**Appendix 1: Supplemental Information Regarding Dynamic Head Mounted Eye-Tracking Methods and Procedures**

**Participants**

Participants included 1 male, Caucasian infant and his mother who participated in an a very early parent mediated intervention for autism spectrum disorder. This caregiver-infant dyad provided dynamic head mounted eye-tracking data at two time points, including prior to the beginning of treatment at the infant age of 14-months and following treatment at the infant age of 18-months. A sample consisting of 13 similarly aged infant peers (average age = 16.8 months, range 14.2 – 19.9 months) and their mothers also participated. Infants were recruited through a university-maintained database of potential participants interested in research and community organizations (e.g., children’s outreach events, libraries) that serve a diverse population.

**Stimuli**

There were six 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 ranged in size from 5 to 8 cm in length, 6 to 12cm in width, and 4 to 6cm in depth when measured from their gravitational upright (flat bottom of object placed on a surface). These were organized into two sets of three so that each object in the set had a unique uniform color.

**Experimental Setup**

Caregivers and infants sat across from each other at a small table. Caregivers sat on the floor such that their eyes and heads were at approximately the same distance from the tabletop as those of the infants. Both caregivers and infants wore head-mounted eye trackers (Positive Science, LLC, http://www.positivescience.com/; also see (Franchak, Kretch, Soska, & Adolph, 2011)). The Positive Science eye tracker was designed for use with infants, and was designed to be attached to the head so as to be stable on the head (even in self-locomoting infants and toddlers, see (Franchak & Adolph, 2010; Franchak, van der Zalm, & Adolph, 2010). Both caregiver and infant eye-tracking systems include 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 108°, providing a broad view but one less than the full visual field - approximately 170° (Smith, Yu, Yoshida, & Fausey, 2015). 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.

**Instructions and Procedure**

Caregivers were told that the goal of the experiment was to better understand how caregivers and infants interact with objects during play, and then asked to engage their infants with the toys as they normally would at home. Each of the two sets of toys was played twice for 1.5 min, resulting in 6 min of play data from each dyad. Order of sets (ABAB or BABA) was counterbalanced across dyads.

**Data Processing**

During postprocessing and before coding, the quality of the eye-tracking videos (with eye images superimposed) for each caregiver-infant dyad was checked to ensure the quality of calibration at multiple time points throughout the recordings. Recalibration was conducted as necessary.

Eye tracking data was collected at a rate of 30 frames per second for approximately 360 s (four trials with 1.5 min per trial) of interaction, yielding potentially 10,800 data points per measure for each participant. The main data for analyses were gaze data directed to four regions-of-interest (ROIs). The four ROIs were defined in the head-camera videos: the three toy objects and the partner’s face. From gaze ROI coding, each dyad provided two gaze data streams containing the four ROIs as shown in Supplemental Figure 1. All results are reported in terms of proportion of the total interaction time - as if the number of recorded frames equaled the total number of possible (i.e., 10,800) frames. Therefore, estimates of percentage time on ROIs are underestimates because they include both off-task time and eye-tracking failures.

**Joint Attention**

Joint attention (JA) was defined as periods during which caregivers and infants were visually fixated on the same object at the same time (Yu & Smith, 2013). Previous research has shown that caregivers, but not infants, monitor the whole scene by often glancing very briefly to other objects or the infant’s face even while more generally attending to the same object as their infant (Yu & Smith, 2013). In addition, given that shared attention should last some amount of time longer than a single video frame (33 ms), we defined a joint attention bout as a continuous alignment of caregiver and infant fixation lasting longer than 500 ms. However, because humans generate three saccades per minute, looks briefer than 300 ms elsewhere where still allowed (i.e., one brief look away before switching back to the target was allowed). Examples of joint attention bouts are shown in Supplemental Figure 1.

**Sustained Attention**

Sustained attention (SA) was defined by consideration of the infant gaze alone. A 3 second period or greater of consistent looking by the infant within an ROI for a single object *without any looks elsewhere* was counted as sustained attention on that object by the infant. The 3 seconds was chosen as the threshold because it was the average duration of concentrated attention for 1-year-olds reported in a recent study using head-mounted eye tracking (Yu & Smith, 2016) and because this is the same threshold used by other researchers as defining a period of sustained attention (Ruff, Lawson, Parrinello, & Weissberg, 1990). Examples of infant sustained attention are shown in Supplemental Figure 1.

Supplemental Figure 1: Illustration of individual data streams color coded by region of interest for each participant and periods of joint attention and infant sustained attention contained within them.

**REFERENCES**

Franchak, J. M., & Adolph, K. E. (2010). Visually guided navigation: head-mounted eye-tracking of natural locomotion in children and adults. *Vision Res, 50*(24), 2766-2774. doi:10.1016/j.visres.2010.09.024

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 Dev, 82*(6), 1738-1750. doi:10.1111/j.1467-8624.2011.01670.x

Franchak, J. M., van der Zalm, D. J., & Adolph, K. E. (2010). Learning by doing: action performance facilitates affordance perception. *Vision Res, 50*(24), 2758-2765. doi:10.1016/j.visres.2010.09.019

Ruff, H. A., Lawson, K. R., Parrinello, R., & Weissberg, R. (1990). Long-term stability of individual differences in sustained attention in the early years. *Child Dev, 61*(1), 60-75.

Smith, L., Yu, C., Yoshida, H., & Fausey, C. M. (2015). Contributions of head-mounted cameras to studying the visual environments of infants and young children. *J Cogn Dev, 16*(3), 407-419. doi:10.1080/15248372.2014.933430

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), e79659. doi:10.1371/journal.pone.0079659

Yu, C., & Smith, L. B. (2016). The Social Origins of Sustained Attention in One-Year-Old Human Infants. *Curr Biol, 26*(9), 1235-1240. doi:10.1016/j.cub.2016.03.026