Head and eyes: Looking behavior in 12- to 24-month old infants

Jeremy I. Borjon[1*], Drew H. Abney[2], Chen Yu[1,3], Linda B. Smith[1,4]

[1]Department of Psychological and Brain Sciences, Indiana University, Bloomington
[2]Department of Psychology, University of Georgia, Athens
[3] Department of Psychology, University of Texas, Austin
[4] School of Psychology, University of East Anglia

*Address correspondence to:
Jeremy I. Borjon
Department of Psychological and Brain Sciences
Indiana University
1101 E. 10th St.
Bloomington IN, 47405
jborjon@iu.edu

Abstract

This study demonstrates evidence for a foundational process underlying active vision in older infants during object play. Using head-mounted eye-tracking and motion capture, looks to an object are shown to be tightly linked to and synchronous with a stilled head, regardless of the duration of gaze for infants 12 – 24 months of age. Despite being a developmental period of rapid and marked changes in motor abilities, the dynamic coordination of head stabilization and sustained gaze to a visual target is developmentally invariant during the examined age range. The findings indicate that looking with an aligned head and eyes is a fundamental property of human vision and highlights the importance of studying looking behavior in freely moving perceivers in everyday contexts, opening new questions about the role of body movement in both typical and atypical development of visual attention.

Keywords: infant vision, active vision, attention, head-eye alignment, motor development, sensorimotor coordination

Head and eyes: Looking behavior in 12- to 24-month old infants

Introduction

Gaze is directed to select targets and is maintained on selected targets to gather relevant information. Thus, looking behavior across the lifespan is intensely studied (e.g. Aslin, 2007; Ballard & Hayhoe, 2009; Brams et al., 2019; Oakes, 2015). However, there is still a great deal not known about looking behavior in freely moving individuals in the purposeful tasks of everyday life (Jovancevic-Misic & Hayhoe, 2009; Lappi, 2016; Schmitow, Stenberg, Billard, & Hofsten, 2013; Tatler, Hayhoe, Land, & Ballard, 2011). This lack of knowledge poses a significant barrier to research on a current topic of interest in developmental science: the ability of newly autonomous toddlers to maintain gaze on a single object in the context of natural play is increasingly implicated as both a biomarker and training ground for later development of the executive functions mediated by the prefrontal cortex (Brandes-Aitken, Braren, Swingler, Voegtline, & Blair, 2019; Fisher, 2019; Rosen, Amso, & McLaughlin, 2019; Werchan & Amso, 2017; Yu & Smith, 2016). Because active looking involves both head and eyes, we used head-mounted eye-tracking and motion-capture sensors to quantify eye and head movements in 12- to 24-month old toddlers as they actively interacted with and directed gaze to objects during play. The main finding is that the duration of gaze to an object, be it brief or sustained, was *synchronous with* decreased head movement. The findings open new questions about the role of body movement in both typical and atypical development of visual attention.

Directing gaze to a target selectively supports visual processing of that target over other information because the retinal area around the gaze point captures a higher resolution image than does the periphery (Dowling, 1987; Lee, 1996; May, 2006; Meister & Tessier-Lavigne, 2013). Thus, when a perceiver sustains gaze on a target, they optimize the extraction of visual information from the target relative to the periphery. The eyes, however, do not operate in isolation. Eyes are located in a head, which is on a body, all of which can move independently. Therefore, stabilizing gaze on a target depends on coordinating eye and head movements (Crawford, Henriques, & Medendorp, 2011; Kretch & Adolph, 2015; Nakagawa & Sukigara, 2013; Regal, Ashmead, & Salapatek, 1983). Active purposeful vision, from making a sandwich (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003) to putting one toy on top of another (Yu & Smith, 2012), often includes large head movements which can be both goal-directed and compensatory to actions such as reaching or posture change (Bertenthal & Von Hofsten, 1998; von Hofsten &

Rosander, 2018; Von Hofsten, Vishton, Spelke, Feng, & Rosander, 1998). The central goal of the present study was to quantify head and eye coordination and sustained gaze in freely moving infants 12 to 24 months of age. This is the developmental period during which active object play strongly predicts long-term outcomes in executive function and self-regulation (Rosen et al., 2019; Werchan & Amso, 2017).

It is well-known that freely moving perceivers, both adults and infants, are strongly biased to direct their gaze towards targets with their eyes and head aligned, turning both the eyes and head in the same direction to the target (Bambach, Crandall, Smith, & Yu, 2018; Bambach, Smith, Crandall, & Yu, 2017; Foulsham, Walker, & Kingstone, 2011; Kretch & Adolph, 2015; Solman, Foulsham, & Kingstone, 2017; Tatler et al., 2011; van Renswoude, van den Berg, Raijmakers, & Visser, 2019; Yoshida & Smith, 2008; Yu & Smith, 2012). When perceivers shift their gaze to a new target, the eyes and head are misaligned for typically less than 500 milliseconds, as either the eyes shift first, followed by the head (typical in adults, Corneil, 2012; Doshi & Trivedi, 2012; Nakashima & Shioiri, 2014), or the head shifts first followed by eyes (frequent in infants and children, Bloch & Carchon, 1992; Funk & Anderson, 1977; Nakagawa & Sukigara, 2013; Regal et al., 1983; Schmitow, Stenberg, Billard, & Hofsten, 2013; Tronick & Clanton, 1971). Notably for natural movements in adults, if planning is possible, the head will frequently move ahead of the eyes (Hayhoe, 2009). Once the shift is accomplished, the extant evidence suggests that purposeful looks occur with eyes and head pointed roughly in the same direction, the perhaps energetic "resting state" for gaze (Seemiller, Port, & Candy, 2018; van Renswoude et al., 2019). This bias is also evident in the spatial distribution of gaze captured by head-mounted eye trackers which find gaze to be pervasively centered in the head-centered field of view (Bambach et al., 2017; Y. Li et al., 2013).

The present study focuses on eye-head coordination and gaze duration in 12- to 24 month old infants because sustained gaze on an object during this period strongly predicts cognitive development more generally; from individual differences in visual attention (Lansink & Richards, 1997; Richards & Casey, 1992; Ruff, 1986), to differences in self-regulation and self-control (Kochanska, Murray, & Harlan, 2000; Reck & Hund, 2011; Ruff, 1986), as well as language development (Welsh, Nix, Blair, Bierman, & Nelson, 2010; Yu, Suanda, & Smith, 2019) and later school achievement (Kannass, Oakes, & Shaddy, 2006; Ruff & Lawson, 1990). Visual attention in active and unrestrained toddlers has been characterized as recruiting the whole body. Toddlers often both move their body closer to objects and hold objects close to their body while looking (Richards & Cameron, 1989; Richards & Casey, 1992; Ruff & Lawson, 1990; Ruff & Rothbart, 1996; Yu & Smith, 2012). During this period of rapid physical growth, infants are also just beginning to control to their bodies and have well documented difficulties in stabilizing their head (Bertenthal & Von Hofsten, 1998; Flatters et al., 2014; Ledebt & Bril, 2000) especially during large body movements, the common context for toddler everyday vision (Adolph, Vereijken, & Shrout, 2003; Bertenthal & Von Hofsten, 1998; Claxton, Haddad, Ponto, Ryu, & Newcomer, 2013; Claxton, Melzer, Ryu, & Haddad, 2012; Claxton, Strasser, Leung, Ryu, & O'Brien, 2014; Flatters et al., 2014; von Hofsten & Rosander, 2018; Von Hofsten et al., 1998). Toddlers, like adult perceivers, primarily direct gaze to the center of the head-center field of view (Bambach et al., 2017). However, the field lacks precise quantification of the relations among eye-head coordination, head stabilization, and gaze to an object in freely moving toddlers during active engagement with objects.

The starting hypothesis is that head *stabilization* is strongly associated with maintained gaze to an object. This hypothesis is suggested by classic studies on focused attention in late infancy (Ruff & Capozzoli, 2003; Ruff, Capozzoli, & Weissberg, 1998; Ruff & Rothbart, 1996). These studies found that long looks by toddlers during object play were associated with a stilled head. In these earlier studies, look durations and head movements were measured by human coders. Here, wearable sensors are used to provide more precise temporal spatial measures of the hypothesized decrease in head movements during gaze to an object. The interest in head movements and gaze duration is also motivated by research on atypically developing children that has shown an association between large head movements during a purposeful task and poor attentional control (Klingberg, Forssberg, & Westerberg, 2002; F. Li et al., 2016; Teicher, Ito, Glod, & Barber, 1996). Together, these observations suggest that *maintaining* gaze to a target is accompanied by an aligned head and eyes and decreased head movements *during the look* to an object.

Methods

Participants

A total of 44 infants (22 male) participated in multiple testing sessions when they were 12, 15, 18, 21 or 24 months of age. Infants possessed no reported visual-acuity or binocular-vision abnormalities. This period of development is under study due to the focus of recent work on sustained attention and its role as a predictor of later developmental outcomes (Brandes-Aitken et al., 2019; Reck & Hund, 2011; Yu et al., 2019). There are no specific *a priori* developmental hypotheses but the broad age range spans a period of marked changes in general sensory-motor skills (Adolph & Franchak, 2017; Libertus & Hauf, 2017; McGraw, 2004; Soska, Robinson, & Adolph, 2015) and is also characterized by the overall shortening of look durations to objects (Bronson, 1991; Colombo, Mitchell, Coldren, & Freeseman, 1991; Helo, Rämä, Pannasch, & Meary, 2016; Wass & Smith, 2014). Both factors could be relevant to the role of head and eye coordination in sustained gaze to an object. Each infant participated at different ages for on average for 2.49 sessions (SD = 1.16) yielding a total of 107 sessions distributed across the 5 ages at testing. Table 1 shows the data for the sessions contributed by each participant. The sample of infants was broadly representative of Monroe County, Indiana (84 % European American, 5% African American, 5% Asian American, 2% Latino, 4% Other) and consisted of predominantly working- and middle-class families. All research was approved by the Human Subjects and Institutional Review Board at Indiana University (Protocol #0808000094) and adhered to the tenets of the Declaration of Helsinki. Caregivers volunteering their infants for the study were fully informed of the study procedures and completed written informed consent and permission forms in advance of the study.

Subject #	**12 Month**	**15 Month**	**18 Month**	**21 Month**	**24 Month**
01					x
02		x			
03		x	x		
04			x	x	x
05				x	
06	x	x	x	x	x
07	x	x		x	x
08	x	x			x
09	x	x		x	x
10	x		x		x
11			x	x	x
12				x	x
13		x		x	x
14		x		x	
15	x	x		x	
16				x	x
17			x	x	x
18	x	x			
19	x	x	x	x	x
20				x	
21			x	x	x
22					
23			x	x	x
24	x	x	x	x	x
25	x		x	x	x
26	x	x			
27					x
28	x		x	x	x
29		x	x	x	x
30	x	x			
31	x				
32	x	x			
33	x				
34	x	x			
35	x	x			
36	x				
37	x		x	x	x
38		x			
39			x	x	x
40				x	x
41			x		x
42			x		
43			x		x
44			x		x
Total	**20**	**19**	**19**	**23**	**26**

Table 1. *Breakdown of subject participation for each age level.* Age at which subject was tested with 'x' indicating when tested.

Stimuli

There were 30 novel objects constructed in the laboratory and pilot-tested to be interesting and engaging to infants. Each object consisted of multiple parts (some moveable) and were of similar size (~280 cm^3) and weight (~95 g). A unique subset of 6 objects were chosen for use in each session and were organized into two sets of 3. Each object in the set of 3 had a unique uniform color (red, blue, green). At each age level, repeating participants received a different set of toys so that no child experienced a repeated set of toys during their participation in the study.

Experimental setup

Infants sat at a small table (61 cm × 91 cm × 64 cm) while their caregiver sat across the table from them (Figure 1). The infant was free to shift, lean and rotate the upper body, head, and to reach for objects in play on the tabletop. The infant wore a head-mounted eye tracker (Positive Science, LLC) designed for use with infants. The tracking system included two cameras: 1) an infrared camera mounted on the head and pointed to the right eye of the participant in order to record eye images and 2) a scene camera which captures the events from the participant's perspective. The scene camera's visual field has a diagonal of 108 degrees, providing a broad view to approximate the full visual field. The eye-tracking system recorded both the egocentric view video and eye-in-head position (x and y coordinates) in the captured scene at a sampling rate of 30 Hz. A wired motion capture sensor was affixed to the eye-tracker on the right temple of the infant's head (Polhemus Liberty, Polhemus). The motion-capture sensor collected rotational position data (roll, pitch, and yaw) at 60 Hz.

Figure 1. *Experimental setup.*

Placing the head gear and eye tracker calibration

Prior to entering the testing room, in the waiting area, a first experimenter desensitized the infant to touches to the head and hair by lightly touching the hair several times when the interest of the infant was directed to a toy. Both the caregiver and the infant entered the experimental room, and a second experimenter and the caregiver engaged the infant with toys not used in the experiment. The infant's head gear was placed while the infant was engaged with the toy. The first experimenter then adjusted the scene camera to ensure the scene camera captured the caregiver across the table and also the manual actions of the infant. The overall success rate for infant wearing of the sensors and calibration is over 70% (see Slone et al., 2018).

Instructions and procedure

Caregivers were told the goal of the experiment was to study how infants manually and visually explored novel objects and that they should encourage their infants to interact with the objects as naturally as possible. Each of the two sets of objects were played with twice for 1.5 min, resulting in 6 minutes of play data per session.

Data processing

During post-processing and before coding, the quality of the eye-tracking video for each infant was checked to ensure the quality of calibration at the end as well as the beginning of the session. If necessary, manual re-calibration was conducted by identifying moments in which the pupil and corneal reflection are accurately detected, and the eye is stably fixated on a clearly identifiable point in space in the scene image. These locations were chosen as re-calibration points. For a more in-depth discussion of the calibration and recording procedure, see Slone et al, 2018.

Looking. Within the study of vision, operational definitions of oculomotor functions such as saccades and fixations vary (Hessels, Niehorster, Nyström, Andersson, & Hooge, 2018) and there have been many debates about space-based versus object-based characterizations of attention (Chen, 2012; Logan, 1996; Scholl, 2001). The present study used object-based measures of attention as it is a better indicator of the duration of visual attention to a target when the targets are 3D objects in a 3-dimensional space and the perceiver is moving. In addition, gaze to an object (not a spatial location) has been the principal measure of sustained attention in studies of freely moving toddlers. Accordingly, *looks* to objects were measured in terms of continuous gaze that fell on an object.

The three regions-of-interest (ROIs) were defined in the head-camera videos as each of the three different and uniformly colored objects. ROI coding was done by highly trained coders who were responsible for coding many different projects and were naïve to the specific hypotheses or experimental questions of this study. Each of the three ROIs was coded separately. Frame-by-frame coders marked when the crosshair indicating gaze fell on a pixel of the object. This was a relatively easy task as each object was a unique color and the experimental room was white and both parent and child wore white smocks. Eye images were rendered via picture-in-picture superimposition at the upper-right corner of a scene frame, which allowed coders to constantly use the eye images as a reference to verify reliability of the crosshair indicating gaze direction in view. If coders detected that the eye-tracking software failed to detect the pupil correctly due to image quality or eye blinks, coders disregarded that frame. An unbroken look was defined as one that fell within a single object (Slone et al., 2018) and lasted a minimum of 15 frames, corresponding to 500 ms (Yu & Smith, 2012). This definition of a look thus includes both saccades and fixations. A second coder independently coded a randomly selected 10% of the frames (111,539 frames) with the inter-coder reliability ranging from 82% to 95% (Cohen's kappa = 0.81).

Analyses were conducted only on looks directed to one of the three objects in play. The head movements from the 44.45% of the play periods excluded from analyses were used for the determination of baseline rotational velocity of the head for each subject (described below).

Gaze clustering. Gaze refers to the eye-tracking data and need not be part of a look directed to an individual object. To measure the dispersion of frame-by-frame gaze across the head camera, the x-y coordinates from head-mounted eye-trackers were normalized for each individual by alignment to their centroid calculated from individual gaze points. Such an approach corrects for any off-center offset due to an imperfect positioning of the scene camera while preserving the original spread of the distribution (Bambach et al., 2018, 2017; Slone et al., 2018). The Euclidean distance from each x-y coordinate of eye position to the center of the scene-camera image, the origin, was then calculated in visual degrees.

To calculate the proportion of gaze points that fell within a radius of 10° and 20° from the center, the degrees per pixel in the head camera image was first determined. Frames from the head camera video were 480 pixels in height by 640 pixels with a diagonal of 108° in visual angle (Smith, Yu, Yoshida, & Fausey, 2015). Therefore, the head camera image is 86.4° in width and 64.8° in height. This results in 7.404 pixels per visual degree. For all analyses, the x-y coordinates of the head-mounted eye-tracker were converted into visual degrees by dividing the normalized x-y coordinates by 7.404.

For some analyses, looks (continuous gaze to an object) were categorized into two classes by duration: short (shorter than 3 seconds) or long (equal to or longer than 3 seconds in duration) as explained in the results section. Multivariate kernel density estimates of the normalized gaze distributions for these categorized long and short looks were independently calculated for each age and each look type using *kde2d* in Matlab and normalizing the resulting density by dividing all values by the maximum density value for that age level and look type. This resulted in a series of numbers between 0 and 1, separately calculated for each age level and look type.

Rotational velocity

Head stabilization in infants is typically measured in terms of the rotational coordinates of the head (Ledebt & Bril, 2000; Ledebt & Wiener-Vacher, 1996; Reisman & Anderson, 1989; Richards & Hunter, 1997; Rosander & Von Hofsten, 2000; Wiener-Vacher, 1996). Participants were equipped with a wired, magnetic motion capture marker (Polhemus Liberty, Polhemus) placed on the right temple of the head to record head rotation (roll, pitch, and yaw) and position (x, y, and z) during the task, at a rate of 60 Hz. The placement of the motion sensor was not consistent between subjects during the experiment due to toddler behavior. Experimenters needed to place the sensor and adjust it in one or two moves, or else the toddler will pull it off. Therefore, small variation was allowed in final placement. While the sensor is at the same location (right temple) the orientation of the sensor varies. Thus, translation is an unreliable measure and rotation was used. Rotational data were converted from millimeters to degrees by calculating the angular rotation between subsequent samples using the following formula in Matlab, where *rpy* represents an n-by-3 matrix where each row is a sample and each column is roll, pitch, or yaw in millimeters; *t* indicates time and *t+1* indicates the subsequent sample.

atan2d(norm(cross(rpy(t,:),rpy(t+1,:))),dot(rpy(t,:),rpy(t+1,:)))

As a measure of head stability, the rotational velocity was then calculated by taking the difference in angular rotation between subsequent samples divided by the change in time between samples. For each individual, rotational velocities exceeding the 99th percentile for that subject at that age level were replaced with NaNs in Matlab and excluded from further analysis.

As the rotational velocity data were captured and 60 Hz and the eye-tracking data was captured at 30 Hz, the rotational velocity was downsampled to 30 Hz to accommodate analyses between the sensors. Data were downsampled using cubic smoothing spline interpolation with *csaps* in Matlab. A smoothing parameter of 1 was used, resulting in minimal smoothing.

A baseline calculation of the rotational velocity of the head was made for each subject by randomly choosing portions of time when the infant was not looking to the objects in play and were exhibiting gaze that was centered within a 20° radius of the center of the head camera image. The median of this randomly selected baseline was taken, and a 95% bootstrapped confidence interval was calculated.

Correlation between head movements and eye movements

The vestibulo-ocular reflex (VOR) refers to rapid eye-movements of equal magnitude in the opposite direction counter small head movements that stabilize gaze on a target (Ornitz, Kaplan, & Westlake, 1985; Poletti, Aytekin, & Rucci, 2015; Rosander & Von Hofsten, 2000; Weissman, DiScenna, & Leigh, 1989). Although these compensatory movements are not the focus of the present study, they may be embedded at a finer temporal and spatial resolution than the head stabilizations and larger head movements of central interest. To measure the extent of the VOR within a look, repeated Spearman correlations were used to calculate the moment-to-moment correlation between the rotational yaw of the head and eye movements along the x-axis, horizontal gaze movements for every subject at every age level. Analyses were conducted on the 30 Hz eye-tracking data and motion-tracking data downsampled from 60 Hz to 30 Hz. An algorithm was constructed to calculate the Spearman correlation on the first 500 ms of data (15 data points). The *r* value and p-value were stored and the bin advanced one data sample and the correlation was estimated again. This was repeated until the end of the time series was reached.

To determine whether the correlation between the head and eyes exceeded chance, **a bootstrapped significance test** was conducted. For each of the 1,000 permutations, a number of random looks was chosen for each session equal in number and duration to the looks exhibited. Randomly selected looks were binned into 500 ms bins and stored. At the end of the simulation, the 2.5 and 97.5 percentiles of each bin were calculated.

For both of the above calculations, only *r* values which had p-values less than or equal to 0.01 were included in subsequent analyses.

Statistical approach

For all the analyses reported in this paper, the alpha level was set at 0.01 to minimize the likelihood of false positives. P-values for each conducted analysis were corrected for multiple comparisons using the Bonferroni-Holm correction (Holm, 1979). Using *lmefit* in Matlab, linear mixed effects (LME) models were constructed for each dependent measure. Dependent measures were: the proportion of time looking to objects, the number of looks to objects, the proportion of looks greater than or equal to 3 seconds in duration, the median distance of gaze to the center of the head-camera image, the proportion of gaze within a 10° radius of the center of the head camera image, the proportion of gaze within a 20° radius of the center of the head camera image, the proportion of fast head movements and the proportion of slow head movements. Subject identity and total number of trials, or trial number, were included as a random effect and infant age level was included as a fixed effect. The formula for these LME were as follows:

dependent variable ~ age + (1|subject identity) + (1|number of trials)

Main effects were determined by running an ANOVA on the LME.

Results

Age-related changes in look durations

During the 6-minute play sessions, children spent a median of 57.12% (SD 3.19%, min. 52.63%, max. 59.59%) of session time looking to one of the three play objects. An LME revealed no main effect of age on the proportion of time infants looked to objects ($F(4, 102) = 2.449$, p = n.s.). The total number of analyzed frames with gaze directed to an object was 647,698. The total number of looks to an object was 11,055 with the minimum look duration being 15 frames (500 ms). Table 2 provides the median and standard deviation of the proportion of time spent looking at objects and the number of looks to an object for each age level.

Although the proportion of time spent looking at the objects did not vary with the age of the infant, the number of looks did (LME, $F(4, 102) = 4.464$, $p < 0.003$), as older infants produced more short looks and younger infants more long looks, a well-known developmental change during this age period (Bronson, 1991; Colombo & Mitchell, 1990; Helo et al., 2016; Wass & Smith, 2014). Figure 2 shows the frequency distribution of look durations less than or equal to 10 seconds in duration, grouped

	Proportion looking time to objects	**Number of looks to objects**
Age level	**Median (SD)**	**Median (SD)**
12 months	0.596 (0.097)	80 (29.328)
15 months	0.571 (0.123)	88 (33.038)
18 months	0.526 (0.112)	117 (35.684)
21 months	0.595 (0.129)	116 (32.473)
24 months	0.538 (0.129)	118 (41.745)

Table 2. *Proportion of Looking Time to Objects*

into 500 ms bins. The data included in these graphs include 98.91% of the data analyzed below. Wilcoxon rank sum tests of subsequent age groups revealed look durations become more skewed (proportionally more short looks) with increasing age from 12-to-15-months ($Z = 5.289$, $p < 0.0001$), 15-to-18-months ($Z = 3.132$, $p < 0.004$), and 21-to-24-months of age ($Z = 5.078$, $p < 0.0001$). There was no difference in look duration from 18 to 21 months of age. ($Z = -0.592$, p = n.s.). Research on infant visual attention often divides looks in to short and long durations (Ruff & Lawson, 1990; Suarez-Rivera, Smith, & Yu, 2019; Wass, Clackson, et al., 2018; Wass, Noreika, et al., 2018; Yu & Smith, 2016; Yu et al., 2019; Yuan, Xu, Yu, & Smith, 2019) using the threshold of a look 3 sec or longer for defining long looks. This threshold is near the flexion point in the frequency distribution for all ages (Ruff & Lawson, 1990; Suarez-Rivera et al., 2019; Wass, Clackson, et al., 2018; Wass, Noreika, et al., 2018; Yu & Smith, 2016; Yu et al., 2019; Yuan et al., 2019). As shown in Figure 2B, proportional frequency of long looks, not just the overall durations, also decline with age (LME, $F(1, 104) = 11.224$, $p<0.0001$). Earlier studies based on human coding of look durations (Ruff & Lawson, 1990) were interpreted as showing steady increases in the frequency of long looks. The more precise measures of the present study suggest that this is not the case.

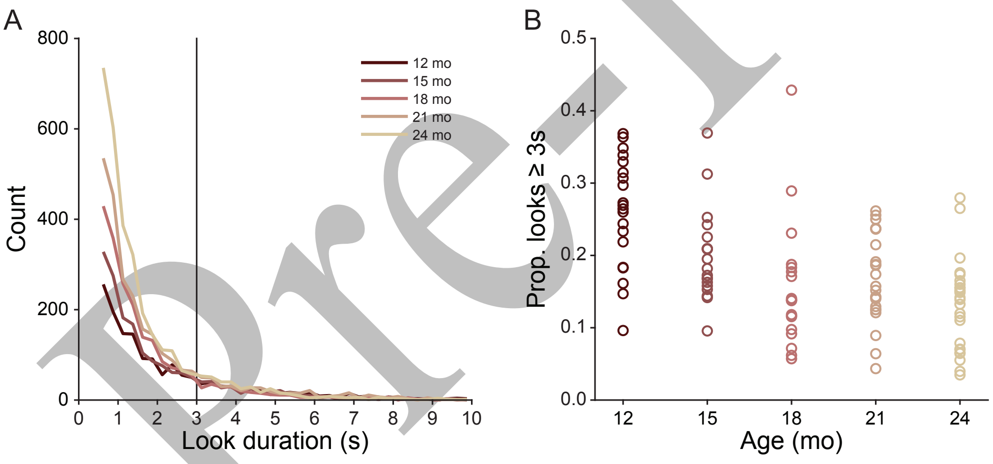

Figure 2. *Look duration decreases from 12-24 months*. (A) Histograms showing the distribution of look durations less than 10 seconds at each age level. Vertical black line indicates the 3 second cutoff for short and long looks. (B) Proportion of looks greater than or equal to 3 seconds in duration for each individual in each age group.

Gaze to the center of the head-centered field of view

Figure 3A and 3B show the distribution of frame-by-frame gaze to objects within the head-centered image for both short and long looks, respectively. As is apparent, both short and long looks are characterized by gaze to the center of the head-centered image. A linear mixed effects model revealed the median distance of gaze points to the center of the head-camera image did not vary as a function of age (*F*(4, 204) = 1.939, p = n.s.) or duration (*F*(1, 204) = 3.593, p = n.s.) and there was no interaction between these factors (*F*(1, 204) = 0.509, p = n.s). Supplementary Table 1 provides the median and standard deviation of the distance of gaze points to the center of the head-camera image for each age level for both long and short looks. Supplementary Table 2 provides the total number and proportion of data points that fell within a radius of 10° and 20° from the center for each age level. The proportion of gaze points within these two defined regions do not vary as function of age (LME, 10° radius *F*(4, 204) = 1.805, p = n.s; LME, 20° radius *F*(4, 204) = 2.974, p = n.s) nor look duration (LME, 10° radius *F*(1, 204) = 5.318, p = n.s.; 20° radius *F*(1, 204) = 2.803, p = n.s) and there were no interactions. Across ages, over 34% of gaze fell within 10° of center and more than 78% fell within the 20° radius, indicating the narrow and centered range of gaze to objects within the head camera image. Thus, the present findings show what is being consistently observed in studies of ego-centric vision and freely moving perceivers of all ages: a strong bias for looking with head and eyes generally pointed in the same direction.

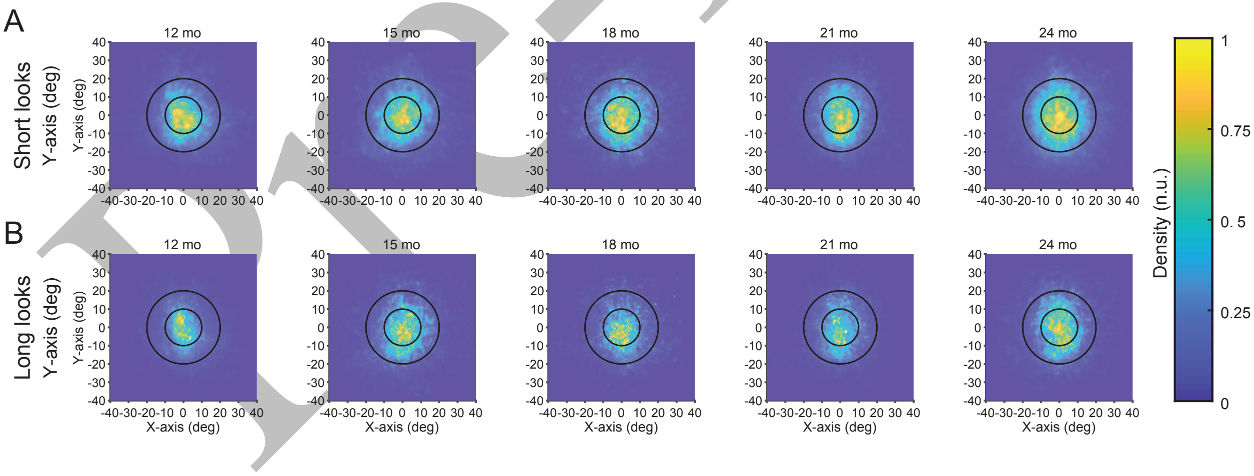

Figure 3. *A bias to look at objects with the head and eyes aligned.* (A, B) Multivariate kernel density estimates of the accumulated x-and y-coordinates of eye gaze using a head-mounted eye-tracker for gaze to novel objects where (A) looks are shorter than three seconds and (B) looks are equal to or longer than three seconds at 12, 15, 18, 21, and 24 months of age. Inner circles

encompass a 10^o radius from the center of the head camera image while outer circles encompass a 20^o radius from the center of the head camera image. Color indicates density of the distribution with more yellow colors indicating greater density.

Decreased head movement within a look

Maintained looks to an object within the center of the head-centered field of view imply the coordination of the head and eyes, and thus some limitation on head movements. Figure 4A and 4B show histograms of the head's rotational velocity when infants were looking to objects compared to a baseline where infants exhibited a centered head, with gaze within 20° of the center of the head camera image but were not looking to objects (method of calculating baseline defined in Methods). Histograms include rotational velocity up to 45 deg/s, which encompasses 99.99% of the observed data. Comparisons of the whole distributions yielded reliable differences between the rotational velocity of the head while looking at objects compared to baseline for each age level (Wilcoxon rank sum test, min. $Z = 31.821$, max. $Z = 54.004$, all Bonferroni-Holm corrected $p < 0.0001$). As shown in Figure 4C, the frequency of fast head movements, defined as movements exceeding the 75th percentile of rotational velocity observed in the dataset (5.283 deg/s) was proportionally greater when infants were exhibiting a centered gaze but not looking at objects than when they were looking at an object (LME, $F(1, 204) = 9.113$, $p< 0.003$) with no main effect of age ($F(4, 204) = 2.866$, p = n.s.) and there was no interaction between looking target and age ($F(4, 204) = 0.230$, p = n.s.). Additionally, as shown in Figure 4D, the frequency of slow head movements, defined as less than the 25th percentile of rotational velocity observed (1.147 deg/s), was comparable between conditions when infants were looking at objects than when they were not (LME, $F(1, 204) = 1.373$, p = n.s.) with no main effect of age ($F(2, 204) = 2.572$, p = n.s.) nor an interaction between looking target and age ($F(2, 204) = 0.071$, p = n.s.). Relative to comparably centered looks, looks to objects exhibited fewer fast head movements and a comparable amount of slower head movements across all ages.

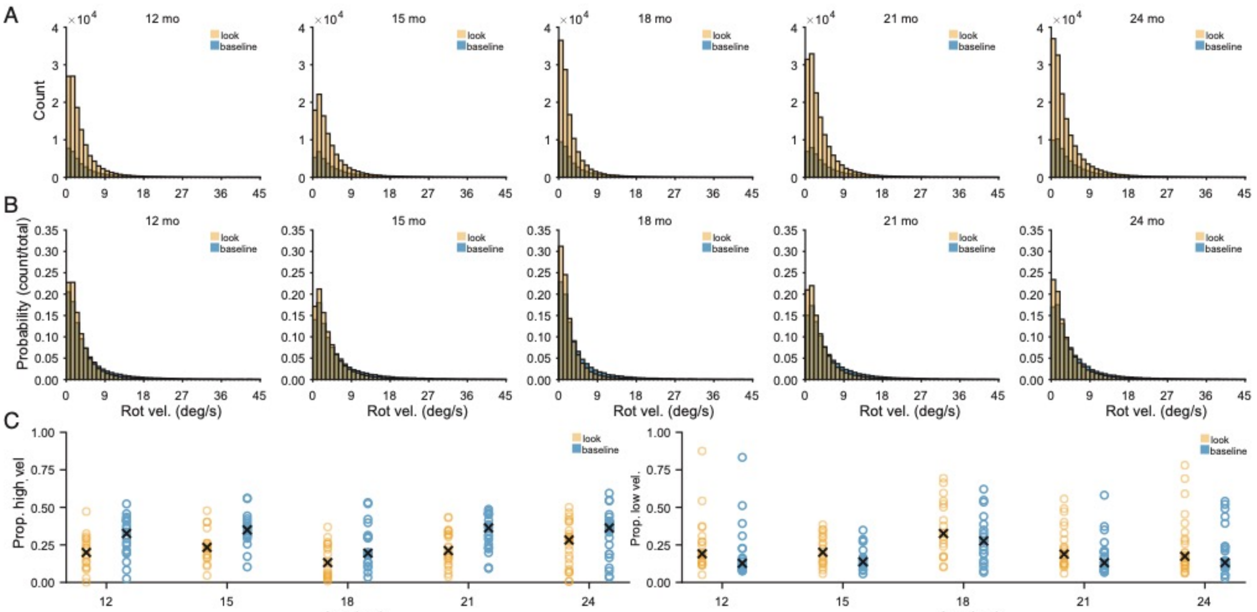

Figure 4. *Centered gaze lowers head movements.* (A, B) Histograms demonstrating the (A) count and (B) probability distribution of rotational velocity for looks to objects (amber) and looks to targets that were not one of the three play objects or the caregiver's face (blue) at each age level. (C) The proportion of the head's rotational velocity during a look which exceeds the 75th percentile of the rotational velocity in the observed dataset for every subject at each age level with 'X' representing the median proportion for that age level. (D) The proportion of the head's rotational velocity during a look which is slower than the 25th percentile of the rotational velocity in the observed dataset for every subject at each age level with 'X' representing the median proportion for that age level.

For all look durations, at all ages, head movements markedly decrease after the onset of a look. As the duration of each look is variable, look duration was binned into 500ms bins, up to a maximum of 4.5 seconds. Such a cutoff includes 99.95% of the observed data. Figure 5A shows the median rotational velocity of the head aligned to the onset and offset of a look. Velocity profiles begin 500 milliseconds before the onset of a look and end 500 milliseconds after the offset of a look. Supplementary Table 3 lists the number and proportion of looks in each bin for each age level. Across all look durations, looks begin with a brief change in velocity followed by a slowing of the head before the look ends with another brief change in velocity at the look's offset. Figure 5B shows the median rotational velocity of the head for the looks in each of the bins in Figure 5A compared to the baseline rotational velocity of the head (method of calculating baseline defined in Methods). Figure 5C shows the standard deviation of rotational velocity of the head for looks to objects and for the baseline. Baseline encompasses moments when infants were not looking to either of the 3 objects and their gaze was within 20 degrees of the center of the head camera image. Error bars for both the baseline and observed rotational velocity indicate the 95% bootstrapped confidence interval. The median rotational velocity of the head was lower during a look than baseline for every bin (Wilcoxon rank sum test, min. Z: -8.102 max. Z: -4.325, all Bonferroni-Holm corrected $p < 0.0001$). The standard deviation of the rotational velocity of the head was lower during a look than baseline for every bin (Wilcoxon rank sum test, min. Z: -33.643 max. Z: -7.940, all Bonferroni-Holm corrected $p < 0.0001$). In sum, infants between the ages of 12 and 24 months consistently and uniformly look to objects with their eyes and head aligned and they maintain alignment throughout the look by slowing their head movement and minimizing its variability.

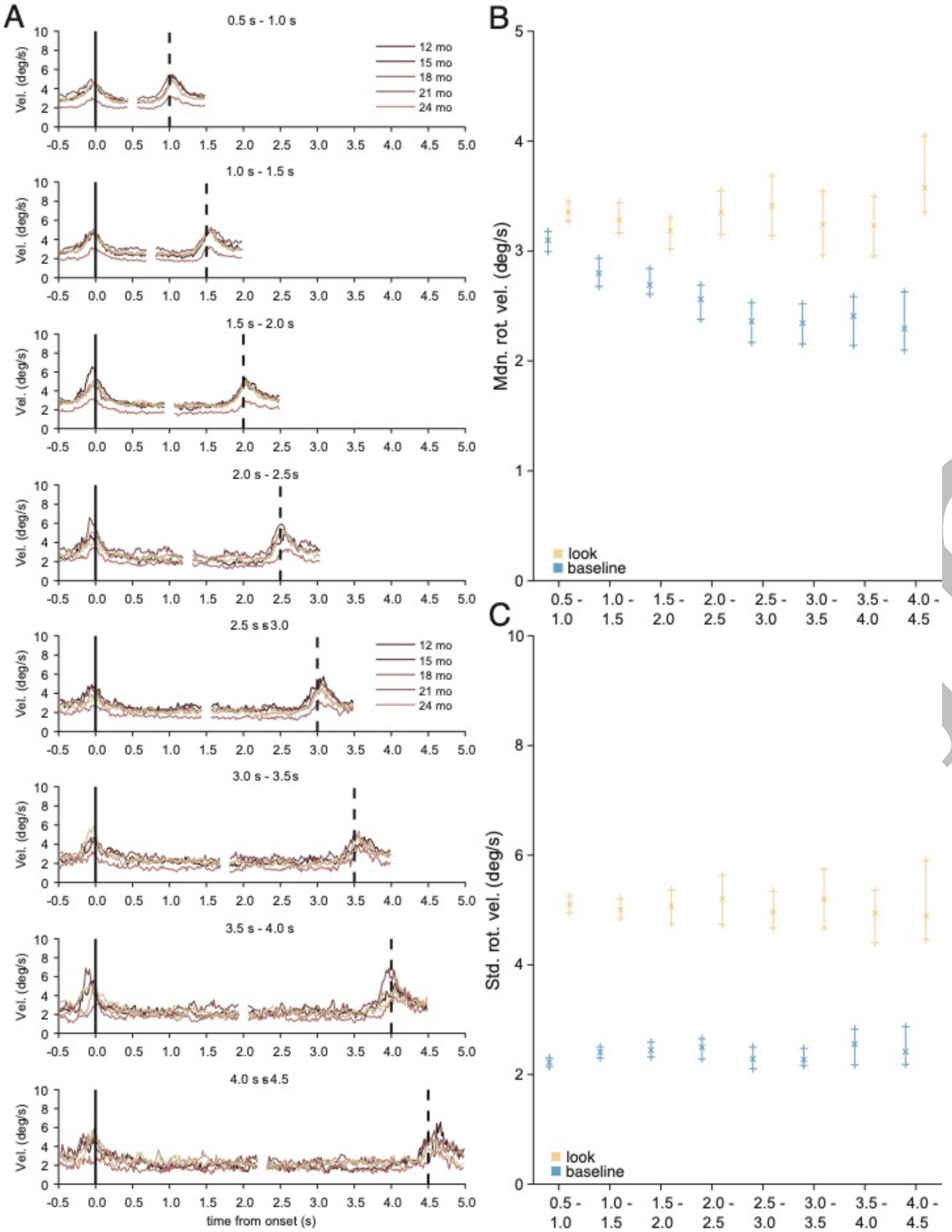

Figure 5. *Head stability is a function of look duration.* (A) Median rotational velocity traces aligned to the onset (vertical solid black line) and offset (vertical dotted black line) of a look with lighter colors indicating older groups. Traces begin 500 milliseconds before the onset and after the offset of the look. As instances of looks to an object vary in duration, the rotational velocity traces were binned into 500ms bins. (B, C) The (B) median and (C) standard deviation of the rotational velocity of the head for the binned looks (amber) with a calculated baseline (blue). Error bars indicate 95% bootstrapped confidence intervals.

The function of a stabilized head

During a look, infants make rapid head movements and minimize variability in head movements. This stabilization does not imply a complete stillness of the head. At no point in time did any subject at any age exhibit head movement that was 0 deg/s. As demonstrated in the distributions of rotational velocity in Figure 4A and B, head movement is continuous and there is no sharp divide between a still and not-still head. The decreased movement characteristic of looks to an object, however, are associated with the spatial location of gaze in the field of view. The head camera image was divided into bins 1 visual degree in height and width and the median rotational velocity of the head was calculated for each bin. Figure 6 shows the median rotational velocity of the head for each eye position in the head camera image across all ages. Gaze to the center of the head camera image coincides with a low head velocity while gaze to the periphery of the head camera image coincides with high velocity head movements. Thus, a slower-moving more stabilized head is strongly associated with the centering of gaze within a head-centered field of view.

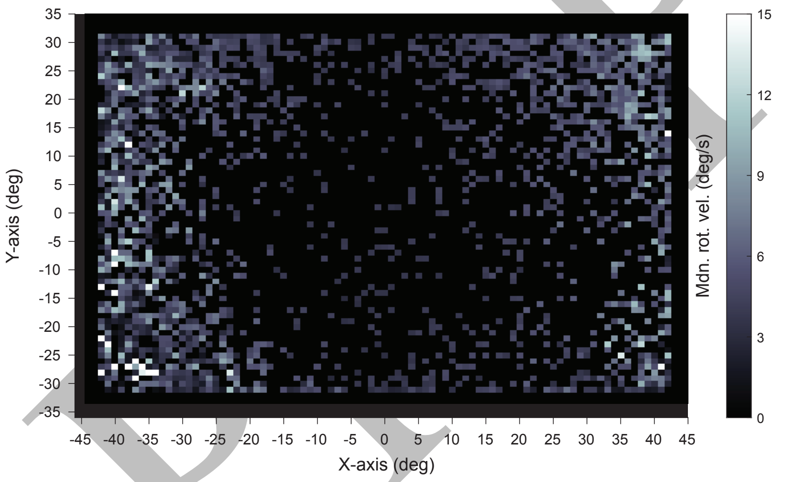

Figure 6. *Head velocity is slower when gaze is at the center of the head camera image.* Median head movement for each position of the eye in the head camera image calculated across all subjects at all ages. Brighter colors indicate a greater head velocity.

A measurable vestibulo-ocular reflex?

Does VOR provide a measurable contribution to the stabilization of the head and eyes during a look? If VOR is present and playing a role within looks to an object, there should be a negative correlation between the horizontal direction of head and eye movements. In an attempt to measure the possible contribution of this reflex in active naturalistic viewing, the yaw rotation of head movement was correlated to horizontal movement of the eyes during looks. Moments when the eyes and head moved in the same direction resulted in positive *r* values while moments when the eyes and head moved in the opposite direction resulted in negative *r* values. For every individual session, the moment-to-moment Spearman correlation between the head and eyes was calculated in successive, overlapping 500 ms bins during a maintained look to an object. As the analysis was conducted on the 30 Hz eye-tracking data and the downsampled motion capture data, 500 ms corresponds to 15 data points. The moment-to-moment *r* value was then calculated for every look, up to 4.5 seconds in duration, at every age and binned into 500 ms bin durations. For the duration of the look, while the correlation between the head and eyes changes over time, the extent of the correlation did not exceed chance. This lack of a measurable VOR in natural viewing is consistent with previous reports (Agtzidis, Startsev, & Dorr, 2019; Fuller, 1996; Meyer, O'Keefe, & Poort, 2020; Tatler, 2007; Tseng, Carmi, Cameron, Munoz, & Itti, 2009; Wang, Koch, Holmqvist, & Alexa, 2018) that were also unable to detect VOR in natural vision. Thus, the role of VOR in active natural viewing remains an unanswered question in need of further study and better measurement approaches.

Discussion

During play, toddlers look to objects with a stilled head but rapidly move their head to begin and end a look. Looking at an object with the head and eyes aligned appears to be the default mode for both short and long looks and does not vary with age during the period between the first and second birthday. For toddlers, gaze sustained on an object for any duration begins with the rapid movement of the head and eye to the object which is then is maintained by limited head movement with the centering of gaze within a head-centered frame of reference. The look ends with another rapid movement of the head and eyes. These findings contribute to the understanding of visual attention in freely moving perceivers in the context of their own self-generated purposeful behavior, which is the context of everyday vision. Within this context, a suite of behaviors appears to form a complex interdependent system of shifting both gaze and head in the same direction then maintaining gaze on an object with limited head movements such that the looked-to object is centered in a head-centered field of view.

The head and eyes can and do move independently: What, then, is the function of the observed strong coordination of the head and eye movement at the start of a look, the joint stabilization of eye and head direction to the attended object such that gaze is centered within the head-centered view, and the synchronous shift of both head and eyes to end a look? Both behavioral (Cicchini, Valsecchi, & De'Sperati, 2008; Corneil & Munoz, 2014; Khan et al., 2009) and neural (Gandhi & Katnani, 2011; Ignashchenkova, Dicke, Haarmeier, & Thier, 2004; Müller, Philiastides, & Newsome, 2005; Stryker & Schiller, 1975; Walton, Bechara, & Gandhi, 2007) evidence indicates that the networks that plan motor behaviors (Desimone & Duncan, 1995; Miller & Cohen, 2001; Miyake & Friedman, 2012) overlap with the networks that internally control the spatial direction of visual attention (Cicchini et al., 2008; Corneil & Munoz, 2014; Khan et al., 2009). Planning and executing the independent movement of different body parts – the head, eyes, and hands – requires the coordination of multiple spatial reference frames (Galati, Pelle, Berthoz, & Committeri, 2010; Lappi, 2016; Schlicht & Schrater, 2007). For example, in looking and reaching to an object, the actor must coordinate the reference frame for the eye by moving gaze from the current eye position to the target and for the hand by moving the hand from its current position, which is different from the eye, to the target. In freely-moving individuals, the reference frames for the eyes, head, torso, and hand must continuously be coordinated (Badde, Röder, & Heed, 2015; Bosco, Piserchia, & Fattori, 2017; Crollen et al., 2017; Crollen, Spruyt, Mahau, Bottini, & Collignon, 2019; Pouget, Deneve, & Duhamel, 2002; Tagliabue & McIntyre, 2014). Considerable research shows this coordination is difficult and imposes a measurable computational burden not just on action but also on visual attention with effects on the detection, discrimination and location of visual events. For example, in adults, the misalignment of the head and eyes destabilizes and disrupts gaze relative to the aligned head and eyes (Einhäuser et al., 2007; Flanders, Daghestani, & Berthoz, 1999; Thaler & Todd, 2009) and goal-directed bodily actions become less spatially precise when the head and eyes point in different directions.

Between their first and second birthday, toddlers are in the midst of mastering many new bodily movements and skills. Considerable research shows that toddlers decrease the degrees of freedom in frames of reference for body movements by limiting or aligning the movement of different body parts when initially walking, carrying objects, or bending over to pick up an object (Claxton et al., 2013, 2012, 2014; Smith & Thelen, 1996). Looking is a motor behavior. Just as toddlers planning and controlling of other actions benefits from synergistic movements, so may the spatially coordinated head and eyes support visual attention. Gaze to the midline of the head and body is the positional resting state and it may take more energy to maintain gaze in eccentric orbital positions, the eyes will naturally return to the center. But an aligned and stabilized head and eyes for the duration of a look to a target may also not just be easy but highly functional by limiting coordination and competition among spatial frames of reference (Einhäuser et al., 2007; Flanders et al., 1999; Thaler & Todd, 2009).

The brief changes in rotational velocity of the head at the onset and offset of a look have been described previously, albeit in very young infants, but also may a key role in toddler visual attention. Infants at 3-months of age exhibit rapid bursts of body movement preceding gaze shifts during screen-based viewing and these have been shown to facilitate ending a look to one target to shift to another (Robertson, Johnson, Masnick, & Weiss, 2007) as young infants have considerable difficulty in disengaging from an attended target. Young infants who exhibit less coordinated bursts in movement during screen-based viewings went on to develop deficits in attention (Friedman, Watamura, & Robertson, 2005). Indeed, similar disruptions in sensory-motor coordination are exhibited by premature infants (Berger, Harbourne, & Guallpa Lliguichuzhca, 2019) and infants with several developmental disorders (Hartman, Houwen, Scherder, & Visscher, 2010; Proudlock & Gottlob, 2007). Toddlers with more well-developed control of eye, head, and body movements may well use head movements to purposely break gaze, a hypothesis worthy of future study.

Toddlers' ability to maintain a look to an individual object during active object play strongly predicts later developments in executive function and self-regulation and has been proposed to be causally related to those developments (Brandes-Aitken et al., 2019; Fisher, 2019; Rosen et al., 2019; Werchan & Amso, 2017; Yu & Smith, 2016). The origins of individual differences in sustained attention has not been identified (see Rosen et al., 2019). The goal here was to determine the mechanics of the behaviors –both the head and eyes – that underlie continuous looks to an object as a first step to understanding potential sources of individual differences. Uncontrolled body movements and specifically head movements have been linked to poor attentional control in older children (Friedman et al., 2005; Hartman et al., 2010; Proudlock & Gottlob, 2007) suggesting the integrative hypothesis that disruptions in sensory-motor coordination of eyes and head lead to disrupted attentional abilities. For example, toddlers with autism spectrum disorders sometimes exhibit difficulties in maintaining the midline position of the head during active attentional tasks (Dawson et al., 2018; Martin et al., 2018), a bias strongly evident in typically developing toddlers (Bambach et al., 2018, 2017). Difficulties in early head and trunk control are also exhibited by children with Down syndrome (Cardoso, De Campos, Dos Santos, Santos, & Rocha, 2015; Rast & Harris, 2008), language delays (Vuijk, Hartman, Scherder, & Visscher, 2010), and other cognitive disorders (Visscher, Houwen, Scherder, Moolenaar, & Hartman, 2007). Many of these disorders occur with concomitant deficits in the control of visual attention.

In conclusion, the present study provides evidence on eye and head coordination in infant looking behavior during active self-generated interactions with objects, the context of children's everyday vision and visual learning. There is much that is not known about looking behavior in this context. The present results provide a first step by showing a tight coordination of head and eyes during toddlers' sustained looks to objects.

Acknowledgements

This work was supported by the National Science Foundation SBE Postdoctoral Research Fellowship under Grant No. 1908519 (JIB), by Indiana University's Emerging Area of Research Initiative (JIB, CY, LBS), the National Institutes of Health grant R01HD074601 and R01HD093792 (CY, LBS). The authors would like to thank Steven Elmlinger, Seth Foster, Charlene Ty, Melissa Hall, and Char Wozniak as well as the past and present members of the Cognitive Development Laboratory and the Computational Cognition and Learning Laboratory for managing the data collection and preparation. The authors would also like to thank Eric Seemiller and Lydia Hoffstaetter for their careful reading of this manuscript. The authors declare that no competing interests exist.

Author Contributions

Jeremy I. Borjon: Data curation, Formal analysis, Investigation, Writing—original draft, Writing—review and editing

Drew H. Abney: Validation, Writing—review and editing

Chen Yu: Conceptualization, Data collection, Funding acquisition, Validation, Writing—review and editing

Linda B. Smith: Conceptualization, Supervision, Funding acquisition, Validation, Writing—original draft, Writing—review and editing

References

Adolph, K. E., & Franchak, J. M. (2017). The development of motor behavior. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1–2).

Adolph, K. E., Vereijken, B., & Shrout, P. E. (2003). What Changes in Infant Walking and Why. *Child Development*, *74*(2), 475–497.

Agtzidis, I., Startsev, M., & Dorr, M. (2019). 360-degree Video Gaze Behaviour: A ground-truth data set and a classification algorithm for eye movements. *MM 2019 - Proceedings of the 27th ACM International Conference on Multimedia*, 1007–1015.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, Vol. 10, pp. 48–53.

Atkinson, J., & Braddick, O. (2012). Visual attention in the first years: Typical development and developmental disorders. *Developmental Medicine and Child Neurology*, Vol. 54, pp. 589–595.

Badde, S., Röder, B., & Heed, T. (2015). Flexibly weighted integration of tactile reference frames. *Neuropsychologia*, *70*, 367–374.

Bahill, A. T., Adler, D., & Stark, L. (1975). Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology*, *14*(6), 468–469.

Ballard, D. H., & Hayhoe, M. M. (2009). Modelling the role of task in the control of gaze. *Visual Cognition*, *17*(6–7), 1185–1204.

Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2018). Toddler-inspired visual object learning. *Advances in Neural Information Processing Systems*, *2018-Decem*, 1201–1210.

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.

Berger, S. E., Harbourne, R. T., & Guallpa Lliguichuzhca, C. L. (2019). Sit Still and Pay Attention! Trunk Movement and Attentional Resources in Infants with Typical and Delayed Development. *Physical and Occupational Therapy in Pediatrics*, *39*(1), 48–59.

Bertenthal, B., & Von Hofsten, C. (1998). Eye, head and trunk control: The foundation for manual development. *Neuroscience and Biobehavioral Reviews*, *22*(4), 515–520.

Bloch, H., & Carchon, I. (1992). On the onset of eye-head coordination in infants. *Behavioural Brain Research*, *49*(1), 85–90.

Bosco, A., Piserchia, V., & Fattori, P. (2017). Multiple coordinate systems and motor strategies for reaching movements when eye and hand are dissociated in depth and direction. *Frontiers in Human Neuroscience*, *11*.

Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Mark Williams, A., & Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. *Psychological Bulletin*, *145*(10), 980–1027.

Brandes-Aitken, A., Braren, S., Swingler, M., Voegtline, K., & Blair, C. (2019). Sustained attention in infancy: A foundation for the development of multiple aspects of self-regulation for children in poverty. *Journal of Experimental Child Psychology*, *184*, 192–209.

Bronson, G. W. (1991). Infant Differences in Rate of Visual Encoding. *Child Development*, *62*(1), 44–54.

Cardoso, A. C. D. N., De Campos, A. C., Dos Santos, M. M., Santos, D. C. C., & Rocha, N. A. C. F. (2015). Motor Performance of Children with Down Syndrome and Typical Development at 2 to 4 and 26 Months. *Pediatric Physical Therapy*, *27*(2), 135–141.

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, and Psychophysics*, *74*(5), 784–802.

Cicchini, G. M., Valsecchi, M., & De'Sperati, C. (2008). Head movements modulate visual responsiveness in the absence of gaze shifts. *NeuroReport*, *19*(8), 831–834.

Claxton, L. J., Haddad, J. M., Ponto, K., Ryu, J. H., & Newcomer, S. C. (2013). Newly Standing Infants Increase Postural Stability When Performing a Supra-Postural Task. *PLoS ONE*, *8*(8).

Claxton, L. J., Melzer, D. K., Ryu, J. H., & Haddad, J. M. (2012). The control of posture in newly standing infants is task dependent. *Journal of Experimental Child Psychology*, *113*(1), 159–165.

Claxton, L. J., Strasser, J. M., Leung, E. J., Ryu, J. H., & O'Brien, K. M. (2014). Sitting infants alter the magnitude and structure of postural sway when performing a manual goal-directed task. *Developmental Psychobiology*, *56*(6), 1416–1422.

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. In *Individual differences in infancy: Reliability, stability, prediction.* (pp. 193–227).

Colombo, J., Mitchell, D. W., Coldren, J. T., & Freeseman, L. J. (1991). Individual Differences in Infant Visual Attention: Are Short Lookers Faster Processors or Feature Processors?

Child Development, *62*(6), 1247–1257.

Corneil, B. D. (2012). Eye-head gaze shifts. *The Oxford Handbook of Eye Movements*, 303–322.

Corneil, B. D., & Munoz, D. P. (2014). Overt responses during covert orienting. *Neuron*, *82*(6), 1230–1243.

Crawford, J. D., Henriques, D. Y. P., & Medendorp, W. P. (2011). Three-Dimensional Transformations for Goal-Directed Action. *Annual Review of Neuroscience*, *34*(1), 309–331.

Crollen, V., Lazzouni, L., Rezk, M., Bellemare, A., Lepore, F., & Collignon, O. (2017). Visual experience shapes the neural networks remapping touch into external space. *Journal of Neuroscience*, *37*(42), 10097–10103.

Crollen, V., Spruyt, T., Mahau, P., Bottini, R., & Collignon, O. (2019). How visual experience and task context modulate the use of internal and external spatial coordinate for perception and action. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(3), 354–362.

Dawson, G., Campbell, K., Hashemi, J., Lippmann, S. J., Smith, V., Carpenter, K., … Sapiro, G. (2018). Atypical postural control can be detected via computer vision analysis in toddlers with autism spectrum disorder. *Scientific Reports*, *8*(1).

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Doshi, A., & Trivedi, M. M. (2012). Head and eye gaze dynamics during visual attention shifts in complex environments. *Journal of Vision*, *12*(2), 1–16.

Dowling, J. E. (1987). *The Retina: An Approachable Part of the Brain*. Cambridge, MA: Harvard University Press.

Einhäuser, W., Schumann, F., Bardins, S., Bartl, K., Böning, G., Schneider, E., & König, P. (2007). Human eye-head co-ordination in natural exploration. *Network: Computation in Neural Systems*, *18*(3), 267–297.

Fisher, A. V. (2019). Selective sustained attention: a developmental foundation for cognition. *Current Opinion in Psychology*, *29*, 248–253.

Flanders, M., Daghestani, L., & Berthoz, A. (1999). Reaching beyond reach. *Experimental Brain Research*, *126*(1), 19–30.

Flatters, I., Mushtaq, F., Hill, L. J. B., Rossiter, A., Jarrett-Peet, K., Culmer, P., … Mon-

Williams, M. (2014). Children's head movements and postural stability as a function of task. *Experimental Brain Research*, *232*(6), 1953–1970.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920–1931.

Friedman, A. H., Watamura, S. E., & Robertson, S. S. (2005). Movement-attention coupling in infancy and attention problems in childhood. *Developmental Medicine and Child Neurology*, *47*(10), 660–665.

Fuller, J. H. (1996). Eye position and target amplitude effects on human visual saccadic latencies. *Experimental Brain Research*, *109*(3), 457–466.

Funk, C. J., & Anderson, M. E. (1977). Saccadic eye movements and eye head coordination in children. *Perceptual and Motor Skills*, *44*(2), 599–610.

Galati, G., Pelle, G., Berthoz, A., & Committeri, G. (2010). Multiple reference frames used by the human brain for spatial perception and memory. *Experimental Brain Research*.

Gandhi, N. J., & Katnani, H. A. (2011). Motor functions of the superior colliculus. *Annual Review of Neuroscience*, *34*, 205–231.

Hartman, E., Houwen, S., Scherder, E., & Visscher, C. (2010). On the relationship between motor performance and executive functioning in children with intellectual disabilities. *Journal of Intellectual Disability Research*, *54*(5), 468–477.

Hayhoe, M. M. (2009). Visual Memory in Motor Planning and Action. In J. Brockmole (Ed.), *The Visual World in Memory* (pp. 117–139). Hove: Psychology Press.

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), 49–63.

Helo, A., Rämä, P., Pannasch, S., & Meary, D. (2016). Eye movement patterns and visual attention during scene viewing in 3-to 12-month-olds. *Visual Neuroscience*, *33*, E014.

Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., & Hooge, I. T. C. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *Royal Society Open Science*, *5*(8).

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, *16*(4), 405–422.

Ignashchenkova, A., Dicke, P. W., Haarmeier, T., & Thier, P. (2004). Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nature Neuroscience*, *7*(1), 56–64.

Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *Journal of Neuroscience*, *29*(19), 6234–6238.

Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A Longitudinal Investigation of the Development of Attention and Distractibility. *Journal of Cognition and Development*, *7*(3), 381–409.

Khan, A. Z., Blohm, G., McPeek, R. M., & Lefèvre, P. (2009). Differential influence of attention on gaze and head movements. *Journal of Neurophysiology*, *101*(1), 198–206.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*(6), 781–791.

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*(2), 220–232.

Kretch, K. S., & Adolph, K. E. (2015). Active vision in passive locomotion: Real-world free viewing in infants and adults. *Developmental Science*, *18*(5), 736–750.

Lansink, J. M., & Richards, J. E. (1997). Heart rate and behavioral measures of attention in six-, nine-, and twelve-month-old infants during object exploration. *Child Development*, *68*(4), 610–620.

Lappi, O. (2016). Eye movements in the wild: Oculomotor control, gaze behavior & frames of reference. *Neuroscience and Biobehavioral Reviews*, *69*, 49–68.

Ledebt, A., & Bril, B. (2000). Acquisition of upper body stability during walking in toddlers. *Developmental Psychobiology*, *36*(4), 311–324.

Ledebt, A., & Wiener-Vacher, S. R. (1996). Head coordination in the sagittal plane in toddlers during walking: Preliminary results. *Brain Research Bulletin*, *40*(5–6), 371–373.

Lee, B. B. (1996). Receptive field structure in the primate retina. *Vision Research*, Vol. 36, pp. 631–644.

Li, F., Zheng, Y., Smith, S. D., Shic, F., Moore, C. C., Zheng, X., … Leckman, J. F. (2016). A preliminary study of movement intensity during a Go/No-Go task and its association with ADHD outcomes and symptom severity. *Child and Adolescent Psychiatry and Mental*

Health, *10*(1).

Li, Y., Fathi, A., & Rehg, J. M. (2013). Learning to predict gaze in egocentric video. *Proceedings of the IEEE International Conference on Computer Vision*, 3216–3223.

Libertus, K., & Hauf, P. (2017). Motor skills and their foundational role for perceptual, social, and cognitive development. *Frontiers in Psychology*, *8*(301).

Logan, G. D. (1996). The CODE theory of visual attention: an integration of space-based and object-based attention. *Psychological Review*, *103*(4), 603–649.

Martin, K. B., Hammal, Z., Ren, G., Cohn, J. F., Cassell, J., Ogihara, M., … Messinger, D. S. (2018). Objective measurement of head movement differences in children with and without autism spectrum disorder. *Molecular Autism*, *9*(1).

May, P. J. (2006). The mammalian superior colliculus: Laminar structure and connections. *Progress in Brain Research*, Vol. 151, pp. 321–378.

McGraw, M. (2004). Motor development. In R. M. Lerner, L. Liben, & U. Muller (Eds.), *Handbook of child psychology and developmental science: Vol. 2: Cognitive processes* (7th ed., pp. 79–95). New York: Wiley.

Meister, M., & Tessier-Lavigne, M. (2013). Low-Level Visual Processing: The Retina. *Principles of Neural Science, Fifth Edition*.

Meyer, A. F., O'Keefe, J., & Poort, J. (2020). Two Distinct Types of Eye-Head Coupling in Freely Moving Mice. *Current Biology*, *30*(11), 2116–2130.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14.

Müller, J. R., Philiastides, M. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(3), 524–529.

Nakagawa, A., & Sukigara, M. (2013). Variable coordination of eye and head movements during the early development of attention: A longitudinal study of infants aged 12-36 months. *Infant Behavior and Development*.

Nakashima, R., & Shioiri, S. (2014). Why do we move our head to look at an object in our peripheral region? Lateral viewing interferes with attentive search. *PLoS ONE*, *9*(3).

Oakes, L. M. (2015). A biopsychosocial perspective on looking behavior in infancy. In *Handbook of Infant Biopsychosocial Development*.

Ornitz, E. M., Kaplan, A. R., & Westlake, J. R. (1985). Development of the vestibulo-ocular reflex from infancy to adulthood. *Acta Oto-Laryngologica*, *100*(3–4), 180–193.

Poletti, M., Aytekin, M., & Rucci, M. (2015). Head-Eye Coordination at a Microscopic Scale. *Current Biology*.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, *3*(9), 741–747.

Proudlock, F. A., & Gottlob, I. (2007). Physiology and pathology of eye-head coordination. *Progress in Retinal and Eye Research*, *26*(5), 486–515.

Rast, M. M., & Harris, S. R. (2008). Motor Control in Infants with Down Syndrome. *Developmental Medicine & Child Neurology*, *27*(5), 682–685.

Reck, S. G., & Hund, A. M. (2011). Sustained attention and age predict inhibitory control during early childhood. *Journal of Experimental Child Psychology*, *108*(3), 504–512.

Regal, D. M., Ashmead, D. H., & Salapatek, P. (1983). The coordination of eye and head movements during early infancy: a selective review. *Behavioural Brain Research*, *10*(1), 125–132.

Reisman, J. E., & Anderson, J. H. (1989). Compensatory eye movements during head and body rotation in infants. *Brain Research*, *484*(1–2), 119–129.

Richards, J. E., & Cameron, D. (1989). Infant heart-rate variability and behavioral developmental status. *Infant Behav Dev*, *12*(1), 45–58.

Richards, J. E., & Casey, B. J. (1992). Development of sustained visual attention in the human infant. *Attention and Information Processing in Infants and Adults: Perspectives from Human and Animal Research*, 30–60.

Richards, J. E., & Hunter, S. K. (1997). Peripheral stimulus localization by infants with eye and head movements during visual attention. *Vision Research*.

Robertson, S. S., Johnson, S. L., Masnick, A. M., & Weiss, S. L. (2007). Robust coupling of body movement and gaze in young infants. *Developmental Psychobiology*, *49*(2), 208–215.

Rosander, K., & Von Hofsten, C. (2000). Visual-vestibular interaction in early infancy. *Experimental Brain Research*, *133*(3), 321–333.

Rosen, M. L., Amso, D., & McLaughlin, K. A. (2019). The role of the visual association cortex in scaffolding prefrontal cortex development: A novel mechanism linking socioeconomic status and executive function. *Developmental Cognitive Neuroscience*, *39*, 100699.

Ruff, H. A. (1986). Components of attention during infants' manipulative exploration. *Child Development*, *57*(1), 105–114.

Ruff, H. A., & Capozzoli, M. C. (2003). Development of Attention and Distractibility in the First 4 Years of Life. *Developmental Psychology*, *39*(5), 877–890.

Ruff, H. A., Capozzoli, M., & Weissberg, R. (1998). Age, individuality, and context as factors in sustained visual attention during the preschool years. *Developmental Psychology*, *34*(3), 454–464.

Ruff, H. A., & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. *Developmental Psychology*, *26*(1), 85–93.

Ruff, H. A., & Rothbart, M. K. (1996). Attention in Early Development: Themes and Variations. In *Attention in Early Development: Themes and Variations*. Oxford University Press.

Schlicht, E. J., & Schrater, P. R. (2007). Impact of coordinate transformation uncertainty on human sensorimotor control. *Journal of Neurophysiology*, *97*(6), 4203–4214.

Schmitow, C., Stenberg, G., Billard, A., & Hofsten, C. Von. (2013). Using a head-mounted camera to infer attention direction. *International Journal of Behavioral Development*, *37*(5), 468–474.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46.

Seemiller, E. S., Port, N. L., & Candy, T. R. (2018). The gaze stability of 4- to 10-week-old human infants. *Journal of Vision*, *18*(8), 1–10.

Slone, L. K., Abney, D. H., Borjon, J. I., Chen, C. H., Franchak, J. M., Pearcy, D., … Yu, C. (2018). Gaze in action: Head-mounted eye tracking of children's dynamic visual attention during naturalistic behavior. *Journal of Visualized Experiments*, *2018*(141).

Smith, L. B., & Thelen, E. (1996). *A dynamic systems approach to the development of cognition and action.* Cambridge: MIT press.

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. *Journal of Cognition and Development*, *16*(3), 407–419.

Solman, G. J. F., Foulsham, T., & Kingstone, A. (2017). Eye and head movements are complementary in visual selection. *Royal Society Open Science*, *4*(1).

Soska, K. C., Robinson, S. R., & Adolph, K. E. (2015). A new twist on old ideas: How sitting reorients crawlers. *Developmental Science*, *18*(2), 206–218.

Stahl, J. S. (1999). Amplitude of human head movements associated with horizontal saccades. *Experimental Brain Research*, *126*(1), 41–54.

Stryker, M. P., & Schiller, P. H. (1975). Eye and head movements evoked by electrical stimulation of monkey superior colliculus. *Experimental Brain Research*, *23*(1), 103–112.

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Developmental Psychology*, *55*(1), 96–109.

Tagliabue, M., & McIntyre, J. (2014). A modular theory of multisensory integration for motor control. *Frontiers in Computational Neuroscience*, *8*(JAN).

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14).

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *Journal of Vision*, *11*(5), 5.

Teicher, M. H., Ito, Y., Glod, C. A., & Barber, N. I. (1996). Objective measurement of hyperactivity and attentional problems in ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(3), 334–342.

Thaler, L., & Todd, J. T. (2009). The use of head/eye-centered, hand-centered and allocentric representations for visually guided hand movements and perceptual judgments. *Neuropsychologia*, *47*(5), 1227–1244.

Trevarthen, C. (1984). How Control of Movement Develops. *Advances in Psychology*, *17*(C), 223–261.

Tronick, E., & Clanton, C. (1971). Infant looking patterns. *Vision Research*, *11*(12), 1479–1486.

Tseng, P. H., Carmi, R., Cameron, I. G. M., Munoz, D. P., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, *9*(7).

van Renswoude, D. R., van den Berg, L., Raijmakers, M. E. J., & Visser, I. (2019). Infants' center bias in free viewing of real-world scenes. *Vision Research*, *154*, 44–53.

Visscher, C., Houwen, S., Scherder, E. J. A., Moolenaar, B., & Hartman, E. (2007). Motor profile of children with developmental speech and language disorders. *Pediatrics*, *120*(1).

von Hofsten, C., & Rosander, K. (2018). The Development of Sensorimotor Intelligence in Infants. *Advances in Child Development and Behavior*, *55*, 73–106.

Von Hofsten, C., Vishton, P., Spelke, E. S., Feng, Q., & Rosander, K. (1998). Predictive action in infancy: Tracking and reaching for moving objects. *Cognition*, *67*(3), 255–285.

Vuijk, P. J., Hartman, E., Scherder, E., & Visscher, C. (2010). Motor performance of children with mild intellectual disability and borderline intellectual functioning. *Journal of Intellectual Disability Research*, *54*(11), 955–965.

Walton, M. M. G., Bechara, B., & Gandhi, N. J. (2007). Role of the primate superior colliculus in the control of head movements. *Journal of Neurophysiology*, *98*(4), 2022–2037.

Wang, X., Koch, S., Holmqvist, K., & Alexa, M. (2018). Tracking the gaze on objects in 3D: How do people really look at the bunny? *SIGGRAPH Asia 2018 Technical Papers, SIGGRAPH Asia 2018*.

Wass, S. V, Clackson, K., Georgieva, S. D., Brightman, L., Nutbrown, R., & Leong, V. (2018). Infants' visual sustained attention is higher during joint play than solo play: is this due to increased endogenous attention control or exogenous stimulus capture? *Developmental Science*, *21*(6), e12667.

Wass, S. V, Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., … Leong, V. (2018). Parental neural responsivity to infants' visual attention: How mature brains influence immature brains during social interaction. *PLoS Biology*, *16*(12).

Wass, S. V, & Smith, T. J. (2014). Individual differences in infant oculomotor behavior during the viewing of complex naturalistic scenes. *Infancy*, *19*(4), 352–384.

Weissman, B. M., DiScenna, A. O., & Leigh, R. J. (1989). Maturation of the vestibulo-ocular reflex in normal infants during the first 2 months of life. *Neurology*, *39*(4), 534–538.

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The Development of Cognitive Skills and Gains in Academic School Readiness for Children from Low-Income Families. *Journal of Educational Psychology*, *102*(1), 43–53.

Werchan, D. M., & Amso, D. (2017). A novel ecological account of prefrontal cortex functional development. *Psychological Review*, *124*(6), 720–739.

Wiener-Vacher, S. R. (1996). Changes in otolith VOR to off vertical axis rotation in infants learning to walk preliminary results of a longitudinal study. *Annals of the New York Academy of Sciences*, *781*, 709–712.

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

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262.

Yu, C., & Smith, L. B. (2016). The Social Origins of Sustained Attention in One-Year-Old Human Infants. *Current Biology*, *26*(9), 1235–1240.

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, *22*(1).

Yuan, L., Xu, T. L., Yu, C., & Smith, L. B. (2019). Sustained visual attention is more than seeing. *Journal of Experimental Child Psychology*, *179*, 324–336.