Developmental Cognitive Neuroscience 64 (2023) 101299

Contents lists available at ScienceDirect

Developmental Cognitive Neuroscience

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

Infant embodied attention in context: Feasibility of home-based head-mounted eye tracking in early infancy

Jessica Bradshaw [a,b,*], Xiaoxue Fu [a,b], Julia Yurkovic-Harding [a,b], Drew Abney [c]

[a] University of South Carolina, 1800 Gervais St., Columbia, SC 29201, USA
[b] Carolina Autism and Neurodevelopment Research Center, University of South Carolina, USA
[c] University of Georgia, 125 Baldwin St., Athens, GA 30602, USA

ARTICLE INFO

Keywords:
Eye tracking
Infancy
Attention
Autism
Social attention

ABSTRACT

Social communication emerges from dynamic, embodied social interactions during which infants coordinate attention to caregivers and objects. Yet many studies of infant attention are constrained to a laboratory setting, neglecting how attention is nested within social contexts where caregivers dynamically scaffold infant behavior in real time. This study evaluates the feasibility and acceptability of the novel use of head-mounted eye tracking (HMET) in the home with $N = 40$ infants aged 4 and 8 months who are typically developing and at an elevated genetic liability for autism spectrum disorder (ASD). Results suggest that HMET with young infants with limited independent motor abilities and at an elevated likelihood for atypical development is highly feasible and deemed acceptable by caregivers. Feasibility and acceptability did not differ by age or ASD likelihood. Data quality was also acceptable, albeit with younger infants showing slightly lower accuracy, allowing for preliminary analysis of developmental trends in infant gaze behavior. This study provides new evidence for the feasibility of using in-home HMET with young infants during a critical developmental period when more complex interactions with the environment and social partners are emerging. Future research can apply this technology to illuminate atypical developmental trajectories of embodied social attention in infancy.

1. Introduction

Infant-caregiver interactions are critical for the development of infant communication and language skills. The form and structure of these interactions shift across early development as infants gain new motor and cognitive abilities. In the first few months of life, infant-caregiver interactions are predominantly dyadic and face-to-face. In the latter half of the first year, as infants' play and motor skills develop, these interactions become triadic, involving the infant, objects, and caregivers. Coordinated infant and caregiver attention, or joint engagement, during triadic interactions has been established as an important context for language learning (Bakeman and Adamson, 1984; Carpenter et al., 1998; Carpenter and Tomasello, 2000). Differences in how infants allocate attention to caregivers and objects during these early dyadic and triadic interactions could have a profound impact on the acquisition of communication skills, especially for infants already at an elevated likelihood for social and language impairments (Adamson et al., 2019; Mundy et al., 2009; Mundy and Burnette, 2005).

Traditionally, studies of infant attention during dyadic or triadic interactions use granular behavioral coding of infant gaze from third-person perspective video. This "naturalistic" social context (i.e., live interactions with people and objects) is ideal for capturing gaze behavior as infants move through and interact with the environment, but the temporal and spatial resolution of gaze coding is limited by the number of cameras and available camera angles. On the other hand, studies of infant attention using screen-based eye tracking can capture gaze behavior with good spatial and temporal resolution, but this method neglects a key component of infant cognition: the infant's dynamic, embodied interaction with people and objects. Dynamic, embodied interactions refer to infants' active and constantly changing bodily engagement with the surrounding environment, which includes their physical interactions with caregivers and objects. It reflects infants' active participation as they explore, engage, and respond to their environment. Given that attention is nested within this embodied sensory-motor system that continuously drives and modifies infant attention (Adolph and Hoch, 2019), research methods and technology that can flexibly capture gaze behavior during dynamic, unrestricted, interactions with good spatial and temporal resolution is needed. Studies

* Corresponding author at: University of South Carolina, 1800 Gervais St., Columbia, SC 29201, USA.
E-mail address: jbradshaw@sc.edu (J. Bradshaw).

https://doi.org/10.1016/j.dcn.2023.101299
Received 16 March 2023; Received in revised form 24 August 2023; Accepted 8 September 2023

that leverage wearable technology to study infant attention in this way highlight the importance of this approach in generating new knowledge about infant social interactions (Adolph and Hoch, 2019; Franchak and Yu, 2022; Smith et al., 2015).

The application of head-mounted eye tracking (HMET) to the study of infant-caregiver interactions has offered insight into how caregiver gaze and action guide infant gaze and action. Franchak et al. (2018) showed that infants' social looking was constrained by their own posture and that of their social partner. Abney et al. (2020) documented the composition of gaze components that make up coordinated attention during triadic interactions, and how caregivers' behaviors are organized around these coordinated attention bouts. Hoch et al. (2020) demonstrated that infants' gaze patterns were drawn to objects within arms' reach for both crawlers and walkers. A longitudinal approach to studying infant embodied attention in early infancy can shed light on developmental trends in gaze patterns and the implications therein for the development of social and language skills. For example, egocentric cameras on very young infants have revealed developmental differences in the frequency and duration of faces within the infants' visual field (Fausey et al., 2016; Jayaraman and Smith, 2019). Fausey et al. (2016) observed that throughout the first two years, the proportion of faces in their visual field decreased, while the proportion of hands in their visual field increased. Jayaraman and Smith (2019) further observed that young infants have higher proportion of faces in their visual field and these faces were also more temporally persistent compared to faces in the visual field of older infants. Egocentric cameras demonstrate the availability of visual information in the infant's field of view, whereas wearable eye trackers can estimate gaze direction and location time-locked to environmental events. This visual input serves as critical data for infant learning, which varies depending on the infant's context, including location, body posture, behavior, and the presence of objects and people in their environment. As infants develop and acquire new skills, their visual input changes, reflecting their evolving interactions with the world. Quantifying early gaze patterns within the infant's familiar context (i.e., their home environment with a caregiver) may be particularly valuable, as this is the everyday context from which infants create statistical data for learning (Smith et al., 2018).

Despite significant advances in HMET for infants, toddlers, and children (Franchak and Yu, 2022; Pérez-Edgar et al., 2020; Slone et al., 2018), there are two significant limitations to research conducted thus far. First, HMET studies to date have been limited to infants who have developed independent head control, typically those older than 5–6 months of age. This is largely because the accuracy of HMET depends on successful headset and camera placement followed by an adequate tracking procedure to be used for online or offline calibration, and subsequent headset and camera stability (Hessels et al., 2022; Niehorster et al., 2020). This process represents a major technical challenge for younger infants still relying on external head support and who spend the majority of their time in supine or supported/reclined positions. Second, HMET systems are time- and labor-intensive to prepare and apply, and thus have been largely limited to laboratory-based studies. However, lab-based HMET studies of infant gaze may provide an inaccurate, or simply different, illustration of how infants interact with and learn from their familiar, daily, home environment. In contrast to the highly-controlled laboratory setting, the home affords a specific, highly familiar context in which infants spend a large proportion of their waking hours and experience consistent, repeated learning opportunities. Removing infant-caregiver dyads from their home is removing a critical context from which they develop attention, play, and communication skills (Franchak, 2020; Smith et al., 2018).

Despite these challenges, HMET technology could be of great value for mapping developmental pathways that lead to atypical social and communication development. Only recently has this approach been applied to the study of atypical development (Yurkovic et al., 2021). Autism spectrum disorder (ASD) represents an ideal application of this approach, as significant evidence points to both motor and attentional abnormalities in young infants at an elevated genetic likelihood of ASD (Bradshaw et al., 2020, 2022, 2023; Leezenbaum and Iverson, 2019; Macari et al., 2020). Screen-based eye-tracking studies largely suggest that infants and toddlers with ASD show decreased social attention, characterized by reduced looking to complex social scenes (Bradshaw et al., 2019; Chawarska et al., 2013; Pierce et al., 2016; Shic et al., 2014, 2020; von Hofsten et al., 2009). However recent findings using HMET during live social interactions present a contrasting picture wherein toddlers with and without ASD show comparable social attention, characterized by similar rates of looking to faces (Yurkovic-Harding et al., 2022). These results highlight the importance of context-specific influences on gaze behavior. Specifically, gaze during live interactions may be influenced by infant movement, object exploration, caregiver movement, and caregiver scaffolding. Child-caregiver interactions are rich with the active co-construction of interaction goals and activities, often with the caregivers scaffolding the interaction by introducing new toys or modifying activities (Sameroff, 2009). HMET studies have documented how child attention is altered by caregiver input as caregivers dynamically scaffold child behavior in real time. For example, one HMET study of infant-caregiver toy play showed that caregiver scaffolding behavior, such as looking to, talking about, and touching a toy, significantly increased the likelihood and duration of infant sustained attention to that object (Suarez-Rivera et al., 2019). Thus, HMET can reveal context-specific effects that shape social attention in young children with and without ASD.

The application of HMET technology to the familiar home setting, with infants younger than six months, and with infants at elevated likelihood for atypical development or ASD therefore could significantly advance our understanding of social attention within nested contexts and over developmental time. However, HMET has yet to be applied in these settings and the feasibility of doing so has not been assessed. One particular concern for collecting HMET data in less constrained environments is HMET data quality (Valtakari et al., 2021). A number of factors, including changes in the distance between the infants and areas of interest (AOI), changes in ambient lighting, changes in infants' posture, and headset slippage, can disrupt the detection of pupil location and corneal reflection. This disruption can introduce errors and compromise data quality (Franchak and Yu, 2022; Niehorster et al., 2020). For young infants with limited cognitive and motor abilities or control, these data quality issues can be exacerbated. This is evidently the case for screen-based eye-tracking studies in which data quality can be affected by experimental factors like infant positioning (Hessels et al., 2015) and demographic factors like infant age, with younger infants demonstrating lower data quality compared to older infants (Hessels and Hooge, 2019; Wass et al., 2014). Given the limitations listed above for approaching this research in non-standardized settings (e.g., the home), with younger infants, and with infants at elevated likelihood for neurodevelopmental disorders, the primary aim of this study is to evaluate the feasibility, acceptability, and data quality of in-home HMET during caregiver interactions with 4- and 8-month-old infants who are at low and elevated likelihood for ASD. In order to establish feasibility and acceptability of this technology in this specific context with this population, we report on data collection procedures, infant and caregiver acceptability of procedures, longitudinal participant retention, eye-tracking data quality, and exploratory analyses of looking behavior.

2. Materials and methods

2.1. Participants

Participants included $N = 40$ infant-caregiver dyads with infants ($N = 18$ female) who were either at elevated genetic likelihood for ASD (EL; $N = 25$) or who had no family history of ASD (low likelihood, LL; $N = 15$). We did not exclude caregivers based on sex or gender, but all caregivers were mothers. Our participant sample was representative of local racial/ethnic demographics and included 35 % African American/

Black ($n = 14$), 7.5 % Asian ($n = 3$), 5 % Hispanic/Latino ($n = 2$), 50 % White (n = 20), and 2.5 % who did not report race/ethnicity ($n = 1$). The highest maternal education for dyads were as follows: 10 % high school/GED ($n = 4$), 40 % some college ($n = 16$), 27.5 % Bachelor's or Master's degree ($n = 11$), 7.5 % professional degree (MD, PhD, etc.; $n = 3$), and 15 % who did not report ($n = 6$). In addition, 50 % of families resided in non-urban areas (RUCA Code > 1; Hart et al., 2005).

Study visits occurred in the home at age 4 and 8 months. Eligibility criteria for all participants included full-term birth (≥37 weeks gestation), no congenital vision or hearing abnormalities, and no known genetic syndromes (e.g., Down syndrome and Fragile-X syndrome). Siblings of EL infants were required to have a documented diagnosis of ASD by a licensed medical provider, confirmed through record review by a licensed psychologist, and completion of autism screeners. Siblings of children with ASD exhibit greater variability in their social-communication development compared to the typical population (Messinger et al., 2013; Ozonoff et al., 2011). For this feasibility study, all participants were analyzed together. All procedures were approved by the University Institutional Review Board and families completed informed consent prior to any procedures.

2.2. Procedures

2.2.1. Experimental procedures

Procedures for each study visit included equipment setup, infant headset placement, object tracking procedure for offline calibration, and a 10-minute video-recorded infant-parent interaction. The 4-month visit also included a two-minute free play period before study procedures began so that experimenters could observe how infants and parents typically interact at home to inform decisions about the best location to complete the infant-parent interaction. At four months, infants were placed in an infant seat that was available in the home and that provided support in a reclined position. Parents were instructed to engage in face-to-face play (e.g., peek-a-boo, songs, tickling) for 3 min followed by toy play with a standard toy set for 5 min. At 8 months, infants and parents were seated on the floor facing each other with a standard set of toys between them. For infants who were not yet sitting independently, parents were told to place their infant in whichever position they feel most comfortable and is most representative of their usual play (e.g., supported sitting with a pillow, prone on stomach, etc.). Parents were instructed to play as they normally would for 10 min, and that they and infants could move about the room as they wish. A second object tracking procedure was administered if needed (i.e., if headset slippage or camera movement was observed) at the 3-minute mark for 4-month-olds and 5-minute mark for 8-month-olds. If the infant fussed out or fell asleep, a second experimental session was scheduled. Following eye-tracking data collection, experimenters completed a data quality form that contained questions about data collection (e.g., specific headset used, number of interaction sessions completed, infant position), data completion and reasons for incomplete data (e.g., fussiness, fell asleep), challenges to data collection (e.g., slippage, infant interference), and details about the child's mood and activity level during the session. Additionally, parents completed a form about acceptability and their perception of how infant behavior was affected by the experimental procedures.

2.2.2. Headset placement and object track procedure

Eye trackers consisted of two tethered systems (Positive Science, LLC) equipped with two cameras attached to a cap or headband that recorded eye and scene images. For 4-month-olds, a series of caps were custom-designed in-house to adequately fit a variety of infant head sizes, head shapes, and hair textures, and to remain stationary (i.e., centered on forehead without slippage) for infants without independent head control. For 8-month-olds, the standard infant Positive Science cap was used. See Fig. 1 for experimental setup and headsets. All interactions were recorded with three cameras: the mounted eye camera, the mounted scene camera, and a third-person perspective room camera placed on a tripod and adjusted occasionally to capture the full interaction. The infant's eye camera was always positioned to record data from the infant's right eye and the scene camera was centered between the eyes and positioned as low as possible on the forehead to best approximate the infant's view (Slone et al., 2018). Data were collected at a consistent frame rate of 29.98 Hz for the external room camera. Eye and scene images were captured via the manufacturer computer and software, resulting in a slightly variable frame rate due to CPU timing, averaging at 29.2 Hz (SD=0.09 Hz). Frame rate variability is not dependent on video content and does not differ as a function of infant behavior.

After headset placement, experimenters adjusted eye-tracking cameras and explained to parents the importance of keeping cameras in place during the interaction so they could help prevent camera movement. Next, an "object track procedure" was administered to be used for later offline calibration completed in Positive Science Yarbus software (described below). For this procedure, infants were placed in the supported reclined position (4 months) or parent's lap (8 months) and the experimenter presented a number of eye-catching toys of equal size (~1.75 in) from a distance of about 4 ft from the infant to achieve a visual angle of approximately 2º. Specifically, infants' gaze was attracted to nine general areas within their field of gaze split by vertical (upper, middle, and lower field of vision) and horizontal (left, center, and right) coordinates. Similar to other studies of this kind (e.g., Slone et al., 2018), this procedure was intended to elicit large angular displacements of the eye, but infants were free to move their eyes and/or heads as needed to attend to the toys. Strategies such as varying the toy velocity during the track procedure were used to elicit isolated eye displacement in the absence of head movement. Looks to each of the nine general areas were elicited at least three times, resulting in approximately 27 locations. This procedure was done to ensure that the resulting calibration accounted for the possible extremes of infant gaze.

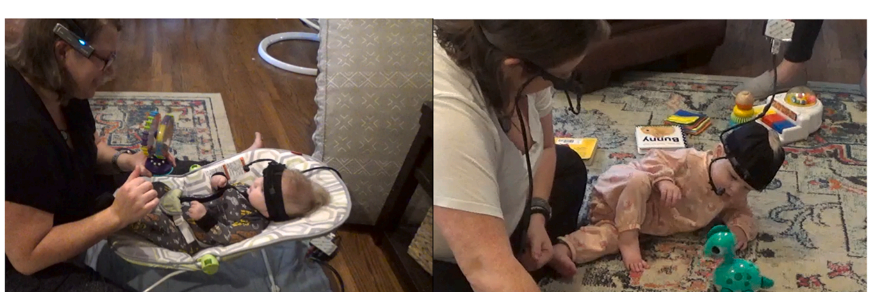

Fig. 1. Experimental Setup: (a) 4-month study visit with infant wearing custom headband in their own infant seat that provided support in a reclined position (left); (b) 8-month study visit with infant wearing standard Positive Science cap and placed on the floor with freedom to move about the room (right).

2.3. Data processing and quality analysis

Eye-tracking data quality was inspected by quantifying accuracy, precision, data loss, frequency and distribution of caregiver faces in the scene, infant gaze distribution, and infant gaze location (Hessels and Hooge, 2019; Holmqvist et al., 2022; Wass et al., 2014).

Accuracy was measured as the difference (in degrees of visual angle) between the point of gaze indicated by the eye tracker and the target location from the calibration track procedure (Holmqvist et al., 2022). Inaccuracy was computed using processed recordings from the object track procedure based on the method described in Franchak and Yu (2022). Precision was calculated as the median root mean square sample-to-sample deviation, reported in degrees, during the object track procedure when infants were fixating on a static target and infant movement was minimal (Hessels et al., 2022). Data loss was defined as the percent of the interaction that gaze location could not be determined from the calibrated gaze replay (i.e., bullseye was not visible; see calibration procedure described below). Possible reasons for data loss include transient occlusion of the eye or looks outside of the scene camera field of view.

The frequency and distribution (i.e., size and centeredness) of caregiver faces visible in the infant scene camera were examined using an open-source artificial intelligence (AI) face detector (RetinaFace; Deng et al., 2020). A Hidden Markov Model was used to smooth the resulting output, correcting moments where RetinaFace either missed a face (false negative) or found a non-face (false positive). To test reliability of RetinaFace, a human coder coded the number of faces present in a subset of video frames. Human and AI face detection agreed on the presence of faces 92 % (+/- 10 %) of the time, suggesting that the AI detector is reliably detecting faces. Frequency was defined as the percent of video frames that contained a face from the total number of frames for the play session. Size of the face was defined as the proportion of the scene that contained a face, such that larger values indicate larger faces from the infant's perspective. This was calculated by dividing the area of the detected face by the area of the entire scene view (i.e., video from the infant's head-mounted scene camera). Centeredness was calculated by computing the Euclidean distance of the center point of each face to the center of the scene view, such that smaller values indicate the face is more centered in the scene camera.

Infant gaze distribution and location were determined after an offline calibration procedure was completed to indicate the point of infant gaze in the scene (Yarbus software; Positive Science, LLC). The manufacturer software allows for calibration using pupil and corneal reflection detection, as well as pupil-only detection. In this study, similar to previous studies and existing guidelines of applying head-mounted eye tracking using similar headgear in young infant populations (e.g., Slone et al., 2018; Yurkovic et al., 2021), pupil-only calibration was used due to persistent challenges with detecting the corneal reflection reliably and consistently. The offline calibration procedure begins with a trained research assistant marking the calibration target locations on the scene image recording when the infant's point of gaze was clearly identifiable. The software applies an algorithm to map the pupil locations with the estimated gaze locations in the scene camera coordinate space. The manual identification of points of gaze and automated mapping procedure is then run iteratively to establish satisfactory calibration (Slone et al., 2018). A satisfactory calibration uses only points where the pupil is accurately detected within the eye camera, where the infant is looking within the bounds of the scene recording, and the linear fit between pupil position and gaze on the screen is greater than 97 % on both the x- and y-dimension. Following offline calibration, a "gaze replay" is created, consisting of the scene image recording with a crosshair and bullseye indicating the infant's gaze.

Gaze distribution was examined using heatmaps built from the gaze replay. A 3-D histogram was created that represented the x-axis location of gaze, the y-axis location of gaze, and the number of gaze moments that fell within that pixel on the z-axis. Each gaze point was represented by a bivariate gaussian kernel, such that the actual recorded x-y coordinate of gaze had a value of 1 and coordinates around the gaze had slightly decreasing values. The kernel allowed for slight error in the model of gaze during the interaction and for visual acuity (i.e., we see more than just the exact point that we are fixating on). Gaze distribution was defined as how far each of the infant's gaze points were from the center of all gaze points. Gaze centeredness was defined as how far each gaze point was from the center of the first-person scene view. Both analyses were represented as root mean squared error of distances.

Gaze location was examined based on established coding methods of HMET data (e.g., Franchak et al., 2011, 2018). First the gaze replay, eye image recording, and room camera recording were synchronized into a single composite video using Adobe Premiere Pro (see Fig. 2). Frame-by-frame synchronization of the eye and scene recordings occurs through PSLiveCapture software (Positive Science, LLC). Frame-by-frame synchronization among the gaze replay, eye, and room recordings was achieved by selecting several moments throughout the video that contained an identifiable action (e.g., infant touches an object) that was clearly visible in all three video streams. If needed, the room camera recording was down-sampled to match the rate of the gaze replay and eye camera recording so that the videos remained synchronized over time. The composite video was exported at a resolution of 1920×1080 pixels at 30 fps for frame-by-frame behavioral coding using Datavyu software (Datavyu Team, 2014). Coders inspected the composite video frame-by-frame to denote the onset and offset time of each valid fixation on one of three AOIs: parent face, parent body, and toys. A valid AOI gaze was defined as ≥ 3 consecutive frames ($\approx$99 ms) of stable gaze on the same AOI. Frames were coded as invalid if a) the gaze point was on an AOI for shorter than 3 consecutive frames, b) no gaze point (i.e., cross hair) was detected, or c) the gaze point was off the field of view. Coders established an average inter-rater agreement of 97 % ($\kappa = 0.96$). Gaze duration to each AOI was summed and divided by the total duration of valid looking for each play session to derive a proportion of looking to each AOI.

2.4. Statistical analysis

Paired-sample *t*-tests were conducted to compare eye tracking data quality metrics between age groups. Linear mixed effects models were used to detect age-related differences for all analyses. Every session from every participant was included in the models. Models treated age (4mo or 8mo) as a fixed effect and included a random intercept for each participant. Statistical significance was set at an alpha of 0.05.

3. Results

3.1. Feasibility and acceptability of data collection procedures

Feasibility and acceptability of data collection procedures were evaluated using rates and proportions of infants who tolerated the headset and contributed usable data, experimenter ratings of the occurrence and impact of headset slippage, parent reports of whether the interaction was typical or atypical, and participant retention rates from the 4-month visit to the 8-month visit. The flow of participants through the longitudinal study procedure is shown in Fig. 3. A total of 49 families were contacted to participate in the study, of which 40 (81.6 %) consented to participate. Of the 40 participants who consented, $n = 6$ were older than 4 months and so only completed an 8-month visit. At both the 4-month and the 8-month visits, 100 % of infants tolerated the placement of the headset, defined by successful cap placement and completion of the object track procedure. Partial or complete data was collected for all infants who attempted a 4-month visit ($N = 34$, 100 %). Complete data, defined as completion of the object track procedure and recorded, likely usable data for at least part of each play session, was collected for 24 of the 34 infants (70.59 %). Partial data collection, defined as completion of the object track procedure and recorded data

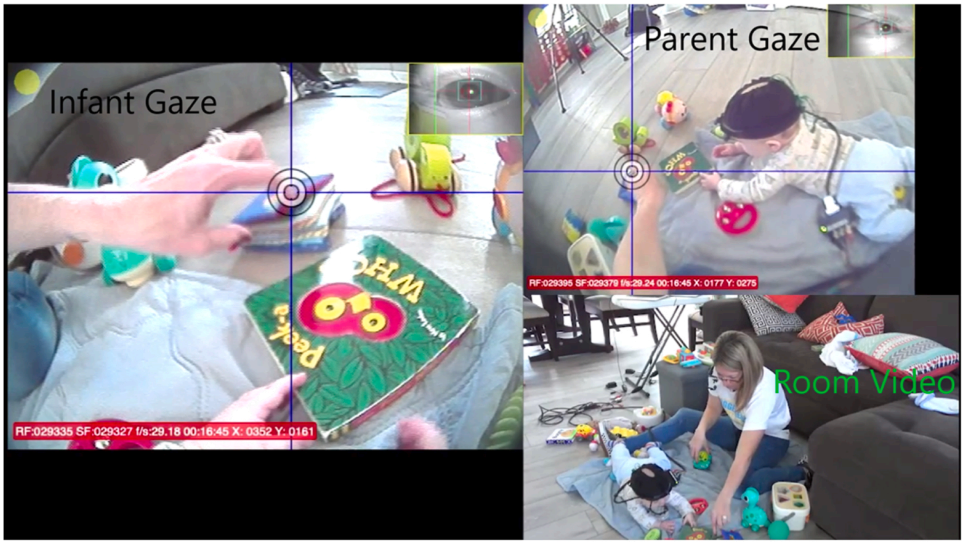

Fig. 2. An example frame of the composite video used for gaze coding. The present paper focuses on the infant gaze data.

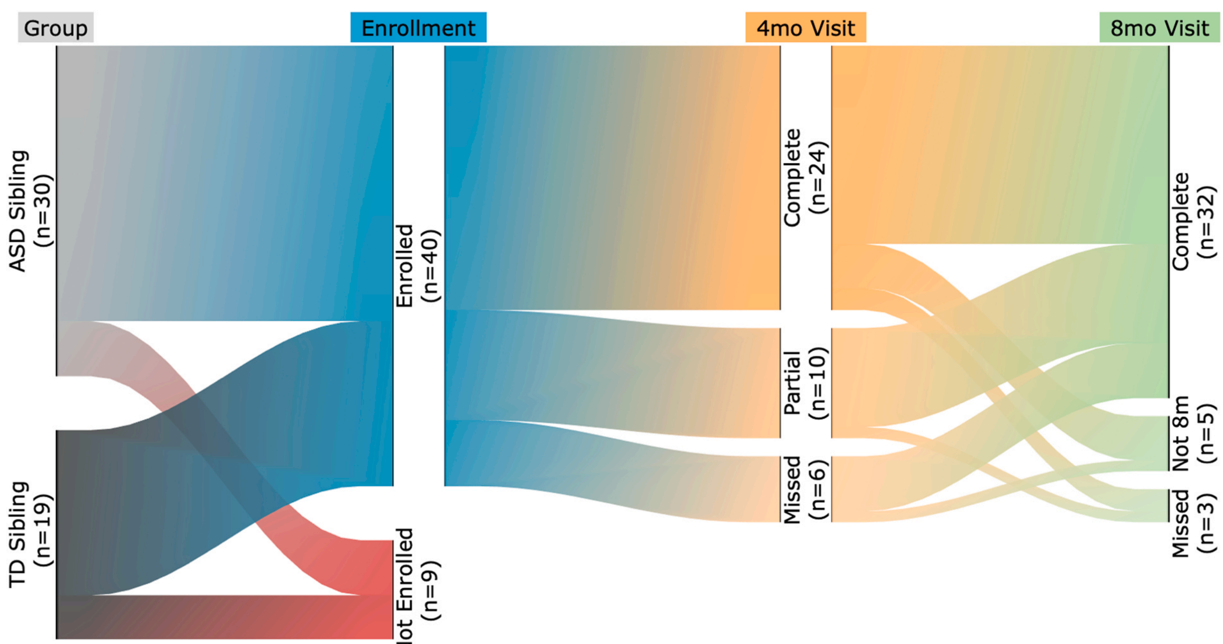

Fig. 3. A flow diagram shows each participant's movement through our study protocol. A total of 49 families were contacted for participation in the study because the infant had an older sibling that either had an ASD diagnosis ($n = 25$) or was TD ($n = 19$). Of these families, 40 enrolled in the study (ASD Sibling: $n = 25$, TD Sibling: $n = 15$). The remaining 9 families either declined participation or did not respond to our attempted contact. At the 4-month visit, we were able to place the eye tracking equipment and collect data from all 34 enrolled families (ASD Sibling $n = 15$, TD Sibling $n = 14$), with 24 of those visits having complete data and 10 having partial data. The remaining 6 families either enrolled after the infant was 4mo or we were unable to schedule a visit within the necessary timeframe. At the 8-month visit, we collected complete data from all $n = 32$ of the enrolled families. An additional $n = 5$ were not yet 8 months, and we were unable to schedule $n = 3$ visits within the necessary timeframe.

for part of at least one play session, was collected for the remaining 10 infants (29.41 %). Reasons for partial data collection included: fussing out ($n = 3$), falling asleep ($n = 2$), parent/sibling interference ($n = 4$), or equipment malfunction ($n = 1$). Of the $N = 40$ participants, $n = 32$ attempted an 8-month study visit, $n = 5$ infants had not yet turned 8 months, and $n = 3$ infants were lost to attrition. Of the infants who attempted an 8-month visit ($N = 32$), 100 % contributed complete data.

During each play session experimenters noted the occurrence and impact of headset slippage on the eye image. Significant slippage usually required a camera adjustment and/or repeated object track procedure, while moderate slippage occasionally warranted a repeated object track procedure, based on the experimenters' judgement. For the 4-month visit, slippage was noted to affect the eye image on at least one occasion in 17.6 % of sessions (significant slippage: 0 %, moderate slippage: 17.6 %). For the 8-month visit, slippage affected the eye image in 18.6 % of sessions (significant slippage: 6.45 %; moderate slippage: 12.9 %). Overall slippage was not significantly different by age ($\chi^2 = 0.43$, $p = .51$).

Across all visits, the majority of parents reported that their infants displayed overall typical behavior during the interaction (71.1 %), including typical interest in toys and parent. For those that reported atypical behavior, a minority of them (25 %) attributed this to the equipment/headset. Participant retention from the 4- to 8-month visit was high at 91.2 %.

3.2. Eye-tracking data quality

3.2.1. Accuracy, precision, and data loss

HMET inaccuracy was calculated as the average difference between the point of gaze indicated by the eye tracker and the target location from the object track procedure (Holmqvist et al., 2022). Results show an average inaccuracy between 2 and 3 degrees and suggest that accuracy increases with age (4mo: M=3.11°, SD=0.62; 8mo: M=2.53°, SD=0.44; $t = 3.29$, $p < .01$; see Supplementary Figure 1). In this study, gaze location was depicted with a bullseye with three rings (see Fig. 2). The middle ring was sized at a radius of 18 pixels, which allows for an inaccuracy of 2.49°, while the outer ring is sized at a radius of 28 pixels, allowing for an inaccuracy of 3.87°. Precision was calculated as the median root mean square sample-to-sample deviation for each participant at 4 months (M=0.25°, SD= 0.13°) and 8 months ($M = 0.37°$, $SD =$ 0.21°), with 8-month-olds showing significantly higher precision ($t = 3.05$, $p < .01$; see Supplementary Figure 2).

In regard to data loss, our data showed < 10 % invalid data from the calibrated gaze replay over the total experiment across all ages (4mo and 8mo; see Table 1 and Supplementary Figure 3) and participant groups (EL and LL), with no main effect of age or group, suggesting satisfactory robustness (Holmqvist et al., 2012). We also tested for possible age- and group-specific differences in data loss for younger vs. older infants and for EL vs LL infants. There were no statistically significant differences in the proportion of invalid looking for data collected at 4mo vs. 8 mo ($t = -0.06$, $p = .95$) or for EL vs. LL infants ($t = -1.07$, $p = .30$).

3.2.2. Frequency and distribution of faces in the scene

The proportion, size, and centeredness of faces detected in the infant head-mounted scene camera during the interactions differed significantly across age groups (see Fig. 4). Compared to 8-month-olds, 4-month-olds had access to significantly more faces (4mo: 76.47% (SD=28.4); 8mo: 17.41% (SD=17.9); F(1,65)= 115.31, $p < 0.01$) that were bigger (4mo: 5.44% (SD=2.8); 8mo: 3.78% (SD=3.1); F(1,63)= 4.84, $p = 0.03$) and more centered (4mo: 38.78% (SD=6.3); 8mo: 45.9% (SD=451.9); F(1,63)= 15.40, $p < 0.01$). Within the 4-month-old interactions, the introduction of toys in the play session reduced the size (Face-to-Face: M=7.15 %, SD=3.32 %; Toy Play: M=4.02 %, SD=1.10 %; F(1,32)= 12.93, $p < 0.01$) and centeredness (Face-to-Face: M=34.44 %, SD=6.03 %; Toy Play: M=42.33 %, SD=3.69 %; F(1,32)= 20.15, $p < 0.01$) of faces.

3.2.3. Gaze distribution

Gaze distribution heatmaps visualize each infant's gaze location within their first-person scene views during the interaction (see Fig. 5). Analyses of these gaze distributions show that gaze was mostly centered in the scene camera image, with gaze of older, 8-month-old infants trending more centered (4mo: M=0.36, SD=0.07; 8mo: M=0.32, SD=0.07; $t(36) = 1.85$, $p = 0.07$) and being significantly less distributed (4mo: M=0.22, SD=0.02; 8mo: M=0.14, SD=0.05; $t(36) = 2.55$ $p < 0.02$) than that of younger, 4-month-old infants.

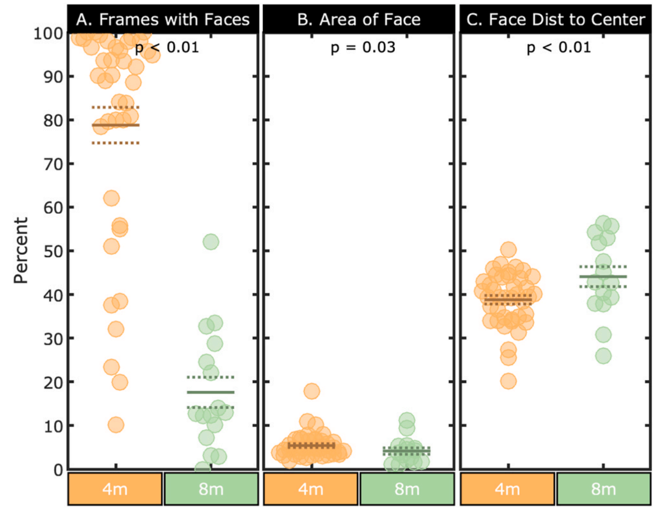

Fig. 4. Faces were automatically detected from the infant's first-person perspective. (A-C) Each dot represents one participant's session (i.e., one dot for 4mo supported dyadic play and one for 4mo supported triadic play). Group means are represented by the solid lines, and standard errors are represented by the dotted lines. (A) We calculated the percentage of all frames where at least one face was detected. Faces were more available to 4mo infants than to 8mo infants. (B) We calculated the area of each face relative to the entire recorded scene view. Each dot is the median area of faces. Faces are bigger in the views of 4mo infants than 8mo infants. (C) We calculated the distance of each face to the center of the recorded scene view. Each dot is the median distance of faces to center. Faces are more centered in the views of 4mo infants than 8mo infants.

3.2.4. Gaze location

Gaze location was examined in terms of the proportion of time spent looking to each of the three AOIs (parent face, parent body, and toys). Descriptive statistics are presented in Table 1 and Fig. 6. Overall, 4-month-olds spent a greater proportion of time looking to the parent's face and body, and less time looking at the toys (see Fig. 6). When considering face-to-face and toy play at 4 months, face looking decreased when toys were introduced to the interaction (B=−12.57, $t = 5.24$, $p < .001$) and decreased even further during toy play at 8 months (B=−7.42, $t = -2.95$, $p < .01$) (see Table 1). Looking to the parent's body followed a similar pattern, decreasing when toys were introduced to the interaction at 4 months (B=−8.07, $t = 4.46$, $p < .001$) and decreasing further during toy play at 8 months (B=−7.76, $t = -4.24$, $p < .001$). Lastly, the proportion of time looking to the toys increased from 4 to 8 months (B=6.32, $t = 2.75$, $p < .05$).

4. Discussion

The primary aim of this study was to evaluate the feasibility of collecting, processing, and analyzing HMET data to capture embodied attention of young infants during social interactions with a caregiver within their familiar home context. In particular, we examined the use of HMET with infants as young as 4 months of age whose lack of independent head control pose specific challenges to collecting good-quality HMET data. We also include a sample of infants at elevated likelihood

Table 1
Descriptive Statistics of Proportion of AOI Looking Duration.

	% Total AOI Looking Duration Over Total Valid Looking Duration					% Invalid Over Coded Duration
	Toy	Face	Body	Self	Background	Invalid
4 months		16.00(16.68)	21.25(11.47)	0.83(1.09)	40.84(21.30)	
F2F 4 months		25.08(19.33)	25.53(11.56)	1.27(1.39)	48.73(23.51)	6.63(5.45)
TP 4 months	41.88(17.72)	7.37(6.44)	17.19(10.04)	0.45(0.56)	33.34(16.20)	5.31(5.80)
TP 8 months	54.92(11.66)	5.10(7.70)	9.68(8.05)	2.65(2.81)	28.52(12.01)	6.44(8.28)

Note: Values reported as M(SD); F2F = face-to-face play; TP=toy play; AOI=area of interest.

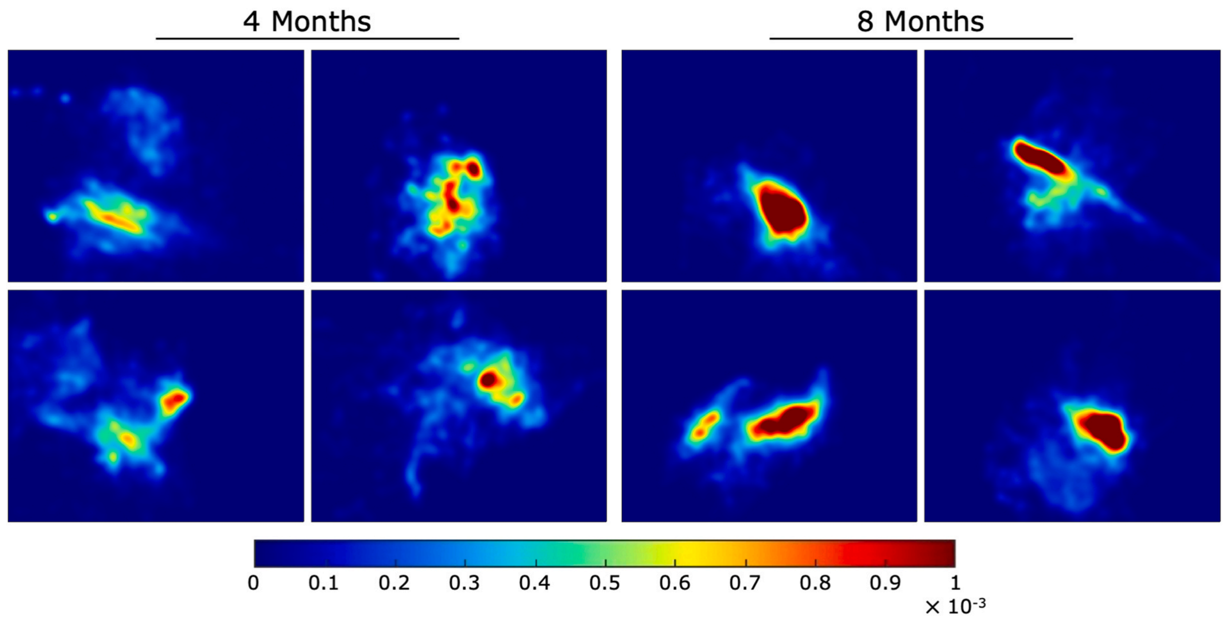

Fig. 5. Heatmaps represent the x-y coordinates of infant gaze for example play sessions. Each gaze point is modeled by a kernel density surrounding the actual x- and y-coordinates of gaze. Red areas represent areas of high gaze frequency, while blue areas represent areas of low gaze frequency.

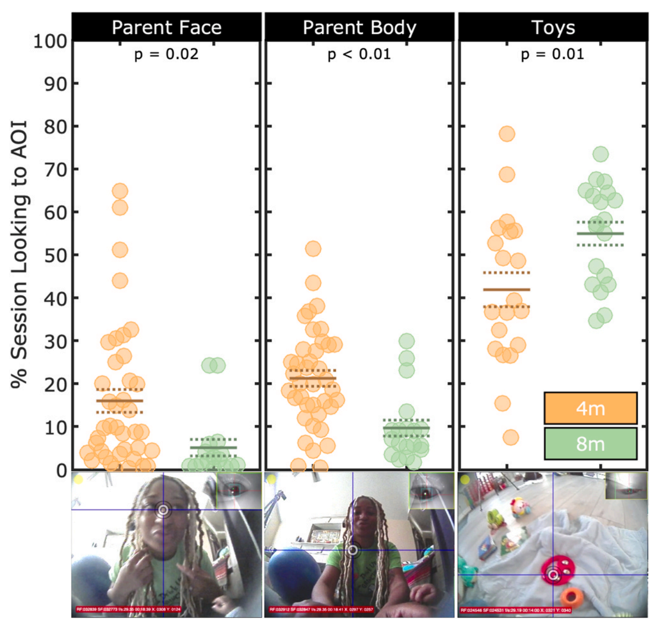

Fig. 6. Percentage of total looking at the areas of interest (AOIs) over the total valid looking duration by AOI type and age. Each dot represents one participant's session. Group means are represented by the solid lines, and standard errors are represented by the dotted lines. Parent face and parent body looking events occurred in sessions where toys were not presented. Examples of AOI gaze events are provided.

for ASD who experience atypical development, such as hypo- and hyper-sensory responses, anxiety, and social and communication delays (Ozonoff et al., 2011), to examine potential group-specific differences in feasibility. Gaze patterns across age and interaction contexts were also explored to further document feasibility of using these procedures to study early infant social attention.

Our results indicate a satisfactory headset solution for the application of HMET to 4-month-old infants who have not yet acquired independent head control. The headset proved to be tolerable and sufficiently stable to capture gaze behavior. This was evident in our 100 % success rate for headset placement, completion of the object track procedure to be used for offline calibration, minimal headset slippage that did not differ by age, and at least partial experimental data collection for all participants. In addition, the majority of parents reported that their infant's behavior was not affected by the experimental procedures, with minimal observed discomfort or fussiness due to the headset. Caregiver acceptability of procedures is also supported by our high retention rate for longitudinal study visits. Overall, these data indicate successful equipment placement and collection of some potentially usable data from every infant in their home. Given the importance of an in-home study that reflects naturalistic interaction behavior within a familiar context, it is also critical that data collected here largely reflect the dyads' "typical" behavior, as perceived by the caregiver. While we are limited to parent-reported ratings of infant behavior, results suggest that these procedures largely capture infants' typical behavior within their home context. Acceptability measures did not differ by participant group, suggesting that infants at elevated likelihood for ASD tolerate HMET study procedures to the same degree as typically developing infants.

Next, the quality of eye-tracking data collected was evaluated in terms of accuracy, data loss, gaze distribution, and gaze location. Accuracy was evaluated by calculating the spatial offset between the point of gaze (i.e., crosshair) after offline calibration and the target location during the object track procedure. This measure of eye-tracking accuracy could be affected by individual infant, environmental, and/or procedural factors including the infant's ability to fixate on the calibration target, camera movement or slippage during the object track procedure, or the number of points available for offline calibration, among others. Our results suggest adequate accuracy within 2–3 degrees, comparable to that of infant stationary eye tracking and adult wearable eye tracking data (Hessels et al., 2015; Hooge et al., 2022). There was a significant effect of age with younger infants showing poorer accuracy, consistent with findings from stationary eye-tracking studies that show poorer data quality for younger infants (Hessels and Hooge, 2019). In regard to data loss, invalid data represented < 10 % of data collected during play sessions and rates of data loss did not differ by age or participant group. Similar to measures of acceptability, eye tracking accuracy and data loss was comparable across infants at elevated and low likelihood for ASD.

Gaze distribution heatmaps revealed generally centered and sufficiently distributed gaze locations that reflect typical real-world visual scanning contextualized within the infant's dynamic head and body movement (Bambach et al., 2014; Foulsham et al., 2011), and suggest good data quality. In addition, gaze became less distributed and possibly more centered as infants aged from 4 to 8 months. This finding may be driven by differences in data quality and/or developmental shifts in visual scanning. Evaluation of gaze centeredness depends on consistent placement of the scene camera relative to the head and eyes, and thus we cannot rule out the possibility that camera placement impacted this finding. Similarly, evaluation of gaze distribution depends on minimal headset slippage, which can artificially increase gaze distribution. Additional analyses of camera placement and slippage can help determine their contribution to gaze distribution results. In addition, the age-related findings are consistent with previously documented developmental gains in ocular motor control of eye movements and gaze shifting (Colombo, 2001) as well as gains in gross and fine motor abilities that have been shown to increase head-eye alignment and centered gaze (Bambach et al., 2017). Overall, these results provide some evidence for adequate data quality and suggest that additional analyses of camera placement and slippage are important for confirming data quality and making firm conclusions about observed developmental changes.

Finally, we explored the distribution of caregiver faces in the infant's scene image as well as infant gaze location during the interactions. Four-month-olds had caregiver faces visible in the scene recording for 76 % of the interaction and gazed to the face for 16 % of the interaction. Eight-month-olds had caregiver faces visible in the scene recording for 17 % of the interaction and gazed to the face for 5 % of the interaction. In addition, faces were found to be bigger and more centered for 4-month-olds. These developmental trends, suggesting a decline in the frequency of faces in the visual field and a decline in gaze to faces from 4 to 8 months, are consistent with existing research (Fausey et al., 2016; Franchak et al., 2018; Jayaraman and Smith, 2019; Suarez-Rivera et al., 2019; Yu and Smith, 2013) and provide additional evidence for adequate data quality. As previous studies suggest, age-related change in face looking may be linked to motoric and visual constraints of 4-month-old infants who have limited freedom of movement and lower visual acuity (Jayaraman and Smith, 2019). These constraints likely shape caregiver behavior wherein caregivers of infants with lower mobility and visual acuity may make their faces more accessible during social interactions by being closer, in view, and more centered, compared to those of older infants. Similarly, 8-month-olds were typically sitting and reaching independently, allowing them to look down and explore toys on their own, often putting caregivers further away and out of view (Kretch et al., 2014). The rate of face looking observed in a recent study of 2–4-year-olds (Yurkovic-Harding et al., 2022) is even lower than that observed in our sample of 4- and 8-month-olds, providing preliminary evidence that face-looking during triadic (child-toy-caregiver) interactions may decline from age 4 months to 4 years. These findings support feasibility of these experimental methods for identifying age-specific changes in gaze patterns. To make stronger interpretations about developmental changes in infant looking behavior, this work could benefit from additional controlled experimental studies that consistently manipulate key variables of interest, such as infant posture and caregiver behavior, across age groups.

4.1. Limitations and future directions

There are several important limitations of this feasibility study that are worth noting. In regard to data quality, it is notable that HMET accuracy was lower for 4-month-olds compared to 8-month-olds. Data with an inaccuracy greater than 3° may be problematic for AOI analyses that only allow for inaccuracies up to about 2.5°, as is the case with this study. Accordingly, we suggest that researchers conduct rigorous analyses of accuracy when collecting HMET data with younger infants and develop AOIs and coding schemes that are well-matched to the constraints of the quality of the data. Using the corneal reflection in addition to the pupil location for calibration may be one way improve accuracy and is worth systematic investigation in this population. While we report precision as a data quality metric, this study was not specifically designed to estimate precision in this population and experimental context. Accordingly, these results serve as an initial step in establishing guidelines for determining and reporting precision using HMET methods with young infants. Given the in-home setting of this study and the inherent challenges in applying automated detection models to dynamically changing scenes, particularly within complex and cluttered environments (Valtakari et al., 2021), comprehensive automated analyses of gaze data fell outside the scope of this study. However, application of automated approaches accompanied by validation of these methods would be a significant contribution in future work.

Data collection was conducted in the infant's home using a standardized toy set selected by the research team. The number and types of toys available in each home can vary greatly from household to household and while the environmental context (i.e., the home) was uniformly familiar across all infants, our standardized toy set was deliberately designed to mitigate the influence of toy volume and motoric/play affordances on gaze patterns. Still, this toy set likely included several toys that were unfamiliar to the infant, which may have led to differences in gaze and interaction behavior compared to a set of familiar toys. This warrants future study of toy novelty effects on infant gaze patterns.

Our assessment of how infant behavior was affected by experimental procedures was largely based on parent report, which has inherent limitations. Several unknown factors could affect parents' judgment of their infant's behavior. A more systematic approach to addressing this issue would be to conduct in-home sessions with and without the elements unique to this study (i.e., eye-tracking equipment, unfamiliar experimenters, novel vs. familiar toys). It was outside the scope of this paper to explore between-group differences (LL vs. EL) and the effect of infant motor skills and position/posture during interactions; these are important areas of future study.

5. Conclusions

This is the first study to describe methods, acceptability, and feasibility of HMET in 4- and 8-month-old infants at low and elevated likelihood of ASD in the home context. Overall, results suggest caregiver and infant acceptability and feasibility of the data collection procedures described. In particular, with the development of tailored equipment and procedures, collection of HMET data is feasible in 4-month-old infants who have not yet developed independent head control. Preliminary analyses provide evidence for feasibility of detecting developmental trends in infant gaze patterns during social interactions that align with existing work. Future directions include illuminating how moment-to-moment infant and caregiver gaze behavior during naturalistic interactions are influenced by context and genetic vulnerability and delineating the role of embodied social attention in the development of ASD.

Funding

This work was supported by the James S. McDonell Foundation (Grant #2021-3177) and the National Institutes of Health [K23MH120476].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

First and foremost, we thank the infants, parents, and families who participated in this research study. We also thank Eilis McLaughlin, Brady Floyd, Anastasia Javoris, and Aaron Beckner for their significant contributions to data collection and processing; Jillian Gulla and Adam Wenzel for their contributions to data collection; and Frederick Shic, Jason Babckock, and Samuel Harding for their helpful comments and suggestions on our analytic approach.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dcn.2023.101299.

References

Abney, D.H., Suanda, S.H., Smith, L.B., Yu, C., 2020. What are the building blocks of parent–infant coordinated attention in free-flowing interaction? Infancy 25 (6), 871–887. https://doi.org/10.1111/infa.12365.

Adamson, L.B., Bakeman, R., Suma, K., Robins, D.L., 2019. An expanded view of joint attention: skill, engagement, and language in typical development and autism. Child Dev. 90 (1), e1–e18. https://doi.org/10.1111/cdev.12973.

Adolph, K.E., Hoch, J.E., 2019. Motor development: embodied, embedded, enculturated, and enabling. Annu. Rev. Psychol. 70 (1), 141–164. https://doi.org/10.1146/annurev-psych-010418-102836.

Bakeman, R., Adamson, L.B., 1984. Coordinating attention to people and objects in mother-infant and peer-infant interaction. Child Dev. 55 (4), 1278. https://doi.org/10.2307/1129997.

Bambach, S., Franchak, J., Crandall, D., Yu, C., 2014. Detecting hands in children's egocentric views to understand embodied attention during social interaction. Proceedings of the Annual Meeting of the Cognitive Science Society 134–139.

Bambach, S., Smith, L.B., Crandall, D.J., Yu, C. (2017). Objects in the center: How the infant's body constrains infant scenes. *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics, ICDL-EpiRob* 2016, 132–137. ⟨https://doi.org/10.1109/DEVLRN.2016.7846804⟩.

Bradshaw, J., Shic, F., Holden, A.N., Horowitz, E.J., Barrett, A.C., German, T.C., Vernon, T.W., 2019. The use of eye tracking as a biomarker of treatment outcome in a pilot randomized clinical trial for young children with autism. Autism Res. 12 (5), 779–793. https://doi.org/10.1002/AUR.2093.

Bradshaw, J., Klin, A., Evans, L., Klaiman, C., Saulnier, C., McCracken, C., 2020. Development of attention from birth to 5 months in infants at risk for autism spectrum disorder. Dev. Psychopathol. 32 (2), 491–501. https://doi.org/10.1017/S0954579419000233.

Bradshaw, J., Schwichtenberg, A.J., Iverson, J.M., 2022. Capturing the complexity of autism: Applying a developmental cascades framework. Child Dev. Perspect. 16 (1), 18–26. https://doi.org/10.1111/CDEP.12439.

Bradshaw, J., Shi, D., Federico, A., Klaiman, C., Saulnier, C., 2023. The pull-to-sit task: examining infant postural development in autism spectrum disorder. J. Pediatr. 253, 225–231.e2. https://doi.org/10.1016/J.JPEDS.2022.09.047.

Carpenter, M., Tomasello, M., 2000. Joint attention, cultural learning, and language acquisition: Implications for children with autism. In: Wetherby, A., Prizant, B. (Eds.), Communication and language intervention series; Vol. 9. Autism spectrum disorders: A transactional developmental perspective. Paul H Brookes Publishing, pp. 31–54.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., Moore, C., 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monogr. Soc. Res. Child Dev. 63 (4), i. https://doi.org/10.2307/1166214.

Chawarska, K., Macari, S., Shic, F., 2013. Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. Biol. Psychiatry 74 (3), 195–203. https://doi.org/10.1016/j.biopsych.2012.11.022.

Colombo, J., 2001. The development of visual attention in infancy. Annu. Rev. Psychol. 52 (1), 337–367. https://doi.org/10.1146/annurev.psych.52.1.337.

Datavyu Team, 2014. Datavyu: A Video Coding Tool. Databrary Project. New York University.

Deng, J., Guo, J., Ververas, E., Kotsia, I., Zafeiriou, S., 2020. RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild 5203–5212.

Fausey, C.M., Jayaraman, S., Smith, L.B., 2016. From faces to hands: changing visual input in the first two years. Cognition 152, 101–107. https://doi.org/10.1016/J.COGNITION.2016.03.005.

Foulsham, T., Walker, E., Kingstone, A., 2011. The where, what and when of gaze allocation in the lab and the natural environment. Vis. Res. 51 (17), 1920–1931. https://doi.org/10.1016/J.VISRES.2011.07.002.

Franchak, J.M., 2020. The ecology of infants' perceptual-motor exploration. Curr. Opin. Psychol. 32, 110–114. https://doi.org/10.1016/J.COPSYC.2019.06.035.

Franchak, J.M., Yu, C., 2022. Beyond screen time: using head-mounted eye tracking to study natural behavior. Adv. Child Dev. Behav. 62, 61–91. https://doi.org/10.1016/BS.ACDB.2021.11.001.

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. https://doi.org/10.1111/j.1467-8624.2011.01670.x.

Franchak, J.M., Kretch, K.S., Adolph, K.E., 2018. See and be seen: infant-caregiver social looking during locomotor free play. Dev. Sci. 21 (4), e12626 https://doi.org/10.1111/desc.12626.

Hart, L.G., Larson, E.H., Lishner, D.M., 2005. Rural definitions for health policy and research. Am. J. Public Health 95 (7), 1149. https://doi.org/10.2105/AJPH.2004.042432.

Hessels, R.S., Hooge, I.T.C., 2019. Eye tracking in developmental cognitive neuroscience – the good, the bad and the ugly. Dev. Cogn. Neurosci. 40, 100710 https://doi.org/10.1016/J.DCN.2019.100710.

Hessels, R.S., Andersson, R., Hooge, I.T.C., Nyström, M., Kemner, C., 2015. Consequences of eye color, positioning, and head movement for eye-tracking data quality in infant research. Infancy 20 (6), 601–633. https://doi.org/10.1111/INFA.12093.

Hessels, R.S., Benjamins, J.S., Niehorster, D.C., van Doorn, A.J., Koenderink, J.J., Holleman, G.A., de Kloe, Y.J.R., Valtakari, N.V., van Hal, S., Hooge, I.T.C., 2022. Eye contact avoidance in crowds: a large wearable eye-tracking study. Atten., Percept., Psychophys. 84 (8), 2623–2640. https://doi.org/10.3758/S13414-022-02541-Z/FIGURES/6.

Hoch, J.E., Rachwani, J., Adolph, K.E., 2020. Where infants go: real-time dynamics of locomotor exploration in crawling and walking infants. Child Dev. 91 (3), 1001–1020. https://doi.org/10.1111/cdev.13250.

Holmqvist, K., Nyström, M., Mulvey, F., 2012. Eye tracker data quality: what it is and how to measure it. Eye Track. Res. Appl. Symp. . (ETRA 45–52. https://doi.org/10.1145/2168556.2168563.

Holmqvist, K., Örbom, S.L., Hooge, I.T.C., Niehorster, D.C., Alexander, R.G., Andersson, R., Benjamins, J.S., Blignaut, P., Brouwer, A.M., Chuang, L.L., Dalrymple, K.A., Drieghe, D., Dunn, M.J., Ettinger, U., Fiedler, S., Foulsham, T., van der Geest, J.N., Hansen, D.W., Hutton, S.B., Hessels, R.S., 2022. Eye tracking: empirical foundations for a minimal reporting guideline. Behav. Res. Methods 2022 55:1 55 (1), 364–416. https://doi.org/10.3758/S13428-021-01762-8.

Hooge, I.T.C., Niehorster, D.C., Hessels, R.S., Benjamins, J.S., Nyström, M., 2022. How robust are wearable eye trackers to slow and fast head and body movements? Behav. Res. Methods 1, 1–15. https://doi.org/10.3758/S13428-022-02010-3/FIGURES/7.

Jayaraman, S., Smith, L.B., 2019. Faces in early visual environments are persistent not just frequent. Vis. Res. 157, 213–221. https://doi.org/10.1016/J.VISRES.2018.05.005.

Kretch, K.S., Franchak, J.M., Adolph, K.E., 2014. Crawling and walking infants see the world differently. Child Dev. 85 (4), 1503–1518. https://doi.org/10.1111/cdev.12206.

Leezenbaum, N.B., Iverson, J.M., 2019. Trajectories of posture development in infants with and without familial risk for autism spectrum disorder. J. Autism Dev. Disord. 49 (8), 3257–3277. https://doi.org/10.1007/s10803-019-04048-3.

Macari, S., Milgramm, A., Reed, J., Shic, F., Powell, K.K., Macris, D., Chawarska, K., 2020. Context-specific dyadic attention vulnerabilities during the first year in infants later developing autism spectrum disorder. J. Am. Acad. Child Adolesc. Psychiatry. https://doi.org/10.1016/j.jaac.2019.12.012.

Messinger, D., Young, G.S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., Landa, R.J., Charman, T., Stone, W.L., Constantino, J.N., Hutman, T., Carver, L.J., Bryson, S., Iverson, J.M., Strauss, M.S., Rogers, S.J., Sigman, M., 2013. Beyond autism: a baby siblings research consortium study of high-risk children at three years of age. J. Am. Acad. Child Adolesc. Psychiatry 52 (3), 300–308.e1. https://doi.org/10.1016/j.jaac.2012.12.011.

Mundy, P., Burnette, C., 2005. Joint attention and neurodevelopmental models of autism. In: Volkmar, F.R., Paul, R., Klin, A.J., Cohen, D. (Eds.), Handbook of autism and pervasive developmental disorders, 3rd ed..,. Wiley, pp. 650–681.

Mundy, P., Sullivan, L., Mastergeorge, A.M., 2009. A parallel and distributed-processing model of joint attention, social cognition and autism. Autism Res. 2 (1), 2–21. https://doi.org/10.1002/aur.61.

Niehorster, D.C., Santini, T., Hessels, R.S., Hooge, I.T.C., Kasneci, E., Nyström, M., 2020. The impact of slippage on the data quality of head-worn eye trackers. Behav. Res. Methods 52 (3), 1140–1160. https://doi.org/10.3758/S13428-019-01307-0/FIGURES/12.

Ozonoff, S., Young, G.S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L.J., Constantino, J.N., Dobkins, K., Hutman, T., Iverson, J.M., Landa, R., Rogers, S.J., Sigman, M., Stone, W.L., 2011. Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. Pediatrics 128 (3), e488–e495. https://doi.org/10.1542/peds.2010-2825.

Pérez-Edgar, K., … L. M.-C. D. in, & 2020, undefined. (2020). Navigating through the experienced environment: Insights from mobile eye tracking. Journals.Sagepub.Com, 29(3), 286–292. https://doi.org/10.1177/0963721420915880.

Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C.C., Malige, A., 2016. Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity. Biol. Psychiatry 79 (8), 657–666. https://doi.org/10.1016/j.biopsych.2015.03.032.

Sameroff, A., 2009. The transactional model. The Transactional Model of Development: How Children and Contexts Shape Each Other 3–21. https://doi.org/10.1037/11877-001.

Shic, F., Macari, S., Chawarska, K., 2014. Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. Biol. Psychiatry 75 (3), 231–237. https://doi.org/10.1016/j.biopsych.2013.07.009.

Shic, F., Wang, Q., Macari, S.L., Chawarska, K., 2020. The role of limited salience of speech in selective attention to faces in toddlers with autism spectrum disorders. J. Child Psychol. Psychiatry Allied Discip. 61 (4), 459–469. https://doi.org/10.1111/jcpp.13118.

Slone, L.K., Abney, D.H., Borjon, J.I., Chen, C.H., Franchak, J.M., Pearcy, D., Suarez-Rivera, C., Xu, T.L., Zhang, Y., Smith, L.B., Yu, C., 2018. Gaze in action: head-mounted eye tracking of children's dynamic visual attention during naturalistic behavior. JoVE (J. Vis. Exp.) 2018 (141), e58496. https://doi.org/10.3791/58496.

Smith, L.B., Jayaraman, S., Clerkin, E., Yu, C., 2018. The developing infant creates a curriculum for statistical learning. Trends Cogn. Sci. 22 (4), 325–336.

Smith, L.B., 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. https://doi.org/10.1080/15248372.2014.933430.

Suarez-Rivera, C., Smith, L.B., Yu, C., 2019. Multimodal parent behaviors within joint attention support sustained attention in infants. Dev. Psychol. 55 (1), 96–109. https://doi.org/10.1037/DEV0000628.

von Hofsten, C., Uhlig, H., Adell, M., Kochukhova, O., 2009. How children with autism look at events. Res. Autism Spectr. Disord. 3 (2), 556–569. https://doi.org/10.1016/J.RASD.2008.12.003.

Wass, S.V., Forssman, L., Leppänen, J., 2014. Robustness and precision: how data quality may influence key dependent variables in infant eye-tracker analyses. Infancy 19 (5), 427–460. https://doi.org/10.1111/INFA.12055.

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. https://doi.org/10.1371/journal.pone.0079659.

Yurkovic, J.R., Lisandrelli, G., Shaffer, R.C., Dominick, K.C., Pedapati, E.V., Erickson, C. A., Kennedy, D.P., Yu, C., 2021. Using head-mounted eye tracking to examine visual and manual exploration during naturalistic toy play in children with and without autism spectrum disorder. Sci. Rep. 11 (1), 3578 https://doi.org/10.1038/s41598-021-81102-0.

Yurkovic-Harding, J., Lisandrelli, G., Shaffer, R.C., Dominick, K.C., Pedapati, E.V., Erickson, C.A., Yu, C., Kennedy, D.P., 2022. Children with ASD establish joint attention during free-flowing toy play without face looks. Curr. Biol. 32 (12), 2739–2746.e4. https://doi.org/10.1016/J.CUB.2022.04.044.