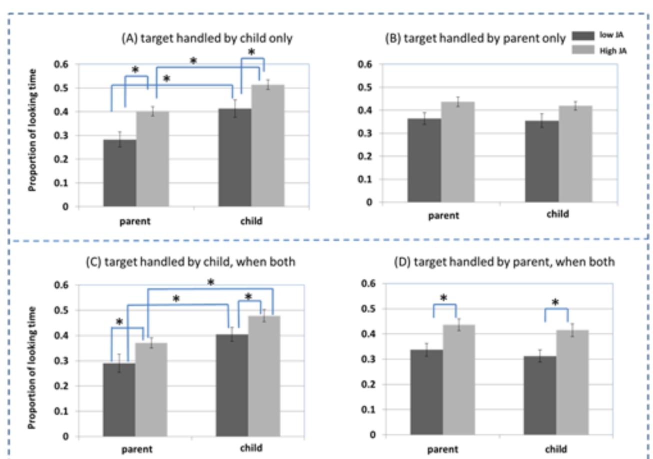

Figure 2. The proportion of total time child and parent visually fixate the target object, when the child is handling the target and the parent is not manually in contact with an object (A), when the parent is handling the target object the child is not handling any object (B), when the child is handling the target and the parent is handling another object (C); and when the parent is handling the target and the child is handling another object (D).

principal individual differences appear due to differences in the pathway through child holding and parent following.

Figure 2(c) and (d) show the findings from the more complicated case in which the infant and parent are each holding different objects: To which object do the partners' look? For the objects held by the infant, a 2 (JA group) x 2 (Participant – parent/infant) ANOVA indicated a main effect of JA group ($F(1,98) = 6.87$, $p<0.01$, $\eta_p^2 = 0.08$), and participant ($F(1,98)=15.74$, $p<0.001$, $\eta_p^2 = 0.14$), but no interaction ($F(1,98) = 0.09$, $p=0.75$, n.s.). High JA parents and infants paid more attention to the objects being handled by the infant than did Low JA parents and infants. The same analyses with respect to the object handled by the parent revealed only a significant effect of JA group ($F(1,98) = 10.48$, $p<0.005$,$\eta_p^2 = 0.11$) with parents and infants in the High JA dyads attending more to the objects handled by the parent than did Low JA infants and parents. There was no difference between parent and infant looking to parent hands, nor the interaction between the two main factors ($F(1,98) =1.13$, $p=0.29$,n.s.; $F_{interaction}(1,98)=0.51$, $p = 0.47$, n.s.). Thus, in the context when the two partners present multiple competing objects in play, High JA infants and parents manage to find a joint solution to selecting an object for attention and attend to the object held by either partner more than do Low JA infants and parents.

These results provide clear support for the hypothesized role of hand following in parent-infant joint attention and a source of individual differences in the development of joint attention in the second year of life. More specifically, High and Low JA dyads are distinguished by infant manual activity, infant attention to their own hand actions, parent attention to infant hand actions, and the joint resolution of competition when the two partners are holding different objects.

General Discussion

The period between 1 and 2 years of age is a time of rapidly emerging motor, attentional and social skills. Consistent with periods of significant developmental changes (Thelen & Smith, 1994), the observed dyad differences in joint attention were larger than age differences. Further, during this period of development and in the context of joint play with multiple toys, success in joint attention was differentially predicted by different behaviors for infants and parents. The infant behavior most strongly associated with dyad success in joint attention was coordination of eyes and hands when handling an object. The parent behavior most critical to joint attention was attention to the infant's hand actions. Further, parent attention to the infant's hands was strongly correlated with infant hand-eye coordination during object manipulations. This last fact suggests that parents whose infants were more coordinated attended to their infant's hand actions more than parents whose infants were less coordinated and thus whose hand actions were less predictive of the infant's looking behavior. The importance of the result is this: For parents to effectively follow their infant's attention, infants must send readable signals; for the hand-following path, that signal requires coordinated hands and eyes so that the easy-to-see hand location can reliably signal looking behavior. The finding that individual differences in joint attention were due principally to the path of infant acting and parent following rather than the path of parent acting and infant following aligns with a large literature on parental responsiveness and the positive effects of parents following on the infant's interests in several developmental domains (Bornstein & Tamis-LeMonda, 1989). Within this context, the present results also highlight the critical importance of the infant's developing sensorimotor system in providing readable cues that parents can follow.

A larger idea behind this research is that the sensorimotor coordination of parents and infants as they jointly interact with objects teaches infants how to rapidly read and respond appropriately to social signals, and how to use their own behavior to send signals to their parent. Because hand actions on objects provide precise and readily perceived cues as to the target of interest, hand actions –and attentional responses to hand actions – may play a critical role in training more precise gaze following (Ullman, Harari, & Dorfman, 2012). By hypothesis, parents who effectively scaffold joint attention with their infants during object play provide the kind of coherent context in which the relevant signals and behavioral responses to those signals are discovered. Thus, parent-infant dyads who for whatever reason have difficultly coordinating attention in object play may put the infant at risk for poorer developmental outcomes. If, as the present results imply, weaker hand-eye coordination on the part of the infant, limits parent ability to effectively scaffold joint attention, then poor hand-eye coordination could cascade into longer term consequences in social development and language learning.

The present proposal about the role of object manipulation and eye-hand coordination in joint attention are also relevant to the well-documented but not well-understood link between atypical sensorimotor development and atypical social and language development. More specifically, infants at risk for significant delays in social and language development have been reported to show atypical patterns of early sensorimotor development include delayed and unusual manual interactions with objects (Baranek, 1999; Koterba, Leezenbaum, & Iverson, 2014; Provost, Lopez, & Heimerl, 2007). If this current proposal is correct – that the infant's own object manipulation skills are a limiting factor in developing social interactions. Thus, the present findings suggest one route through which atypical sensorimotor development may impact social development and joint attention. In addition, the present findings offer new and testable hypotheses about how the development of socially coordinated attention is supported by – as well as supports – other developmental achievements.

Acknowledgments

We thank Seth Foster, Melissa Elston, Char Wozniak, Steven Elmlinger for collection of the data. This research was supported by National Institutes of Health R01 HD074601 and R21 EY017843.

References

Argyle, M. (2007). *Social interaction*: Aldine.

Aslin, R. N. (2009). How infants view natural scenes gathered from a head-mounted camera. *Optometry and Vision Science, 86*(6), 561-565.

Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9–12 months of age. *Journal of autism and developmental disorders, 29*(3), 213-224.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind* Cambridge, MA: The MIT Press.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *New Directions for Child and Adolescent Development*(43), 49-61.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language *Developmental Science, 8*(6), 535-543.

de Barbaro, K., Johnson, C. M., Forster, D., & Deak, G. O. (2013). Methodological Considerations For Investigating the Microdynamics of Social Interaction Development. *Autonomous Mental Development, IEEE Transactions on, 5*(3), 258-270. doi: 10.1109/TAMD.2013.2276611

Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science, 17*(2), 270-281. doi: 10.1111/desc.12122

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-Mounted Eye Tracking: A New Method to Describe Infant Looking. *Child development, 82*(6), 1738-1750.

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the Development of Social Attention Using Free-Viewing. *Infancy, 17*(4), 355-375.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin, 133*(4), 694-724.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in cognitive sciences, 16*(2), 114-121.

Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development, 7*(3), 381-409.

Koterba, E., Leezenbaum, N. B., & Iverson, J. M. (2014). Object exploration at 6 and 9 months in infants with and without risk for autism. *Autism, 18*(2), 97-105.

Land, M., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25-26), 3559-3565.

Landa, R. J., Gross, A. L., Stuart, E. A., & Faherty, A. (2013). Developmental trajectories in children with and without autism spectrum disorders: the first 3 years. *Child development, 84*(2), 429-442.

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant behavior and development, 21*(3), 469-482.

Pereira, A. F., Smith, L. B., & Yu, C. (2008). Social coordination in toddler's word learning: interacting systems of perception and action. *Connection Science, 20*(2-3), 73-89.

Provost, B., Lopez, B. R., & Heimerl, S. (2007). A comparison of motor delays in young children: Autism spectrum disorder, developmental delay, and developmental concerns. *Journal of autism and developmental disorders, 37*(2), 321-328.

Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science, 33*(8), 1468-1482.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science, 1*(2), 305-319.

Shockley, K., Santana, M., & Fowler, C. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326-332.

Smith, L., Yu, C., Yoshida, H., & Fausey, C. M. (2014). Contributions of Head-Mounted Cameras to Studying the Visual Environments of Infants and Young Children. *Journal of Cognition and Development*(just-accepted).

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science, 14*(1), 9-17.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action* (Vol. 10): MIT Press.

Ullman, S., Harari, D., & Dorfman, N. (2012). From simple innate biases to complex visual concepts. *Proceedings of the National Academy of Sciences, 109*(44), 18215-18220.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy, 13*(3), 229-248.

Yu, C., & Smith, L. B. (2012). Embodied Attention and Word Learning by Toddlers. *Cognition, 125*(2), 244-262.

Yu, C., & Smith, L. B. (2013). Joint Attention without Gaze Following: Human Infants and Their Parents Coordinate Visual Attention to Objects through Eye-Hand Coordination. *PLOS ONE, 8*(11).