Active Vision in the Perception of Actions: An Eye Tracking Study in Naturalistic Contexts

Ryan E. Peters (ryerpete@iu.edu)
Dian Zhi (dianzhi@iu.edu)
Matthew Petersen (matthew208@gmail.com)
Chen Yu (chenyu@indiana.edu)
Department of Psychological and Brain Sciences
Indiana University, Bloomington, IN 47405

Abstract

Infants' ability to attend actively and selectively to naturalistic stimuli is critical to early learning. Most studies on infant visual attention use screen-based paradigms wherein infants view stimuli on computer screens. Little is known about how infants observe others' activities in everyday contexts. Using head-mounted eye-tracking, this study examined how infants distributed attention when observing their parents perform an everyday task – making peanut-butter and jelly sandwiches – in a home-like environment. Infant observers attended to parents' activities less than adult observers in the same situation. However, when infants were engaged in action observation, their gaze patterns were distributed on task-relevant objects similarly to adult observers, suggesting they actively obtained rich visual input in this free-viewing situation. Moreover, infant-parent dyads coordinated visual attention during the food preparation task in similar ways as observed in other everyday tasks, such as toy play, suggesting sensorimotor processes play a critical role in coordinated attention.

Keywords: action observation; coordinated attention; eye-tracking; parent-child interaction; selective attention

Introduction

Infants' ability to attend actively and selectively to specific stimuli in the world is critical to early development and learning. This ability to select information from the environment is particularly important for human infants when they observe others' actions, because action observation is one of the primary ways for infants to learn about the world. Previous research has shown that infants are capable of learning the statistical regularities within action sequences and accurately predict future actions before they begin (Monroy, Gerson, & Hunnius, 2017). To obtain useful information through learning from observation, young learners need to look at the right place at the right time as a sequence of actions unfolds quickly in real time.

Most studies on action observation have focused on infants' early action understanding and their object knowledge by analyzing anticipatory gaze during action observation (Cannon & Woodward, 2012; Cannon, Woodward, & Gredebäck, 2012; Falck-Ytter, Gredebäck, & von Hofsten, 2006; Monroy, et al., 2017). We now have a great deal of information about action prediction from well-controlled laboratory paradigms designed to measure infants' eye movements when they watch action stimuli on a computer screen (for a review see Gredebäck & Falck-Ytter, 2015). However, these paradigms have fundamental differences from the ways in which infants observe actions in everyday contexts. First, no matter how realistically visual stimuli are created to closely approximate what we see in the real world, looking at a 2D computer screen with minimal head and body movements differs dramatically from most real-world everyday activities that infants engage in. Second, the experiments in the above studies are composed of discrete trials repeated over time, wherein some simple actions are performed on one or more objects within each trial (for example, bringing a phone to the ear or a cup to the mouth). In contrast, everyday activities usually involve a succession of actions, each different from the previous. Third, the goal of most action observation experiments is to reveal infants' expectations about action events and their growing knowledge about the behavior and goals of other people (Hunnius & Bekkering, 2014). However, in the real world, infants' looking behaviors don't primarily serve to signal existing knowledge, but instead support the acquisition of new knowledge through observing others' activities.

The goal of the present study is to examine how infants distribute their attention when they observe parents perform an everyday task – making peanut-butter jelly sandwiches (PBJ) – in a home-like environment. We chose this task for two reasons. Frist, food preparation is a routine task at home. Previous studies have shown the functional importance of such routine tasks for learning and memory, suggesting that those everyday tasks play a critical role in organizing children's cognitive and language development (Tamis-LeMonda, Custode, Kuchirko, Escobar, & Lo, 2019). Second, making PBJ sandwiches was used in a classic study on eye movements because the task itself involves a sequence of distinct actions, each of which demands visual attention to serve different functionalities (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003).

Given that very little of the infant research has been concerned with infant attention when observing ordinary activities in everyday settings, the present study aimed at addressing three fundamental questions on visual attention in naturalistic contexts. First, we ask whether and to what

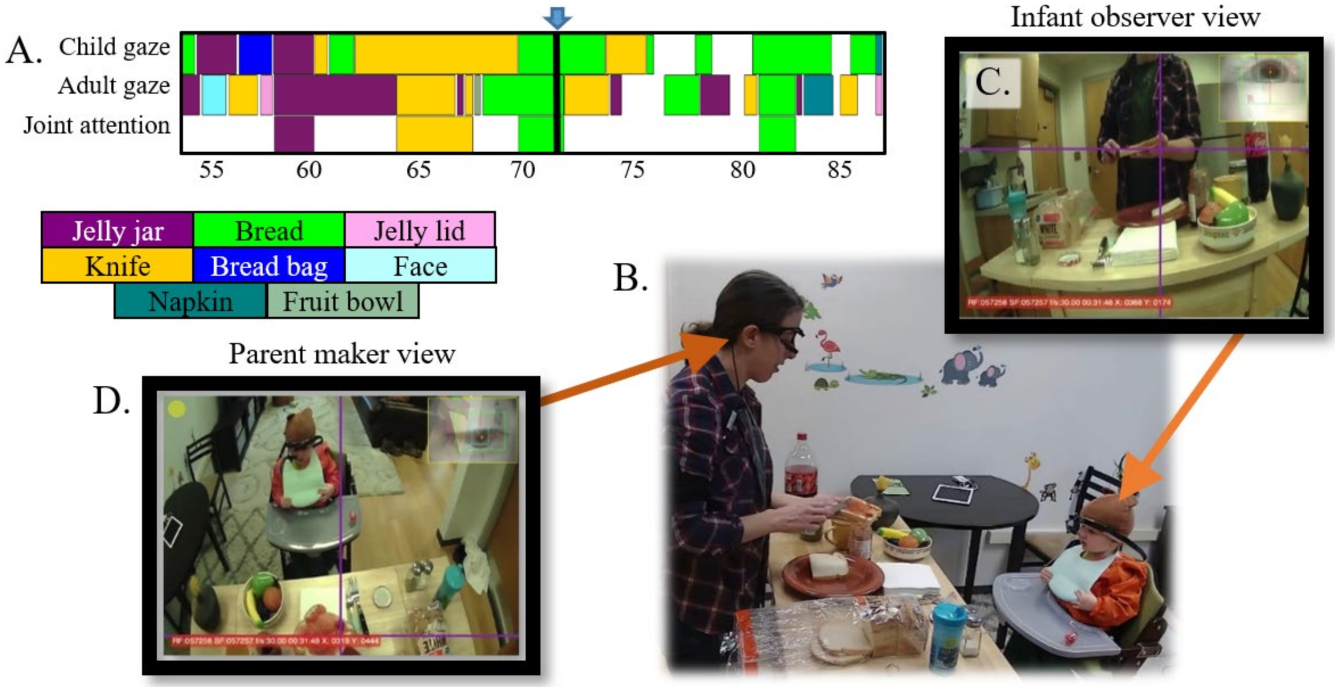

Figure 1. Data stream visualization (A), and third- (B) and first-person views for both infant observer (C) and parent maker (D) capturing a moment of coordinated attention on the bread during the jelly spreading subtask of the PBJ sandwich making task. Cross hairs on first-person views designate fixation locations.

degree infants pay attention to ongoing events when parents prepare food. Second, when infants do pay attention, are they attending to task-relevant objects and do they generate gaze patterns similar to adult observers in the same context? Finally, this food preparation task also allows us to ask questions about how well infants and their parents coordinate their visual attention in this everyday context. The ability to coordinate attention to objects of mutual interest is often regarded as an important developmental milestone.

Methods

Experimental setup

Parent-infant and adult-adult dyads were brought into a home-like laboratory environment to make PBJ sandwiches while wearing head-mounted eye-trackers. The lab is designed to resemble an apartment – with a living room, a play space, and a kitchenette – to allow for the capture of naturalistic behaviors in a controlled setting. The PBJ sandwich making task took place in the kitchenette, with a setup designed to resemble real-world situations in which infants might observe their parents making food (Figure 1B). Parents were asked to make sandwiches at a counter-height table while their infants sat across from them in a high-chair. For adult dyads, one adult was designated as the maker, while the other was the observer. Adult observers sat in a small chair selected so that eye level was roughly equivalent across infant and adult observers.

At the start of the experiment, the PBJ sandwich making supplies and a set of task-irrelevant distractor objects were arranged on the table as shown in Figure 2. Distractor objects were selected and positioned so as to be visually salient but naturalistic and unobtrusive – with the goal of measuring whether infants would be distracted by such task-irrelevant objects situated within the task space.

Figure 2. Initial setup of experimental objects for which gaze regions of interest were coded. Objects intended to be relevant to the PBJ sandwich making task in blue. Objects intended to be task-irrelevant in red.

Data collection

Five toddlers (mean age = 15.26 mos [12.7-19.3]; female = 3) and their parents, and nine adult dyads (recruited from undergraduate psychology classes) were recruited to participate in a study on interactions during naturalistic tasks. Data from parent-infant dyads is part of a larger ongoing project, while adult dyads were recruited as comparisons for this particular task, with a sample size equivalent to the original PBJ experiment by Hayhoe and colleagues (2003). Parents were asked to make four sandwiches, while adult makers were asked to make one (n=5) or two (n=4). Makers were given the option to give sandwiches they had made to the observers. Parents gave the first (n=4) or second (n=1) sandwich to their children, while no adults ate sandwiches during the experiment. Two trials from one parent-infant dyad were excluded from analyses due to poor eye-tracker calibration, resulting in a total of 18 sandwich making trials for parent-infant dyads and 13 for adult-adult dyads. The resulting dataset contained over 70,000 gaze data points (infant-parent dyads: 41,348, adult-adult dyads: 31,550.)

Both adult and infant participants wore head-mounted eye-trackers (Pupil Lab LLC) with infrared cameras aimed at their right eyes to record fixations (60Hz sampling rate) and scene cameras positioned to capture their first-person perspectives (Figure 1C & D). Infant eye-trackers were on hats, while adult eye-trackers were on a pair of glasses. Additional cameras were positioned in the room to capture third-person views. The eye-tracker calibration procedure followed a validated procedure to do head-mounted eye tracking with infants (e.g., Yu & Smith, 2017). Researchers watched the experiment from an adjoining room. If either of the eye cameras was moved during the experiment, the researchers reentered the room and adjusted the camera.

Gaze, subtask and object subtask-relevance coding

After the experiment, all eye-tracking and third-person videos were synchronized, and software was used to generate crosshairs on the parent and toddler first-person views estimating fixation locations. These videos were then used to manually code 21 regions of interest (ROI; 18 experimental objects and social partner's face, torso and empty hand) using an in-house program.

The videos were also used to divide each sandwich making trial into four subtasks: 1) bread preparation – starting with the first physical contact with an experimental object and going until the onset of the movement towards either the peanut butter or jelly jar, 2) peanut butter spreading – starting with onset of movement toward jar and going until either the onset of movement towards the jelly jar or, if following jelly spreading, the offset of the last spreading motion, 3) jelly spreading – starting with onset of movement towards the jelly jar and going until either the onset of movement towards the peanut butter jar or, if following peanut butter spreading, the offset of the last spreading motion and 4) sandwich cutting – starting with the offset of the last spreading motion and going until the offset of the last cutting motion.

Finally, for each subtask, experimental objects were categorized as 1) relevant or 2) irrelevant. Objects were categorized as relevant for a given subtask, across all subjects, if they were used by any participant to complete the subtask. Objects that were never used by a participant for a subtask were categorized as irrelevant. For example, for the peanut butter spreading subtask, in course of completing the task, at least one maker actively used the bread slices, PB jar, PB lid, knife, fork, plate and napkin, and thus all of these items were categorized as relevant to the subtask. Notably, while the fork was initially intended as an irrelevant object, it was actually used by multiple makers and thus categorized as relevant. The remaining items (jelly jar, jelly lid, flower, cola cap, cola bottle, sippy cup, kid fork, and bread bag) were not actively used by even a single maker to complete the peanut butter spreading subtask and thus were categorized as subtask-irrelevant.

Analyses

In the current study we restrict our analyses to subtasks 2 (peanut butter spreading) and 3 (jelly spreading). We made this choice for two reasons. First, due to differences in the number of sandwiches made and eaten between parent-infant and adult-adult dyads, there were consequently differences in subtasks 1 (bread preparation) and 4 (sandwich cutting) between dyad types that were unrelated to our research questions. For example, parents were more likely to take out slices of bread for multiple sandwiches in the first trial and thus did not interact with the bread bag in subsequent trials as often as adult makers. Parent makers were also more likely to interact with previously made sandwiches during the sandwich cutting subtask, such as moving them out of the way or giving them to the infant. Second, there are no theoretical reasons why we might expect differences between spreading peanut butter versus spreading jelly. Thus, focusing on these two subtasks provides the opportunity to collapse across them – maximizing our data while simplifying our results and inferences.

Results

We begin by presenting analyses of duration and relative ordering of the PBJ spreading and Jelly spreading subtasks to justify collapsing across them. Next, we compare infant and adult observer gaze patterns to determine whether and to what degree infants pay attention to ongoing events when parents prepare food. We then explore gaze proportions to subtask-relevant objects, normalized for on-task looking time, to determine if when infants do pay attention, are they attending to the same task-relevant objects as adult observers in the same context. Finally, to ask how well infants and their parents coordinate their visual attention in this everyday context, we explore the proportions and temporal patterns characterizing how infant observers and parent makers attended the same subtask-relevant objects.

Descriptions of the two subtasks performed by actors during making PBJ

We present the durations for subtasks 2 (peanut butter spreading) and 3 (jelly spreading) in table 1. T-tests confirmed Infant-Parent subtask durations were not significantly different from Adult-Adult subtask durations, nor were peanut butter spreading subtask durations significantly different from jelly spreading subtask durations (all p-values greater than .2).

Table 1. Peanut butter spreading and Jelly spreading subtask duration statistics for Infant-Parent and Adult-Adult dyads.

Subtask	Duration (s): mean (range), SD	
	Infant-Parent	Adult-Adult
PB	30.68 (15.67-66.2), 14.22	32.08 (9.23-51.33), 13.59
Jelly	25.93 (13.36-50.8), 9.4	26.11 (8.1-51), 11.36

Note. PB=Peanut butter spreading subtask. Jelly=Jelly spreading subtask.

Likewise, chi-square tests of the counts for the respective orders in which the peanut butter spreading and jelly spreading subtasks occurred did not reveal significant differences for Infant-Parent, $\chi2(1)$=.111, p=.74, nor Adult-Adult dyads, $\chi2(1)$=0, p=1. Based on the null results for comparisons of peanut butter spreading and jelly spreading subtask durations and relative ordering, we collapse across these two subtasks for all following analyses.

Overall gaze patterns from infant and adult observers

In table 2 we show descriptive statistics for observer gaze durations and frequencies in the peanut butter spreading and jelly spreading subtasks. Gaze mean duration and frequency were first calculated at the subtask level. Both infant and adult observers attended to an object for about 2 seconds before switching their attention to the next object. Infant and adult observer gaze durations did not significantly differ, $t(55.98)=0.79$, $p=.43$. However, infant observers made significantly fewer looks to coded regions of interest than adult observers, $t(41.45)=-2.58$, $p<.05$.

Table 2. Gaze duration and frequency by observer type.

Measure	Gaze statistics by dyad type: mean (range), SD	
	Infant Observer	Adult Observer
Duration (s)	2.08 (1.03-4.53), 0.91	1.9 (0.69-4.21), 0.84
Freq. (pm)	24.83 (0-49.31), 9.79	33.21 (14.01-74.15), 14.28

The Distribution of attention in infant and adult observers

In table 3 we present the proportions within subtasks during which infant and adult observers gazed at 1) experimental objects categorized as relevant to the current subtask, 2) experimental objects categorized as irrelevant and 3) social ROIs (i.e., face, torso, empty hands). Infant observers spent approximately half of the total time attending to the objects on the table. When they did so, they rarely looked at task-irrelevant objects (flower, bowl of fruit, etc.) and instead attended to task-relevant objects in the majority of the time. They also rarely looked at the social partner's face which is consistent with recent findings using head-mounted eye trackers to measure infants' attention in other everyday contexts such as toy play (Yu & Smith, 2013; 2017).

It is worth noting that in this naturalistic task of action observation, infants didn't attend to the task all the time while their parents were making PBJ right in front of them. In roughly half of the total time when infants didn't attend to neither the objects on the table nor the social partner's face, they usually attended to objects in their own hands or looked around the room. Adult observers, on the other hand, attended to the task for nearly the whole time, as instructed to do so. Thus, compared with adult observer, infants gazed at relevant objects for significantly smaller proportions of subtasks than adult observers, $t(53.77)=-7.11$, $p<.001$. However, they did not significantly differ in proportion of time looking at irrelevant objects, $t(46.99)=1.93$, $p=.06$, or social ROIS, $t(31.76)=-0.31$, $p=.76$.

Table 3. Proportion looking for subtask relevant and irrelevant objects, and social ROIs, by observer type.

ROI group	Proportion looking: mean (range), SD	
	Infant Observer	Adult Observer
Relevant	.43 (0-.87), .28	.81 (.4-.96), .14
Irrelevant	.07 (0-.51), .11	.04 (0-.16), .04
Social	.03 (0-.22), .03	.04 (0-.48), .1

When infant observers were on tasks, did they attend to the same objects as those attended by their adult counterparts? To answer this question, we calculated the proportions within subtasks for which observers looked at each of the relevant experimental objects normalized for the total proportion of the subtask for which they observed relevant object (Figure 3). A Wilcoxon signed rank test did not find a significant difference between the two distributions, $V=21$, $p=.91$. The results here suggest that when infant observers were engaged in the observation task, they not only spent substantial time on task-relevant objects, they also distributed their attention among those task-relevant objects in similar ways as experienced adult observers.

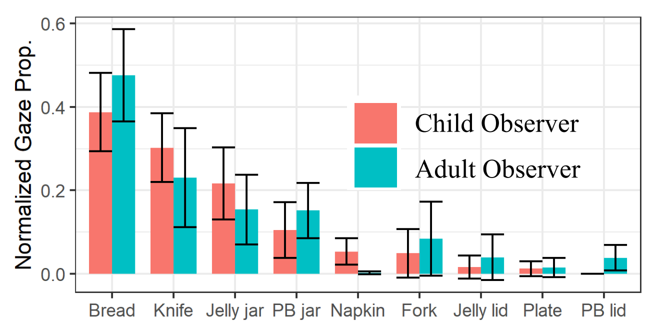

Figure 3. Infant (pink) and adult observer (blue) gaze proportions to relevant objects in peanut butter spreading and jelly spreading subtasks, normalized for proportion of time within subtasks looking at relevant objects.

Coordinated attention between infant observers and adult actors

Recent research using head-mounted eye tracking has examined coordinated attention between infants and their parents in joint toy play and found that infants and their parents are equal social partners as both lead the other's attention and both respond to the other's initiation of attend bid. Different with joint play, infants were primarily observers and parents were actors in the food preparation context. How did they coordinate their visual attention in this task? In table 4 we present the proportions within subtasks during which infant-parent and adult-adult dyads engaged in coordinated attention – looking at the same ROI at the same time – for 1) experimental objects categorized as relevant to the subtask, 2) those categorized as irrelevant and 3) social ROIs. Because infants were not on task in half of the total time, as expected, infant-parent dyads engaged in significantly less coordinated attention on object relevant to the subtask compared to adult-adult dyads, $t(34.95)=-6.02$, $p<.001$. However, the dyad types did not differ in the proportion of coordinated attention on objects irrelevant to the subtask, $t(41.24)=1.32$, $p=.2$, or social ROIs, $t(50.58)=0.82$, $p=.42$. Thus, both infant-parent and adult-adult dyads rarely looked at each other's face in this naturalistic task.

Table 4. Proportion in coordinated attention (CA) on subtask relevant and irrelevant objects, and social ROIs, by dyad type.

	Prop. CA: mean (range), SD	
ROI group	Infant-Parent	Adult-Adult
Relevant	.12 (0-.44), .11	.4 (.07-.84), .21
Irrelevant	.01 (0-.09), .02	0 (0-.02), 0
Social	0 (0-.06), .01	0 (0-.03), 0

Next, to explore whether infant-parent and adult-adult dyads engage in coordinated attention on the same objects when they were in coordinated attention on relevant objects, we calculated the proportions within subtasks for which dyads engaged in coordinated attention on each of the relevant experimental objects normalized for the total proportion of the subtask for which they engaged in coordinated attention on relevant objects (Figure 4). A Wilcoxon signed rank test did not find a significant difference between the two distributions, $V=10$, $p=1$.

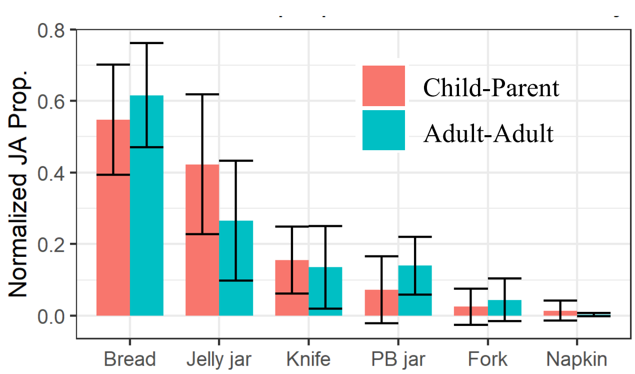

Figure 4. Infant-parent (pink) and adult-adult dyads (blue) proportions in coordinated attention on relevant objects in peanut butter spreading and jelly spreading subtasks, normalized for proportion of time within subtasks in coordinated attention on relevant objects.

Episodes of coordinated attention are initiated when a leader first looks at an object, after which a follower joins them in an episode of coordinated attention by subsequently looking at the same object. To explore how infants and their parents coordinate their visual attention in this everyday context, we next characterize the temporal patterns of coordinated attention episodes on subtask-relevant objects, for cases in which infants are followers and cases in which parents are followers, in terms of 1) duration, 2) frequency (pm), 3) the proportion of an individual's looks that follow their partner in coordinated attention out of all of their looks to subtask-relevant objects and 4) the lag at which followers join leaders in attending an object, all calculated at the subtask level (Table 5). We limit our analyses to coordinated attention episodes with a follower lag of less than 3 seconds, based on precedent in the literature (Yu & Smith, 2017).

Table 5. Coordinated attention (CA) durations, frequency, proportion following and lag following in peanut butter spreading and jelly spreading subtasks for infant-parent dyads, by CA follower, limited to subtask-relevant objects.

Measure	Values by CA follower: mean (range), SD	
	Infant-follower CA	Parent-follower CA
Dur. CA (s)	1.53 (0.03-11.7), 2.13	0.83 (0.03-1.93), 0.65
Freq CA (pm)	4.31 (0.91-9.69), 2.24	3.38 (1.13-7.33), 1.5
Prop. follow	.36 (0-.75), .22	.09 (0-.5), .11
Lag follow (s)	1.04 (0.04-3), 0.79	0.75 (0.03-1.83), 0.57

Looking at duration of episodes of coordinated attention, we first note the similarity between the results shown here and those presented in past work focusing on other everyday tasks. In particular, the mean duration of 1.53 seconds for cases in which infants followed their parents in coordinated attention is nearly identical to the mean duration of parent-led coordinated attention episodes reported in Yu and Smith (2017) for 12 and 18 mos engaged in free toy play with their parents. Comparing across followers, durations were significantly longer for episodes of coordinated attention in which infant observers followed compared to cases where parent-makers followed, $t(61.21)=2.11$, $p<.05$.

Second, turning to the frequency with which episodes of coordinated attention occurred, we once again first note the striking similarity with previous work – the frequency of 4.31 episodes per minute in which infants followed their parents in coordinated attention is once again nearly equivalent to values presented in Yu and Smith (2017). However, there was no significant difference between episodes of coordinated attention in which infants versus parents were followers, $t(44.99)=1.71$, $p=.09$.

Next, for the proportion of looks that follow the partner in coordinated attention out of all looks to subtask-relevant objects, there was a significantly larger proportion for infants versus their parents, $t(40.08)=5.75$, $p<.001$. This result highlights the fact that, although there are cases in which infant observers lead, the parent makers are the ones driving these situations via their control of the action sequences.

Finally, looking at the lags at which followers joined episodes of coordinated attention, we again first note the value of about 1 second is very similar to the results presented in Yu and Smith (2017). However, there was no significant difference comparing episodes in which infants versus parents were followers, $t(66.03)=1.86$, $p=.07$. In other words, children are following their parents in coordinated attention during action observation at a relatively timely manner, hinting that they are actively following their parents' ongoing sequences of actions.

Discussion

The current study was designed to explore three questions: 1) Do infants pay attention to ongoing sequences of events in naturalistic settings? 2) When infants do pay attention, do they attend to task-relevant objects and do they generate gaze patterns similar to adult observers in the same context? 3) How well do infants and their parents coordinate visual attention in such naturalistic settings? To answer these questions, we brought infant-parent and adult-adult dyads into a homelike environment and asked them to perform an everyday task – making peanut butter and jelly sandwiches – while wearing head-mounted eye-trackers.

We found that infant observers did not attend to their parents' activities as much as adult observers in the same situation. However, when infants were in engaged in action observation (i.e., attending to objects relevant to the ongoing task), their gaze patterns were distributed across task-relevant objects in a nearly identical way to adult observers. This suggests that the infants actively directed their attention to the most relevant objects for learning via observation about the peanut butter and jelly making sandwich task. Likewise, although infants and their parents coordinated their visual attention during the food preparation task significantly less compared to adult dyads, when they were engaged in coordinated attention, their focus was once again distributed across task-relevant objects similarly to adult dyads. Finally, the temporal patterning in which infants and their parents coordinated their attention was similar to as observed in other everyday tasks, such as toy play. Importantly, the toy play task differs from the sandwich making task in numerous ways, perhaps most crucially regarding the degree of active control of the situation that the infant versus parent has. The fact that we see such striking similarity regardless of such differences suggests that sensorimotor processes, such as low-level visual observation of sequences of actions, play a critical role in coordinated attention.

We next consider some limitations of our findings. First, while this project aims to explore how infants observe others' activities, here we limited our analyses to eye gaze patterns. Comparisons with adult observers and exploration of coordinated attention allowed for some inferences regarding action observation, but we are currently working on coding maker actions using a coding scheme based on the scheme employed by Hayhoe and colleagues (2003). With maker actions coded, we will be able to explore how attention relates to what is being acted on at a more fine grained level, and whether there are differences in maker actions demonstrated in front of an adult versus in front of a child – as might be expected based on evidence that caregivers modify their actions when interacting with infants versus adults (Brand, Baldwin, & Ashburn, 2002). Second, in the current study we limited our analyses to the peanut butter spreading and jelly spreading subtasks due to unintended differences in the bread preparation and sandwich cutting subtasks between infant-parent and adult-adult dyads. Indeed, given the relatively small sample size for the work presented here, these results should be seen as exploratory. In ongoing work, we plan to collect a much larger sample size, allowing for comparisons with our adult-adult dyads using only the first trials of infant-parent dyad data and within dyad comparison of trials where the child does versus does not have a sandwich. Finally, while our initial experimental setup was chosen so that relevant and irrelevant objects were relatively similar in terms of saliency, in future work we plan to employ saliency maps to explore whether this is actually the case, and whether there are differences between observers and makers or over the course of the experiment as makers actively modify the layout of objects in the course of completing the sandwich making task.

In conclusion, the current work extends what we know about how infants observe others' activities in everyday contexts. We found that children paid less attention to the sandwich making task than adults. However, when they were on-task, they displayed remarkably similar patterns of attention as adult observers – focusing on the objects and actions necessary for them to learn about the task. Moreover, infants and their parents coordinated their visual attention during the food preparation task in similar ways as those observed in other everyday tasks, such as toy play, suggesting that sensorimotor processes play a critical role in coordinated attention.

Acknowledgements

This research was supported in part by NICHD grants R01HD074601 and R01HD093792.

References

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, 5(1), 72-83.

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental science*, *15*(2), 292-298.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental science*, *15*(1), 35-42.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature neuroscience*, *9*(7), 878-879.

Gredebäck, G., & Falck-Ytter, T. (2015). Eye movements during action observation. *Perspectives on Psychological Science*, *10*(5), 591-598.

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of vision*, *3*(1), 6-6.

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130490.

Monroy, C. D., Gerson, S. A., & Hunnius, S. (2017). Toddlers' action prediction: Statistical learning of continuous action sequences. *Journal of experimental child psychology*, *157*, 14-28.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. v. 3.6.2.

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech directed to infants during home activities. *Child development*, *90*(6), 2135-2152.

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).

Yu, C., & Smith, L. B. (2017). Multiple sensory-motor pathways lead to coordinated visual attention. *Cognitive science*, *41*, 5-31.