

# What children with and without ASD see: Similar visual experiences with different pathways through parental attention strategies

Autism & Developmental Language

Impairments

Volume 7: 1–16

© The Author(s) 2022

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/23969415221137293

[journals.sagepub.com/home/dli](https://journals.sagepub.com/home/dli)



**Elizabeth Perkovich** 

Department of Psychology, University of Houston, Houston, TX, USA

**Lichao Sun**

Department of Psychology, University of Houston, Houston, TX, USA

**Sarah Mire**

Educational Psychology Department, Baylor University, Waco, TX, USA

**Anna Laakman**

Department of Psychological Health and Learning Sciences, University of Houston,  
Houston, TX, USA

**Urvi Sakhuja**

Department of Psychology, University of Houston, Houston, TX, USA

**Hanako Yoshida**

Department of Psychology, University of Houston, Houston, TX, USA

## Abstract

**Background and aims:** Although young children's gaze behaviors in experimental task contexts have been shown to be potential biobehavioral markers relevant to autism spectrum disorder (ASD), we know little about their everyday gaze behaviors. The present study aims (1) to document early gaze behaviors that occur within a live, social interactive context among children with and without ASD and their parents, and (2) to examine how children's and parents' gaze behaviors are related for ASD and typically developing (TD) groups. A head-mounted eye-tracking system was used to record the frequency and duration of a set of gaze behaviors (such as sustained attention [SA] and joint attention [JA]) that are relevant to early cognitive and language development.

**Methods:** Twenty-six parent-child dyads (ASD group = 13, TD group = 13) participated. Children were between the ages of 3 and 8 years old. We placed head-mounted eye trackers on parents and children to record their parent- and child-centered views, and we also recorded their interactive parent-child object play scene from both a wall- and ceiling-mounted camera. We then annotated the frequency and duration of gaze behaviors (saccades, fixation, SA, and JA) for different regions of interest (object, face, and hands), and attention shifting. Independent group t-tests and ANOVAs were used to observe group comparisons, and linear regression was used to test the predictiveness of parent gaze behaviors for JA.

**Results:** The present study found no differences in visual experiences between children with and without ASD. Interestingly, however, significant group differences were found for *parent* gaze behaviors. Compared to parents of ASD children, parents of TD children focused on objects and shifted their attention between objects and their children's

---

## Corresponding author:

Elizabeth Perkovich, Department of Psychology, University of Houston, Houston, TX 77204, USA.

Email: [esperkovich@uh.edu](mailto:esperkovich@uh.edu)



faces more. In contrast, parents of ASD children were more likely to shift their attention between their own hands and their children. JA experiences were also predicted differently, depending on the group: among parents of TD children, attention to *objects* predicted JA, but among parents of ASD children, attention to their *children* predicted JA.

**Conclusion:** Although no differences were found between gaze behaviors of autistic and TD children in this study, there were significant group differences in parents' looking behaviors. This suggests potentially differential pathways for the scaffolding effect of parental gaze for ASD children compared with TD children.

**Implications:** The present study revealed the impact of everyday life, social interactive context on early visual experiences, and point to potentially different pathways by which parental looking behaviors guide the looking behaviors of children with and without ASD. Identifying parental social input relevant to early attention development (e.g., JA) among autistic children has implications for mechanisms that could support socially mediated attention behaviors that have been documented to facilitate early cognitive and language development and implications for the development of parent-mediated interventions for young children with or at risk for ASD.

Note: This paper uses a combination of person-first and identity-first language, an intentional decision aligning with comments put forth by Vivanti (Vivanti, 2020), recognizing the complexities of known and unknown preferences of those in the larger autism community.

### Keywords

Autism spectrum disorders, cognition (attention, memory, learning), interaction, social behavior, social cognition

## Introduction

Young children with autism spectrum disorder (ASD) demonstrate the social communication difficulties characteristic of this neurodevelopmental disability in a variety of ways. Specifically, diagnosed children's difficulties may include less social play, exploratory behavior during play, and recognition of social cues (e.g., facial expressions and body language) compared to peers without ASD (Falck-Ytter et al., 2013; Jellema et al., 2009; Jordan, 2003; Ozonoff et al., 2008). Previous studies concerning autistic children's gaze behaviors documented that atypical visual attention, such as limited rates of sustained attention (SA) to referential cues (e.g., social gaze, gaze and point following, gestures, and joint attention [JA], defined as socially shared attentional experiences such as the parent and child visually attending to the same referent), have been proposed as a key contributor to these social communication development difficulties (Murias et al., 2018).

Since the line of research exploring differences in gaze behaviors between autistic children and TD children often (1) use stationary eye tracking in a computer viewing task or (2) use observing researchers to make inferences about the direction of looking, we know little about the everyday visual experiences of children with ASD and how such visual experiences are organized in a social context. Human gaze behaviors are constantly shaped and supported by incoming input from the environment, including social partners' movements and changes generated by one's own body movement. Technological advancements—wearable head-mounted camera/eye-tracking systems—allow us to study moment-to-moment gaze behaviors while parent

and child interact in a naturalistic play context, and recent work indicates how TD children's visual experiences are tightly coupled with parent's social input (Burling & Yoshida, 2019; Smith et al., 2015).

More specifically, in a parent-child interaction context, previous head-camera studies have shown that parents of young TD children actively hold objects close to their child's head and eyes, creating a clear view of the object. These moments are what we call "optimal viewing moments" (OVM). When parents also name the objects during these moments, we call these "visually optimal naming moments" (VONM). Another study further documented that 18-month-old TD infants are more likely to learn those object names when tested immediately after the object play (Yu & Smith, 2012). The increased SA is also attributable to the early scaffolding in which parents generate these clear viewing experiences (Burling & Yoshida, 2019; Sun & Yoshida, 2022). We also know that joint attention (JA) is essential for language and cognitive development (Brooks & Meltzoff, 2005; Tomasello & Farrar, 1986) and that caregivers, through their actions, provide effective nonverbal cues to intended referents (Gogate et al., 2006; Gogate & Bahrick, 1998, 2001; Koterba & Iverson, 2009). To understand how these critical visual experiences are generated for children with ASD, it is essential to characterize how the gaze behaviors of children with ASD are organized by parental input.

The few studies that have investigated these gaze behaviors in a social context have documented similarities in children's gaze behaviors across ASD and TD groups, such as attention to parent's face, toy objects, SA, and JA (Yoshida et al., 2019; Yurkovic et al., 2021). However,

the results of one of the studies focusing on 36–60-month-old children and their parents also suggested group differences in *parental input*—in that parents of ASD children utilized more gestures, more closely monitored their child’s face, and provided more scaffolding for their child’s visual experience (Yoshida et al., 2019). One of the other studies using a head-mounted eye-tracking system directly measuring attentional behaviors in 24–48-month-old children found no significant differences in how autistic children and TD children distributed their visual attention, generated manual action, or coordinated their visual and manual experiences during play with their parents (Yurkovic et al., 2021).

The initial attempts of documenting everyday visual experiences of young children with ASD in a social context suggest no significant differences between autistic children and TD children, which is in contrast to the documented ASD-specific gaze behaviors in computer screen-based studies. Why are these group differences not shown in a live social interactive context? Previous observations of parent–child social interactions have documented that parents of autistic children and parents of TD children scaffold their child’s play differently (Freeman & Kasari, 2013; Yoshida et al., 2019). For example, during an observation of a 10-min free play with children between 20 and 60 months old, parents of children with ASD were more actively engaged with their children than were parents of TD children (Freeman & Kasari, 2013). In this study, parents of autistic children were more likely to initiate play schemes (sequences made up of connected play-acts with connected toys), suggest and command more play-acts, and respond to their children’s play-acts with a higher level play-act. In the same vein, another study with children 3–5 years old found that mothers of children with ASD used more physical contact, more high-intensity behaviors, and fewer social verbal approaches (Doussard-Roosevelt et al., 2003).

Communication abilities may also make a difference in terms of parents’ behaviors, as Kasari and colleagues (1988) found that among children aged 42–65 months, parents of children with ASD spent more time physically orienting their children to tasks and adjusting their own behavior differentially to their children’s specific ASD symptoms (i.e., regulating their child’s behavior less often and engaging in more positive feedback and mutual play when the autistic child had higher communication abilities). Similarly, in 10-min semistructured play interactions, more severe ASD symptoms predicted increased adversely affected parent–child interactions among children aged 4–14 years, including decreased coordination, communication, and responsivity (Beurkens et al., 2013). Finally, parents of autistic children may provide more scaffolding for their children by producing more gestures and more closely monitoring their children’s faces during play compared to parents whose children do not have this diagnosis,

among children between the ages of 36 and 60 months (Yoshida et al., 2019).

The current literature raises the possibility that early visual experiences of children with ASD might be differentially supported in a naturalistic social context when compared with their TD counterparts. The present study explored this possibility by studying similarities and differences in children’s (ASD and TD) and their parent’s moment-to-moment dynamic gaze behaviors during parent–child interactive play, including micro-gaze behaviors such as saccades, fixation, and sustained attention. Studying early gaze behaviors during parent–child interactive play, where everyday communication and learning occurs, and testing how parents’ social input organizes early visual experiences of autistic children may provide the foundation for developing novel interventions that may ultimately support social communication outcomes.

### *Current aims and research questions*

Our overarching aim in this exploratory study was to document the visual experiences of autistic children and TD children during live social interaction with their parents and explore how parents through their looking behaviors may differentially scaffold their children’s visual experiences as a function of child’s diagnostic outcomes (ASD vs. TD group). The present study recorded observations of parent–child play using head-mounted camera systems. These methods allowed for the measurement of precise, moment-by-moment visual experiences while also allowing the child to actively move around their environment and interact with a variety of stimuli and with their parents. Specific questions investigated were: 1) How are visual experiences different and similar between *children* with ASD and TD children during live social interaction with their parents? We hypothesized that visual experiences—the rate of SA, JA (and precise gaze behaviors considered in computer viewing tasks such as saccades, fixation, and attention shifting)—would also be similar across the groups. 2) How are visual experiences different and similar between *parents* of children with ASD and TD during live social interaction? We hypothesized that parents of ASD children will pay more attention to their children, while parents of TD children will pay more attention to objects. 3) Do parental visual experiences impact the rate of their child’s JA? We hypothesized that parental visual experiences would predict JA experiences for both groups.

## **Methods**

### *Participants*

Observation data from 26 parent–child dyads with children between the ages of 3 and 8 years old were collected by our

team and included in the present study, 13 of whom had parent-reported diagnoses of ASD ( $M = 5.69$  years,  $SD = 1.81$  years, 3 females). The other 13 were reported by their parents to have no history of language or developmental delays ( $M = 4.20$  years,  $SD = 0.92$  years, 7 females) and were considered to be TD children. The racial and ethnic backgrounds of the parents in the ASD group (12 mothers, 1 father) included White (5), Black or African American (1), Hispanic (5), and Asian (2). The ethnic backgrounds of the parents in the TD group (13 mothers) included White (6), Black or African American (2), Hispanic (3), and Asian (2). All families provided informed consent and the study and its procedures were approved by the Institutional Review Board of the university where the research took place.

Five of the children in the ASD group had their ASD diagnosis confirmed with the *Autism Diagnostic Observation Schedule* (ADOS) (Lord et al., 2000) prior to their lab visit and the other eight children had their ASD diagnosis confirmed with the *Autism Diagnostic Observation Schedule, Second Edition* (ADOS-2; Lord et al., 2012) during their lab visit by research-reliable administrators. Both the ADOS and ADOS-2 are semistructured assessments that allow for the observation of symptoms related to ASD during social, communicative, and play interactions with trained clinicians. The ADOS-2 is a revision of the ADOS, featuring updated coding items and descriptions, and algorithm items. The general structure, administration guidelines, and goals of the ADOS-2 remain largely unchanged from the ADOS. Children were administered ADOS/ADOS-2 Module 1, 2, or 3 based on their expressive language ability. All 13 participants' ADOS/ADOS-2 scores exceeded the cutoff score for inclusion in the ASD group for this study (i.e., their scores were consistent with those of children who have been diagnosed with ASD) (Lord et al., 2012).

We used three vocabulary measures to characterize our sample further: the MCDI; (Fenson et al., 1993, 2007), the Peabody Picture Vocabulary Test Fourth Edition, which measures receptive vocabulary (PPVT-4) (Dunn & Dunn, 2007), and the Expressive Vocabulary Test Second Edition, which measures expressive vocabulary (EVT-2) (Williams, 2007). The MCDI is a parental vocabulary checklist consisting of early-learned words and sentences by TD children and has been used in populations of older ASD children to report expressive language ability (Plesa Skwerer et al., 2016). Although parents of all 26 children completed the MCDI, not all participants completed the PPVT-4 and EVT-2 due to changes in the experimental protocol (six children from the ASD group and eight children from the TD group completed the PPVT-4 and five children from the ASD group and eight children from the TD group completed the EVT-2). These measures were collected to characterize our sample and were not included in the current analyses (i.e., to test effects). Child age, gender,

ADOS/ADOS-2 Module, and language scores for individual participants are reported in Table 1.

### Equipment

There were two generations of head-mounted camera systems used to collect the data analyzed for this study. A small head-mounted camera (Watec WAT-230A, miniature (25 g) color video camera) was initially used to capture child-centered views alongside a wall-mounted camera that captured an overall view of the scene. The focal length of the lens is f2.1 mm, the numbers of effective pixels are 512 (H) · 492 (V), and the camera's visual field is 84 degrees (V) and 115 (H). This camera system was used for nine of the twenty-six parent-child dyads (five ASD and four TD). Furthermore, the second head-mounted camera system included two small cameras and an infrared light-emitting diode (LED) and weighed 51g in total. One of the cameras faced the child's or parent's right eye and recorded the pupil movements and corneal reflection while a head-mounted camera placed on the forehead recorded the visual field from the child's or parent's perspective (FPV: 54.4° horizontal by 42.2° vertical) (Figure 1). This head-mounted eye-tracking system was used for the other seventeen parent-child dyads (8 ASD and 9 TD). In addition, a wall-mounted camera and a ceiling-mounted camera were also used to capture an overall view of the scene for all of the parent-child dyads. All of these videos were later synced with the videos captured from the head-mounted cameras. All videos recorded by both camera systems were processed at a rate of 33 ms per frame; hence, the use of different camera systems does not impact the capacity of testing the present set of hypotheses. On average, the non-upgraded devices generated 9,477 ( $SD = 290.9$ ) frames, while the upgraded devices generated 10,010 ( $SD = 155.6$ ) frames for child gaze data during the live social interaction task. There was no difference in the number of frames for the play session between the two recording devices. In addition, the upgraded devices generated 10,007 frames ( $SD = 160.3$ ) for parent gaze data during the live, social interaction task for a more precise estimate of gaze patterns and JA. Long and lightweight cables were used in both device types for the study to allow for free, unrestrained engagement during the play session.

### Procedure

The small room in which the play session took was located on the university campus of the principal investigator's institution and contained a child-sized table with a child-sized chair on either side for the parent and child, as well as the computer and eye-tracking equipment. The walls, door, and two-way mirror were covered by black curtains to limit distraction. The computer and the research

**Table 1.** ADOS:n autism diagnostic observation schedule. ADOS-2: Autism Diagnostic Observation Schedule, Second Edition.<sup>a</sup>

Development group	Subject #	Sex	Age (months)	ADOS/ADOS-2 module	MCDI (%)	PPVT-4	EVT-2
Autism spectrum disorder (ASD)							
n = 13	1	Male	59.7	ADOS Module 2	54	-	-
	2	Male	44.8	ADOS Module 2	88	-	-
	3	Female	46.1	ADOS Module 1	27	-	-
	4	Female	61.7	ADOS Module 2	70	-	-
	5	Male	63.5	ADOS Module 2	96	-	-
	6	Male	65.4	ADOS-2 Module 1	54	-	-
	7	Male	107.0	ADOS-2 Module 3	99	104	94
	8	Male	83.3	ADOS-2 Module 3	99	106	106
	9	Male	96.0	ADOS-2 Module 3	99	103	101
	10	Male	56.6	ADOS-2 Module 1	11	-	-
	11	Female	47.0	ADOS-2 Module 1	0	55	-
	12	Male	107.2	ADOS-2 Module 3	99	110	107
	13	Male	49.1	ADOS-2 Module 1	63	90	100
Mean (SD)			<b>68.3 (22.7)</b>		<b>66.1 (35.3)</b>	<b>94.7 (20.6)</b>	<b>101.6 (5.2)</b>
Typically developing (TD)							
n = 14	15	Female	33.3	-	91	-	-
	16	Female	57.3	-	57	-	-
	17	Female	56.1	-	99	-	-
	18	Male	47.9	-	90	-	-
	19	Female	37.8	-	88	-	-
	20	Female	61.9	-	99	125	107
	21	Male	56.4	-	68	96	103
	22	Male	48.4	-	99	140	116
	23	Male	41.4	-	64	95	95
	24	Male	55.6	-	99	110	106
	25	Female	74.2	-	99	143	135
	26	Male	56.0	-	98	98	100
	27	Female	68.0	-	99	109	108
Mean (SD)			<b>53.4 (11.6)</b>		<b>88.5 (15.2)</b>	<b>114.5 (19.3)</b>	<b>109.8 (12.7)</b>

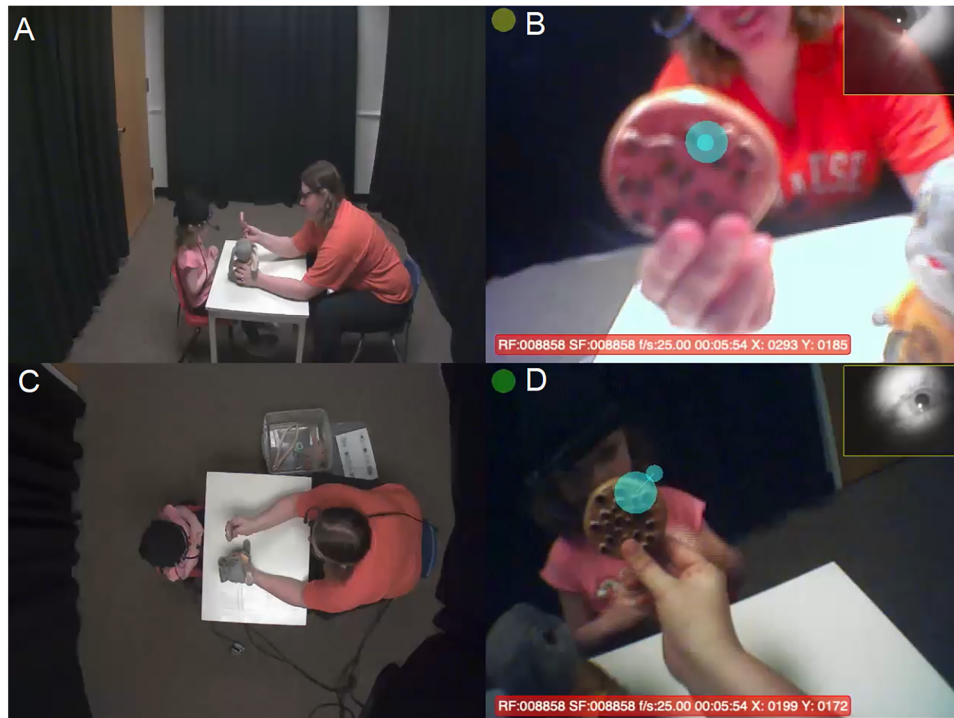
<sup>a</sup>MCDI: MacArthur Communicative Development Inventory: Words and Sentences (reported as the percentage of words known by each child as reported by the parent using a normative sample of words that are known by 50% of children by 30 months). PPVT-4: Peabody Picture Vocabulary Test Fourth Edition (reported as a standard score based on age). EVT-2: Expressive Vocabulary Test Second Edition (reported as a standard score based on age). Bolded numbers represent the Mean (SD) as noted by the first column.

assistants monitoring the equipment were on the opposite side of the room, further blocked by a black curtain during the duration of the play session. During the play session, the parent and child dyads sat at the table with a box of objects and were instructed to play as they normally would at home. Play sessions were 320 s (5 min and 20 s) long, which is consistent with previous studies utilizing head-mounted eye-trackers with young children during play with their parents (Borjon et al., 2021; Deák et al., 2018; MacNeill et al., 2021; Yu & Smith, 2013, 2017; Yurkovic et al., 2021; Yurkovic-Harding et al., 2022). Each play session featured eight objects, four of which were considered early learned nouns (bottle, bunny, car, and cookie), which are listed on the *MacArthur Communicative Development Inventory: Words and Sentences* (Fenson et al., 1993, 2007), a developmental vocabulary checklist for children under 30 months old, and four of which were considered unfamiliar words

(caliper, nylon, pipette, and strainer) that are not among the most frequently used English words. The unfamiliar words were added to the play session in order to increase task complexity for the purpose of maintaining the child's interest in the play session, as children are more likely to select and attend to unfamiliar objects (Henderson, 1981; Horst et al., 2011).

### Data processing and coding

The recordings from the head-mounted camera systems and scene view videos (wall-mounted and ceiling-mounted camera views) were synchronized and imported into the Datavyu coding software where each of the target behavior variables was annotated by trained research assistant coders who had achieved >84% reliability in coding, were blind to the hypotheses and had over 70 h of prior coding training each (Datavyu Team, 2014). The trained coders went



**Figure 1.** (A) Wall-mounted view of the parent–child play session. (B) The first-person view of the parent–child play session of the child. The blue calibration dot indicates the ROI to which the child looked during any given frame. (C) Ceiling-mounted view of the parent–child play session. (D) The first-person view of the parent–child play session of the parent. The blue calibration dot indicates the ROI to which the parent looked during any given frame.

frame-by-frame of both the child and parent head-camera views and annotated the frequency and duration of each instance the child or parent looked at a region of interest (region of interest [ROI]: objects, child hands, parent hands, child face, and parent face). These behaviors were further annotated for saccades (<200 ms in duration), fixation (200–2000 ms), SA (>2000 ms), and frequency of shifts between ROIs (Burling & Yoshida, 2019; Campbell et al., 2014; Devillez et al., 2020; van Beers, 2007; Yu et al., 2019). JA was determined to be temporally aligned looking instances where both the child and parent focused on the same ROI. Operationalization of these variables was based on previous head-mounted eye-tracking studies in developmental and eye-tracking literature (Deák et al., 2018; Yoshida & Smith, 2008; Yu & Smith, 2013).

### Data analysis

The first and second hypotheses concerning the group comparisons in the visual experiences (e.g., saccades, fixation, SA, and JA) to various ROIs of children and parents were tested by general linear modeling. The third hypothesis concerning the predictiveness of parents' attention to ROIs upon their child's JA was tested by a set of linear regressions. An alpha level of 0.05 was used for all statistical tests.

## Results

### Children's gaze behaviors: Gaze location

We first examined how children distributed their attention during the play session among ROIs in terms of frequency and duration measurements per group and summarized in Table 2.

Autistic children, on average, spent 43.17% of the play-session duration looking at ROIs compared to non-ROIs (e.g., the wall, table, and ceiling) while TD children, on average, spent 50.01% of the play-session duration looking at ROIs compared to non-ROIs ( $t(24) = -0.67$ ,  $p = .53$ ).

We examined the effects of group (TD vs. ASD) and the attention allocations on ROIs (objects, parent's hands, child's hands, parents' face) by linear mixed models on both frequency and duration of children's attention. First, the mixed model on frequency counts had fixed effects for groups and types of ROIs and random effects for participants and ROIs within participants to account for the variability in dyads across ROIs. The frequency model revealed no group difference in attention throughout the play session,  $B = 20.96$ ,  $t = 0.63$ ,  $p = .528$  (see the full model in S1 in Supplementary Materials). Children with and without ASD shared a similar ROI distribution in the following order, from greatest to least: objects, parent's

**Table 2.** Means and standard deviations (SD) of frequency (number of discrete looks) and duration (ms) of observed child visual experiences by group, averaged across participants within the ASD and TD groups.

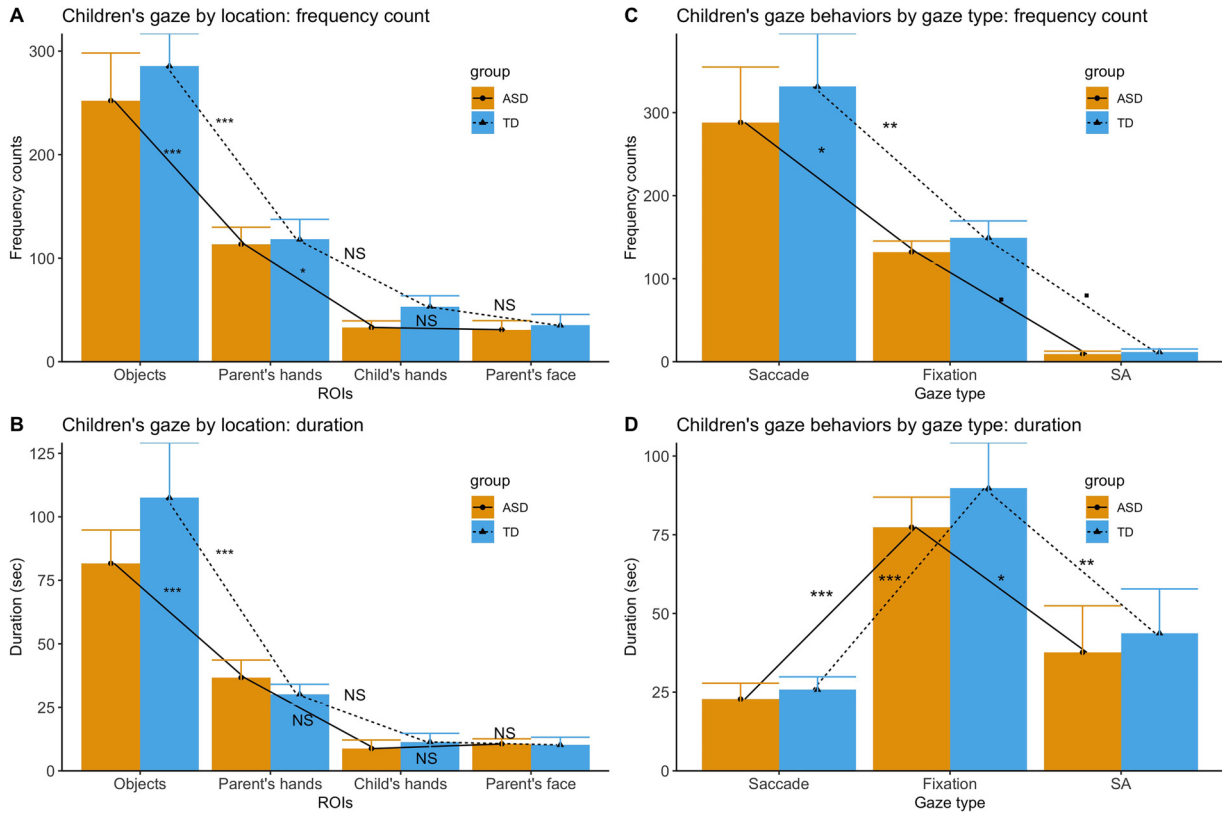
Child behavior	Group	Frequency M (SD)	p value	Duration M (SD)	p value
Attention to Objects	ASD	252.2 (166.0)	.55	81661.5 (47376.1)	.32
	TD	285.6 (113.0)		107499.4 (78125.7)	
Attention to Parent Face	ASD	30.9 (31.5)	.76	10631.5 (7126.6)	.91
	TD	35.2 (37.8)		10239.5 (10812.8)	
Attention to Parent Hands	ASD	113.5 (59.4)	.85	36701.0 (24932.1)	.42
	TD	118.2 (69.5)		30152.4 (14139.0)	
Attention to Child Hands	ASD	33.1 (23.0)	.11	8794.7 (12076.7)	.59
	TD	53.3 (37.4)		11390.9 (12152.0)	
Saccades	ASD	288.2 (240.5)	.65	22779.0 (18246.0)	.65
	TD	331.3 (230.2)		25780.6 (14717.1)	
Fixation	ASD	132.2 (47.5)	.48	77380.4 (34498.6)	.48
	TD	149.4 (72.8)		89787.7 (52056.7)	
SA	ASD	9.3 (12.3)	.64	37629.2 (53345.3)	.77
	TD	11.7 (13.1)		43713.8 (50744.5)	
Saccades to Objects	ASD	170.0 (154.4)	.79	13435.7 (12118.6)	.82
	TD	185.3 (133.1)		14374.7 (8775.1)	
Fixation to Objects	ASD	76.1 (37.1)	.42	43456.2 (23977.5)	.33
	TD	90.7 (52.4)		56430.4 (40476.8)	
SA to Objects	ASD	6.0 (8.2)	.34	24769.6 (35478.5)	.45
	TD	9.6 (10.7)		36694.3 (43430.9)	
Saccades to Parent Face	ASD	17.5 (22.3)	.54	1517.5 (1969.7)	.60
	TD	24.1 (30.3)		1999.3 (2603.4)	
Fixation to Parent Face	ASD	12.9 (10.6)	.59	7751.3 (5510.8)	.68
	TD	10.6 (11.3)		6627.4 (7806.4)	
SA to Parent Face	ASD	0.5 (1.0)	.84	1362.6 (2885.1)	.83
	TD	0.5 (1.0)		1612.8 (3067.3)	
Saccades to Parent Hands	ASD	76.5 (62.8)	.81	5913.2 (4242.1)	.71
	TD	82.8 (70.7)		6592.2 (4882.1)	
Fixation to Parent Hands	ASD	34.7 (16.3)	.97	21346.6 (13727.6)	.77
	TD	34.4 (19.2)		19848.2 (11562.6)	
SA to Parent Hands	ASD	2.3 (3.7)	.26	9441.3 (15297.2)	.24
	TD	1.0 (1.5)		3711.9 (7249.5)	
Saccades to Child Hands	ASD	24.0 (23.0)	.20	1910.1 (1727.6)	.25
	TD	39.0 (36.9)		2806.5 (2132.4)	
Fixation to Child Hands	ASD	8.5 (7.4)	.31	4828.9 (5345.1)	.50
	TD	13.8 (16.4)		6889.6 (9251.5)	
SA to Child Hands	ASD	0.5 (1.9)	1.00	2055.7 (7411.9)	.88
	TD	0.5 (1.4)		1694.8 (4460.0)	
JA	ASD	67.1 (43.9)	.62	81962.6 (105517.8)	.51
	TD	76.5 (49.2)		113908.2 (128078.3)	

hands, child's hands, and parent's face. Specifically, the post-hoc Tukey contrasts showed that both groups attended the target objects significantly more than other ROIs, following with attention to their parent's hands. In addition, children with ASD attended more on parent's hands than their own hands ( $\beta = 80.39$ ,  $t = 2.86$ ,  $p = .028$ ), while TD children had no attention difference between parent's and child's hands ( $\beta = 67.15$ ,  $t = 2.30$ ,  $p = .109$ ; also see Figure 2A and S2).

Similarly, the mixed model on gaze duration with fixed effects for groups and types of ROI and a random effect for participants was run. Children from both groups had no

difference in the duration of attention on ROIs:  $\beta = -6.32$ ,  $t = -0.47$ ,  $p = .639$ . Notably, there was an age effect in the duration of ROIs, that older children attended significantly less time on ROIs,  $\beta = -0.59$ ,  $t = -2.94$ ,  $p = .007$  (see the full model in S3). Post-hoc Tukey tests revealed that children with and without ASD shared a similar ROI distribution in the following order, from greatest to least: objects, parent's hands, parent's face, and child's hands (see Figure 2B). Specifically, both groups attended to the target objects significantly longer than the other ROIs, following with attention to their parent's hands. There was no difference in overall duration





**Figure 2.** The distribution of children's gaze behaviors. (A) Children's gaze by ROIs counted by frequency. (B) Children's gaze by ROIs counted by duration (second [sec]). (C) Children's gaze by gaze type counted by frequency. (D) Children's gaze by gaze type counted by duration (sec). A, Children's gaze counted by frequency across ROIs. Significance levels: marginal, \*  $<.05$ , \*\*  $<.01$ , \*\*\*  $<.001$ .

between the child's attention on the parent's face and on the child's own hands.

### Children's gaze behaviors: Gaze type

We focused on children's attention on ROIs only and classified these by the type of gaze: (1) saccades, (2) fixations, and (3) SA. First, S4 summarizes the mixed model with fixed effects for group and gaze type and random slope of gaze type within participants. It revealed no group difference in frequency counts in the variety of gaze behaviors,  $\beta = 12.09$ ,  $t = 0.21$ ,  $p = .835$ . Moreover, gaze type accounted for frequency distribution, accompanying the significant fixed effects of SA ( $\beta = -122.85$ ,  $t = -2.21$ ,  $p = .031$ ) and saccade ( $\beta = 156.00$ ,  $t = 2.80$ ,  $p = .007$ ). The follow-up Tukey contrasts on gaze type also showed that in both groups, saccades represented the highest frequency, followed by fixations and SA as the least; all the comparisons were statistically significant (see Figure 2C).

Similarly, another mixed model with fixed effects for group and gaze type and a random effect for participants confirmed no group effect on attention duration,  $\beta = 0.46$ ,  $t = 0.03$ ,  $p = .976$ ; in other words, both groups had

comparable duration counts on different types of gaze. Meanwhile, gaze type accounted for attention duration, with significant fixed effects of SA ( $\beta = -39.75$ ,  $t = -2.73$ ,  $p = .009$ ) and saccades ( $\beta = -54.60$ ,  $t = -3.75$ ,  $p < .001$ ). The follow-up Tukey comparisons showed that children from both groups exhibited their longest attention in fixations, followed by comparable durations of SA and saccades (see Figure 2D). In addition, the child's age was controlled in the model and showed that older children attended for significantly less time across ROIs,  $\beta = -0.53$ ,  $t = -3.00$ ,  $p = .005$  (see the full model in S5).

### Children's gaze behaviors: Gaze shifting

We also counted children's gaze shifting among ROIs, in specific, how likely the child tended to shift attention to target objects from other ROIs. The generalized linear model fitting with a quasi-Poisson distribution predicted the likelihood of gaze shifting to objects as a function of groups and prior gaze locations (i.e., parent's face, parent's hands, child's hands). There was no main group difference in the likelihood of gaze shifting to objects ( $\beta = 0.17$ ,  $t = 1.15$ ,  $p = .253$ ), while the effects of gaze shifting



**Table 3.** The model summary of children's gaze shifting to target objects.

Term	Estimate	Standard error	Statistic	p value	Conf.low	Conf.high
(Intercept)	3.12	0.16	19.31	<.001	2.79	3.42
Group:TD	0.17	0.15	1.15	.253	-0.12	0.46
Parent's face to object	-1.05	0.28	-3.79	<.001	-1.62	-0.53
Parent's hands to object	0.87	0.17	5.20	<.001	0.55	1.21

from face to objects ( $\beta = -1.05$ ,  $t = -3.79$ ,  $p < .001$ ) and gaze shifting from parent's hands to objects were significant ( $\beta = 0.87$ ,  $t = 5.20$ ,  $p < .001$ , see the model in Table 3). The model was performed in log scale and transformed to the incident rate ratios for group comparisons: the incident rate for parent's hands-to-object was 2.39 times the incident rate for child's hands-to-object ( $z = 5.20$ ,  $p < .001$ ); the incident rate for child's hands-to-object was 2.85 times the incident rate for parent's face-to-object ( $z = 3.79$ ,  $p < .001$ ), and the incident rate for parent's face-to-object was 0.15 times the incident rate for parent's hands-to-objects ( $z = -7.54$ ,  $p < .001$ ; also see Figure 4A).

### Parent's gaze behaviors: Gaze location

We first examined how parents distributed their attention during the play session among ROIs in terms of frequency and duration measurements per group and summarized in Table 4.

Secondly, we examined how *parents* distributed their attention during the play session as a function of group and type of ROIs; we ran the linear mixed models on both frequency and duration counts of the parent's gaze behaviors throughout the play session.

Similar to the models on the children's gaze, the mixed model on parent gaze included a random slope by participants to account for the variability of each dyad. Table 2 presents the model summary and reveals a significant interaction between the TD group and gaze on objects,  $\beta = 191.29$ ,  $t = 2.84$ ,  $p = .007$  (see the full model in S6). The follow-up group contrasts by location indicated that parents of TD children had significantly more gaze on target objects than parents of ASD children did,  $\beta = 189.55$ ,  $t = 3.38$ ,  $p = .002$ , but no attention differences between groups on other ROIs. Besides, parents of both groups had a similar ROI distribution in the following order from greatest to least: objects, child's hands, child's face, and parent's hands. Specifically, the post-hoc Tukey contrasts showed that both groups attended the target objects significantly more than the other ROIs, and they had comparable counts of attention on the child's hands, child's face, and parent's own hands (see Figure 4A).

Moreover, the duration model with fixed effects of group and type of ROI and a random effect of participants also showed a significant interaction between TD group and gaze on objects,  $\beta = 72.49$ ,  $t = 3.33$ ,  $p = .002$  (see the

model in S7). In specific, the post-hoc comparisons demonstrated that parents of TD children generated longer attention on target objects than parents of children with ASD did,  $\beta = 59.83$ ,  $t = 3.57$ ,  $p < .001$ , while there was no group difference in other ROIs (see Figure 4B). To be noted, the follow-up ROI contrasts revealed various distribution by groups: (1) among parents of children with ASD, the duration of attention on parent's own hands was significantly less than on objects ( $\beta = -58.14$ ,  $t = -3.91$ ,  $p = .002$ ) than faces ( $\beta = -52.20$ ,  $t = -3.51$ ,  $p = .006$ ), while attention on target objects and child's face were comparable ( $\beta = 5.94$ ,  $t = 0.40$ ,  $p = .978$ ); (2) among the parents of TD children, the duration of attention on objects was significantly longer than the other three ROIs (see also all the contacts in S8).

### Parent's gaze behaviors: Gaze type

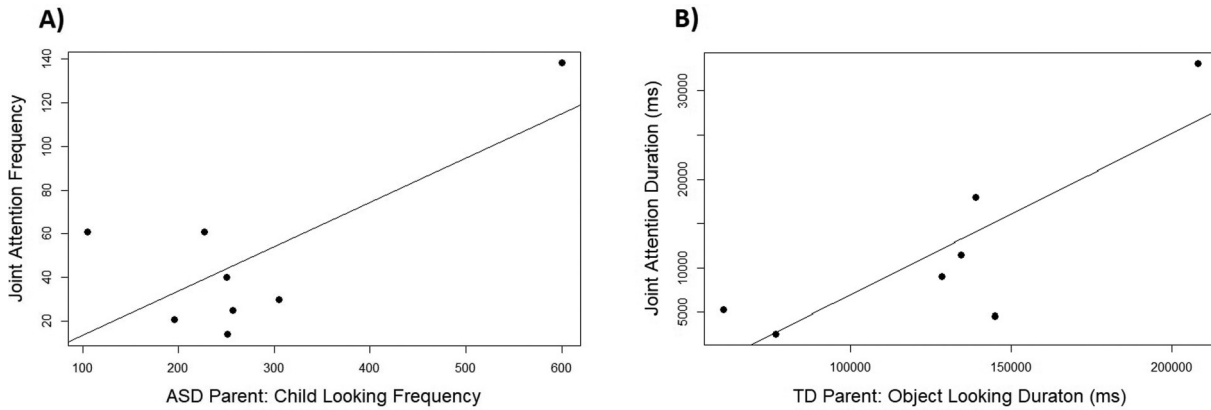
We then examined the types of gaze behaviors parents displayed to ROIs only during the play session; the statistics summary on both frequency and duration measurements are listed in Table 4. First, the frequency model with fixed effects of group and type of gaze and a random effect of participant revealed no group differences on parent's attention,  $\beta = 105.45$ ,  $t = 1.44$ ,  $p = .159$ ; the type of gaze accounts for gaze frequency, with a significant effect of SA,  $\beta = -175.37$ ,  $t = -2.79$ ,  $p = .010$  and a significant effect of saccades,  $\beta = 285.25$ ,  $t = 4.54$ ,  $p < .001$  (also see S9). Similar to the children's frequency distribution on various gaze types, parents of both groups also showed a similar distribution order, that saccades represented the highest frequency, followed by fixations and SA as the least, and all the comparisons were significant (see Figure 4C).

Notably, the duration model with fixed effects of group and type of gaze and a random effect of participants found two interaction effects: (1) a significant interaction between TD group and SA,  $\beta = -64.10$ ,  $t = -2.94$ ,  $p = .007$ ; and (2) a significant interaction between TD group and saccades,  $\beta = -46.16$ ,  $t = -2.12$ ,  $p = .044$  (see the full model in S10). To better interpret the interaction effect, the post-hoc Tukey contrasts revealed that parents of TD generated significantly longer fixations than parents of children with ASD did,  $\beta = 54.80$ ,  $t = 2.96$ ,  $p = .006$ . Besides, both groups generated similar duration distribution as follows, from greatest to least: fixations, saccades, and SA by the least. Specifically, parents of both groups had significantly longer fixations, followed by comparable saccades and SA (see Figure 4D).

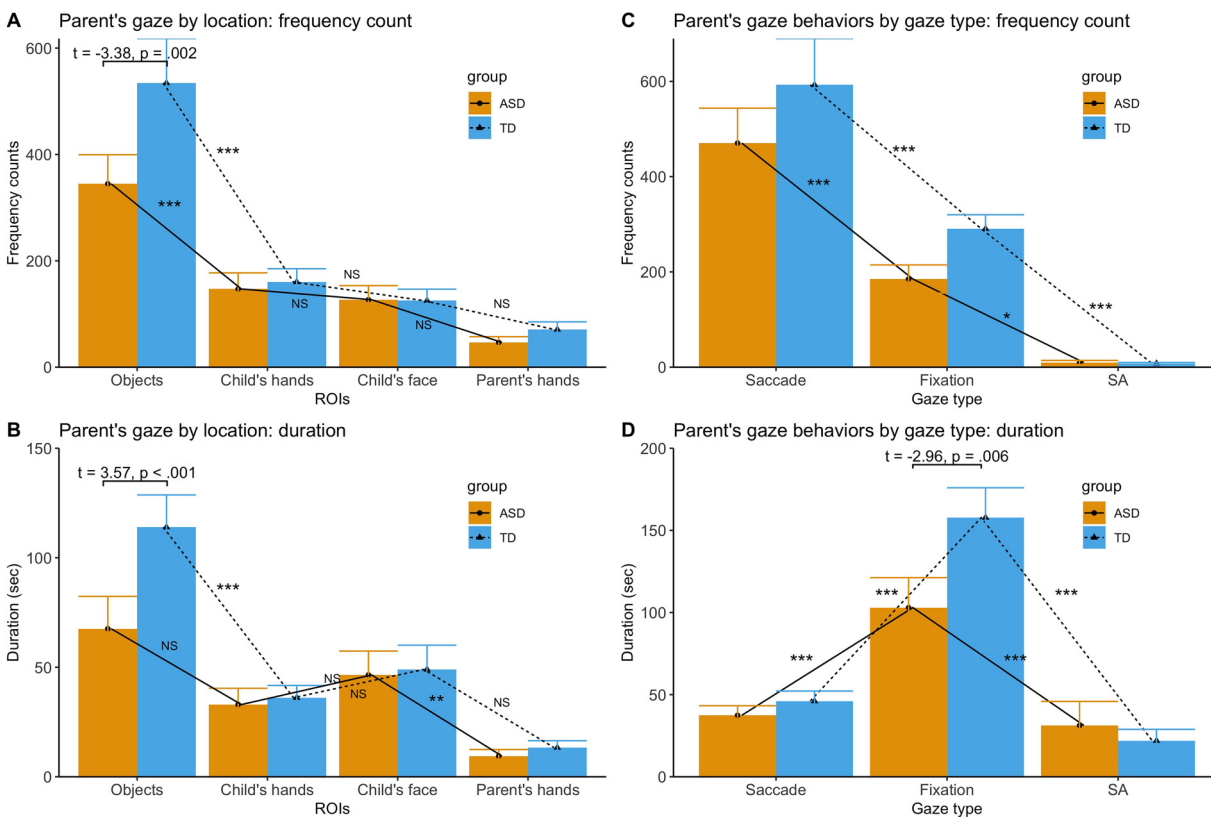
**Parent's gaze behaviors: Gaze shifting**

We also counted parents' gaze shifting to objects. The generalized linear model fitting with the quasi-Poisson distribution was used to predict the likelihood of gaze shifting as a function of groups and prior gaze locations. Table 5

summarizes the model in a log scale and reveals the main effect of groups that parents of TD children were significantly more likely to shift their gaze toward the target objects than parents of children with ASD did,  $\beta = 0.38$ ,  $t = 2.35$ ,  $p = .024$ . In addition, parents of both groups shared a similar shifting trend:



**Figure 3.** The frequency distribution of children's gaze shifting to objects (A) and parents' gaze shifting to objects (B). A, No difference between children with and without ASD in the amount of gaze shifting; children in both groups were significantly more often to shift attention when the prior location was the parent's hands than the child's hands, and the parent's face was the least. B, Parents of TD children had significantly more gaze shifting than parents of children with ASD did; parents in both groups had significantly more shift to objects when the prior location was child's hands, followed by comparable amounts of shifting from the child's face and the parent's hands.



**Figure 4.** The distribution of parents' gaze behaviors. (A) Parents' gaze by ROIs counted by frequency. (B) Parents' gaze by ROIs counted by duration (sec). (C) Parent's gaze by gaze type counted by frequency. (D) Parent's gaze by gaze type counted by duration (sec). Significance levels: \* < .05, \*\* < .01, \*\*\* < .001.

**Table 4.** Means and standard deviations (SD) of frequency (number of discrete looks) and duration (ms) of observed parent visual experiences by group, averaged across participants within the ASD and TD groups.

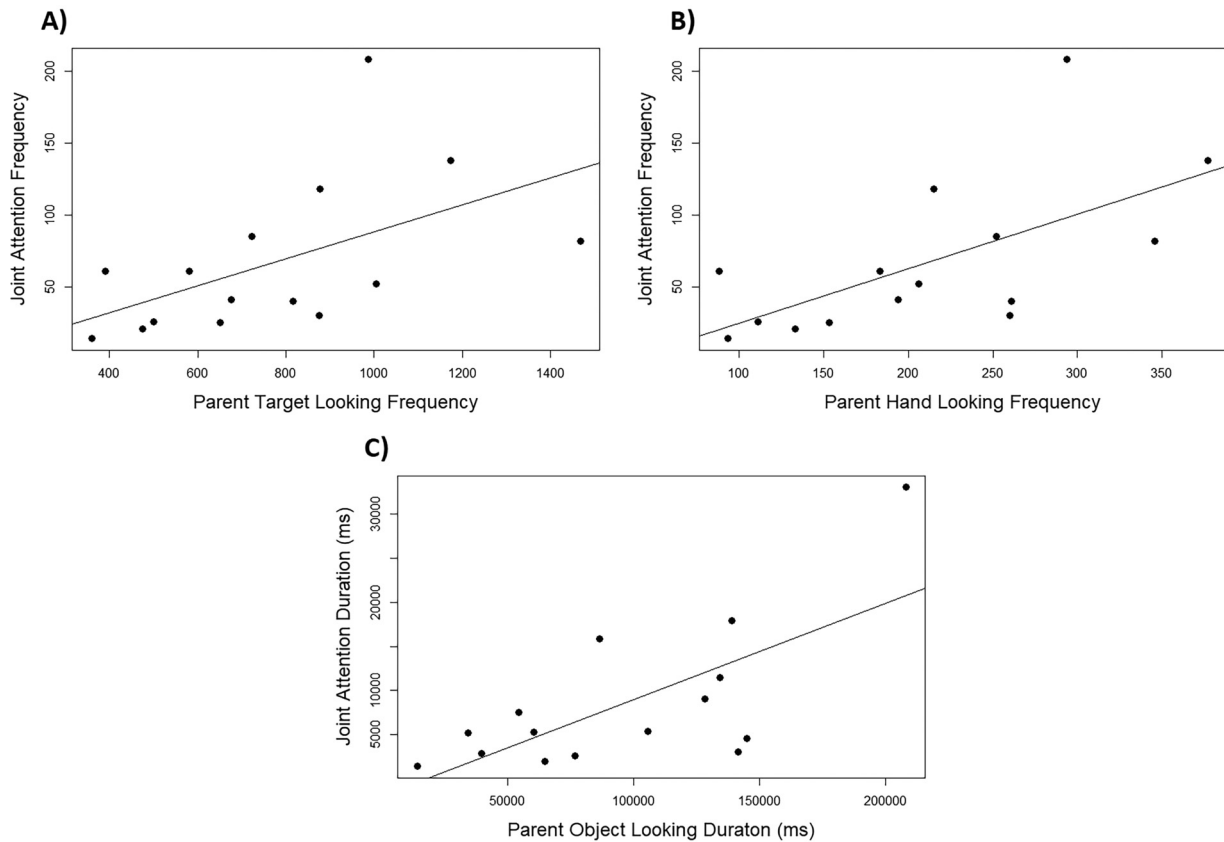
Parent behavior	Group	Frequency <i>M</i> (SD)	<i>p</i> value	Duration <i>M</i> (SD)	<i>p</i> value
Attention to Objects	ASD	344.9 (154.6)	.08	67612.8 (41743.8)	<b>.03</b>
	TD	534.4 (22.06)		127440.9 (48436.8)	
Attention to Child Face	ASD	126.9 (74.7)	.96	61676.6 (50537.4)	.71
	TD	125.1 (57.0)		49014.4 (29355.0)	
Attention to Child Hands	ASD	147.0 (85.7)	.74	32991.8 (20980.3)	.74
	TD	160.1 (66.1)		36097.3 (14815.4)	
Attention to Parent Hands	ASD	46.5 (30.7)	.20	9476.0 (8426.6)	.39
	TD	71.0 (37.7)		13341.4 (8300.8)	
Saccades	ASD	470.4 (207.7)	.33	37459.1 (16384.1)	.32
	TD	592.7 (256.8)		46095.9 (16119.0)	
Fixation	ASD	185.1 (83.6)	<b>.03</b>	103014.0 (51578.7)	.05
	TD	290.6 (78.0)		46095.9 (16119.0)	
SA	ASD	9.8 (12.4)	.65	31284.0 (41271.3)	.58
	TD	7.4 (6.0)		21982.7 (18305.4)	
Saccades to Objects	ASD	259.3 (112.3)	.22	20196.0 (9235.1)	.20
	TD	367.1 (190.7)		27738 (11874.4)	
Fixation to Objects	ASD	84.1 (48.5)	<b>.01</b>	43469.1 (27241.5)	<b>.02</b>
	TD	163.0 (52.5)		86903.1 (32893.8)	
SA to Objects	ASD	1.5 (2.8)	.19	3947.6 (7684.3)	.16
	TD	4.3 (4.6)		12799.3 (13602.6)	
Saccades to Child Face	ASD	63.9 (55.5)	.91	5560.5 (4426.3)	.98
	TD	61.0 (40.0)		6505.3 (3389.2)	
Fixation to Child Face	ASD	56.4 (26.9)	.72	34488.8 (18327.3)	.78
	TD	62.3 (34.3)		37587 (23543.6)	
SA to Child Face	ASD	6.6 (10.9)	.26	21627.4 (37325.7)	.28
	TD	1.9 (2.1)		5822.1 (7818.4)	
Saccades to Child Hands	ASD	112.3 (67.6)	.96	8889.4 (4841.5)	.86
	TD	110.7 (51.9)		8499.9 (3214.7)	
Fixation to Child Hands	ASD	33.3 (20.2)	.14	18665.6 (13267.5)	.33
	TD	48.4 (17.2)		24839.6 (10353.2)	
SA to Child Hands	ASD	1.5 (1.5)	.48	5436.8 (5618.3)	.27
	TD	1.0 (1.2)		2757.9 (3046.6)	
Saccades to Parent Hands	ASD	35.0 (22.4)	.20	2813.3 (1749.5)	.20
	TD	53.4 (29.4)		4242.9 (2210.7)	
Fixation to Parent Hands	ASD	11.4 (10.4)	.35	6390.5 (6922.0)	.53
	TD	17.1 (12.5)		8495.1 (5820.5)	
SA to Parent Hands	ASD	0.1 (0.35)	.49	272.3 (770.0)	.50
	TD	0.3 (0.5)		603.4 (1032.1)	

**Table 5.** The model summary of parents' gaze shifting to target objects.

Term	Estimate	Standard error	Statistic	<i>p</i> value	Conf.low	Conf.high
(Intercept)	4.33	0.14	30.05	<.001	4.04	4.60
Group:TD	0.38	0.16	2.35	<b>.024</b>	0.06	0.70
Child's face to object	-0.59	0.19	-3.14	<b>.003</b>	-0.97	-0.23
Parent's hands to object	-0.89	0.21	-4.27	<b>&lt;.001</b>	-1.31	-0.49

the incident rate for child's hands-to-object was 1.80 times the incident rate for child's face-to-object ( $z = 3.14$ ,  $p = .005$ ), the incident rate for child's hands-to-object was 2.42 times the incident rate for

parent's hands-to-object ( $z = 4.27$ ,  $p < .001$ ), and the incident rate for child's face was 1.34 times the incident rate for parent's hands-to-object ( $z = 1.28$ ,  $p = .405$ ; also see Figure 4B).



**Figure 5.** Significant parent attention predictors for JA instances. (A) Parents' attention to all of the regions of interest (objects, parent hands, child face, and child hands) frequency as a significant predictor of A frequency. (B) Parents' hand-looking frequency as a significant predictor of JA frequency. (C) Parents' object-looking duration as a significant predictor of JA duration.

### Parental impact on joint attention

We examined how parent attention to ROIs was predictive of JA. Parents' attention to combined ROIs- ( $\beta = 3.11$ ,  $p = .04$ ) (See Figure 5A) and hand- ( $\beta = 1.04$ ,  $p = .01$ ) looking (See Figure 5B) frequency was significantly predictive of the frequency of JA experiences. Parents' object-looking duration ( $\beta = 4.36$ ,  $p = .004$ ) was significantly predictive of the duration of JA experiences (see Figure 5C).

Parents of ASD children's *child-looking* frequency was predictive of the frequency ( $\beta = .2$ ,  $p = .04$ ) of JA experiences (see Figure 6A) though this was not the case for the TD parents. Rather, parents of TD children's *object-looking* duration were significantly predictive of the duration of JA experiences,  $\beta = .18$ ,  $p = .02$  (see Figure 6B).

### Discussion

In the current study, we used head-mounted camera techniques during live, social interaction between children and their parents engaged in object play to explore the visual experiences of children with and without ASD and how they distribute their visual attention in relation to parent

visual experiences. Overall, we found similar looking behaviors in play contexts where parents are actively influencing a child's visual experiences but found parents of ASD children and parents of TD children do so by taking potentially different pathways.

First, despite the previous findings from well-controlled experiments and observational studies of parent-child interaction, we did not find differences in viewing experiences of children between ASD and TD groups: children with and without ASD had similar gaze distributions corresponding with the ROIs and types of gaze, and they both were most likely to follow parent's hands to shift attention to objects. These seemingly contradicting results suggest the possible impact of social context on children's early viewing experiences through parents' corresponding scaffolding. Thus, we addressed this potential impact of social context by observing parents' looking behaviors. The results point to the different attention focuses and gaze distributions between parents in the TD group and the ASD group. Specifically, parents of TD children primarily focused on objects in terms of increased gaze shifting and longer attention to objects compared to parents of children with ASD. Parents of children with ASD had

comparable amounts of attention to objects and their child's face, which can be due to their enhanced sensitivity to children's reactions—child's gaze direction—during social interactions. These differences in parental gaze patterns, combined with the similar visual experiences of children in TD and ASD groups, may indicate potentially different processes by which parents scaffold children's visual experiences—suggesting differential attention strategies used by parents from TD and ASD groups. We speculated that parents in the two groups pay attention to different information to support their child's viewing experiences, and thus the children's visual experiences were shown to be comparable despite the previous studies documenting different JA between ASD and TD groups which used computer screen-based tasks (Amso et al., 2014; Chita-Tegmark, 2016; Wang et al., 2015). Third, this speculation was addressed by testing the relationship between parents' looking behaviors and their JA experiences with their children. This is our third key finding. The results not only support the link between parents' gaze behaviors and children's gaze behaviors but also suggest that parents of children with and without ASD differentially support JA experiences for their children. We found that while TD parents' attention to objects significantly predicted JA experiences, ASD parents' attention to their children significantly predicted JA experiences. One may further speculate that while parents of TD children focus on demonstrating the objects clearly to the child (e.g., which object when, and how those objects are facing to the child), parents of ASD children may attend to whether or how the objects are perceived by the child (e.g., child's direction of looking). To be noted, the present findings highlight that children with and without ASD were both sensitive to parents' hands rather than parent's face before shifting their attention to objects, implying a correspondence between hand actions and gaze behaviors by parents and various scaffolding strategies to support constant and stable-looking experiences. A semi-longitudinal study (from 5 to 24 months old) with a head-mounted eye tracker in an object play context indicates that parents significantly change the way they support child's object viewing and that TD children maintained their effective viewing experiences and strengthen their SA throughout the development period (Burling & Yoshida, 2019; Sun & Yoshida, 2022). Taken together, all of these findings suggest that parents can use a variety of social cues that may work separately and together to accommodate the child's developmental changes and scaffold children's visual experiences (Sun et al., 2022).

The potential importance of social scaffolding in terms of early learning is only the beginning. However, the present study revealed the potential mechanism of the impact of task contexts in early visual experiences and unique pathways in which social scaffolding supports similar developmental experiences and/or achievements.

This work further illustrates that while it is important to continue investigating details about how social scaffolding shapes attentional, linguistic, and cognitive development and developmental significance, it is also critical for us to keep in mind whether these experiences are equally important across typically and children who have been diagnosed with ASD.

### *Limitations and future directions*

Although this study offers novel information about the visual experiences of children with and without ASD along with their parents, our findings must be considered in light of at least three important limitations. First, the small sample size and large age range. Given our exploratory study concerning visual experiences, our needed sample size was calculated using a few similar studies with TD children, and the study used a relatively wider age range. Given the nature of how play structures change by age, we controlled for age and found a significant effect of age on children's attention duration only. Therefore, the generalizability of the current study findings is limited to similar groups. Future studies with larger samples with more narrow age ranges will allow researchers to include precise measurements of children's developmental level (e.g., scores reflecting IQ, language levels, and ASD symptom severity), which would allow exploration of how these factors may or may not influence both child and parent behaviors and how their behaviors are related. For example, previous research has shown that parents of autistic children have differences in behaviors toward their children based on their child's ASD symptom severity (Beurkens et al., 2013; Kasari et al., 1988; Strid et al., 2013); further exploration of this in regard to visual attention may reveal a graded effect—variation among ASD children. Nonetheless, the current study extended the previous head-mounted camera studies with microlevel behavioral approaches to ASD samples and successfully generated 7,000–12,000 frames from collected samples of 20–40 parent–child dyads (Deák et al., 2018; Falck-Ytter, 2015; Falck-Ytter et al., 2015; Wass et al., 2018; Yoshida et al., 2019; Yu et al., 2019; Yurkovic et al., 2021) and revealed the potentially unique links between child's and parent's attention with two important groups.

Furthermore, we acknowledge that there are a number of factors influencing the current results. One factor is gender—like most studies, we had limited participation of females diagnosed with ASD; recent research has suggested that there are differences in visual attention between females diagnosed with ASD and males diagnosed with ASD that are not easily captured when using predominantly male samples (Harrop et al., 2018, 2019, 2020). Another factor is an individual's treatment history. For example, prior enrollment in applied behavior analysis (ABA) therapies may influence children's behaviors due to the emphasis

placed on attending to items of interest and other people (Martins & Harris, 2006; Murza et al., 2016); this may also explain children's similarities in attention to different ROIs. Further, parents' tightly coupled behaviors to their child's visual experiences may be different as a function of background factors, such as socioeconomic status and bilingualism (Abels & Hutman, 2015; Sun et al., 2022), and this is worthy of future exploration, as well.

With these factors in mind, future studies in this vein will continue increasing the clinical relevance of research results and allow for more individualized approaches to improving intervention efforts for children with ASD. Nonetheless, the present study made a critical first step forward in our understanding of the mechanisms underlying the different ways in which parental scaffolding may impact visual experiences for children with typical and atypical developmental trajectories. Targeted parental scaffolding—parental gaze behaviors—have been documented to create potential far-reaching effects. This study adds to the current literature by documenting the similar nature of visual experiences of children with and without ASD in everyday visual experiences and how such visual experiences are also tightly linked to parental social scaffolding—specifically, the study provides insights into the potential training ground for early attention development at home, which has potentially great relevance for developing new intervention approaches.

## Conclusion

In the current study, we used head-mounted camera systems to study child (and parent)-centered viewing and gaze behaviors of parent and child (ASD and TD groups) in parent-child interactive play and documented similarities and differences in the distribution of visual attention and the impact of parental gaze behaviors on visual attention behaviors among children with and without ASD. The results of this innovative study concerning the shared visual experiences of children with and without ASD revealed that their parents use differential attention strategies to support their children's visual experiences. Our work sheds light on the impact of social context and parental involvement on children's attention, which is a foundation of early learning, and ultimately may contribute to the development and improvement of parent-mediated interventions to support developmental outcomes.

## Declaration of conflicting interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article:

This work was supported by the Exploring Visual Fixation Patterns in Young Children as a Potential Biomarker for Autism Division of Research/Provost Faculty Research Invigoration Programs: High Priority Area Research Seed Grants (Accessible Healthcare) University of Houston, Social-Embodied Attention and Language Development in Young Children at risk for Autism Spectrum Disorder. CLASS Research Progress Grant, University of Houston.

## ORCID iD

Elizabeth Perkovich  <https://orcid.org/0000-0002-2368-5183>

## Supplemental material

Supplemental material for this article is available online.

## References

- Abels, M., & Hutman, T. (2015). Infants' behavioral styles in joint attention situations and parents' socio-economic status. *Infant Behavior & Development*, *40*(6), 139–150. <https://doi.org/10.1016/j.infbeh.2015.05.004>
- Amso, D., Haas, S., Tenenbaum, E., Markant, J., & Sheinkopf, S. J. (2014). Bottom-Up attention orienting in young children with autism. *Journal of Autism and Developmental Disorders*, *44*(3), 664–673. <https://doi.org/10.1007/s10803-013-1925-5>
- Beurkens, N. M., Hobson, J. A., & Hobson, R. P. (2013). Autism severity and qualities of parent-child relations. *Journal of Autism and Developmental Disorders*, *43*(1), 168–178. <https://doi.org/10.1007/s10803-012-1562-4>
- Borjon, J. I., Abney, D. H., Yu, C., & Smith, L. B. (2021). Head and eyes: Looking behavior in 12- to 24-month-old infants. *Journal of Vision*, *28*(1), 18. <https://doi.org/10.1167/jov.21.8.18>
- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>
- Burling, J. M., & Yoshida, H. (2019). Visual constancies amidst changes in handled objects for 5- to 24-month-old infants. *Child Development*, *90*(2), 452–461. <https://doi.org/10.1111/cdev.13201>
- Campbell, B. A., Hayne, H., Richardson, R., & Campbell, B. A. (2014). *Attention and information processing in infants and adults: Perspectives from human and animal research*. Taylor & Francis.
- Chita-Tegmark, M. (2016). Attention allocation in ASD: A review and meta-analysis of eye-tracking studies. *Review Journal of Autism and Developmental Disorders*, *3*(3), 209–223. <https://doi.org/10.1007/s40489-016-0077-x>
- Datavyu Team (2014). *Datavyu: A video coding tool*. Databrary Project.
- Deák, G. O., Krasno, A. M., Jasso, H., & Triesch, J. (2018). What leads to shared attention? Maternal cues and infant responses

- during object play. *Infancy*, 23(1), 4–28. <https://doi.org/10.1111/inf.12204>
- Devillez, H., Guyader, N., Curran, T., & O'Reilly, R. C. (2020). The bimodality of saccade duration during the exploration of visual scenes. *Visual Cognition*, 28(9), 484–512. <https://doi.org/10.1080/13506285.2020.1830325>
- Doussard-Roosevelt, J. A., Joe, C. M., Bazhenova, O. V., & Porges, S. W. (2003). Mother-child interaction in autistic and nonautistic children: Characteristics of maternal approach behaviors and child social responses. *Development and Psychopathology*, 15(2), 277–295. <https://doi.org/10.1017/s0954579403000154>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test, fourth edition*. NCS Pearson.
- Falck-Ytter, T. (2015). Gaze performance during face-to-face communication: A live eye tracking study of typical children and children with autism. *Research in Autism Spectrum Disorders*, 17, 78–85. <https://doi.org/10.1016/j.rasd.2015.06.007>
- Falck-Ytter, T., Bolte, S., & Gredeback, G. (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders*, 5(1), 28. <https://doi.org/10.1186/1866-1955-5-28>
- Falck-Ytter, T., Carlström, C., & Johansson, M. (2015). Eye contact modulates cognitive processing differently in children with autism. *Child Development*, 86(1), 37–47. <https://doi.org/10.1111/cdev.12273>
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur communicative developmental inventories: User's guide and technical manual*. Singular Publishing Group.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Brookes.
- Freeman, S., & Kasari, C. (2013). Parent-child interactions in autism: Characteristics of play. *Autism*, 17(2), 147–161. <https://doi.org/10.1177/1362361312469269>
- Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, 69(2), 133–149. <https://doi.org/10.1006/jecp.1998.2438>
- Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and 7-month-old infants' memory for arbitrary syllable-object relations. *Infancy*, 2(2), 219–231. [https://doi.org/10.1207/S15327078IN0202\\_7](https://doi.org/10.1207/S15327078IN0202_7)
- Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to maternal multimodal naming by 6- to 8-month-old infants and learning of word-object relations. *Infancy*, 9(3), 259–288. [https://doi.org/10.1207/s15327078in0903\\_1](https://doi.org/10.1207/s15327078in0903_1)
- Harrop, C., Jones, D. R., Sasson, N. J., Zheng, S., Nowell, S. W., & Parish-Morris, J. (2020). Social and object attention is influenced by biological sex and toy gender-congruence in children with and without autism. *Autism Research*, 13(5), 763–776. <https://doi.org/10.1002/aur.2245>
- Harrop, C., Jones, D., Zheng, S., Nowell, S., Boyd, B. A., & Sasson, N. (2018). Circumscribed interests and attention in autism: The role of biological sex. *Journal of Autism and Developmental Disorders*, 48(10), 3449–3459. <https://doi.org/10.1007/s10803-018-3612-z>
- Harrop, C., Jones, D., Zheng, S., Nowell, S., Schultz, R., & Parish-Morris, J. (2019). Visual attention to faces in children with autism spectrum disorder: Are there sex differences? *Molecular Autism*, 10(1), 28. <https://doi.org/10.1186/s13229-019-0276-2>
- Henderson, B. B. (1981). Exploration by preschool children: Peer interaction and individual differences. *Merrill-Palmer Quarterly of Behavior and Development*, 27(3), 241–255.
- Horst, J. S., Samuelson, L. K., Kucker, S. C., & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition*, 118(2), 234–244. <https://doi.org/10.1016/j.cognition.2010.10.015>
- Jellema, T., Lorteije, J., Van Rijn, S., Wout, M., Haan, E., Engeland, H., & Kemner, C. (2009). Involuntary interpretation of social cues is compromised in autism spectrum disorders. *Autism Research*, 2(4), 192–204. <https://doi.org/10.1002/aur.83>
- Jordan, R. (2003). Social play and autistic Spectrum disorders: A perspective on theory, implications and educational approaches. *Autism*, 7(4), 347–360. <https://doi.org/10.1177/1362361303007004002>
- Kasari, C., Sigman, M., Mundy, P., & Yirmiya, N. (1988). Caregiver interactions with autistic children. *Journal of Abnormal Child Psychology*, 16(1), 45–56. <https://doi.org/10.1007/BF00910499>
- Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development*, 32(4), 437–444. <https://doi.org/10.1016/J.INFBEH.2009.07.003>
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223. <https://doi.org/10.1023/A:1005592401947>
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule* (2nd ed.). Western Psychological Services.
- MacNeill, L. A., Fu, K. A., Buss, & Pérez-Edgar, K. (2021). *Do you see what I mean?: Using mobile eye tracking to capture parent-child dynamics in the context of anxiety risk*. *Development and Psychopathology*. Cambridge Core. <https://www.cambridge.org.ezproxy.lib.uh.edu/core/journals/development-and-psychopathology/article/doi-you-see-what-i-mean-using-mobile-eye-tracking-to-capture-parentchild-dynamics-in-the-context-of-anxiety-risk/32732C037360DCAE1879786204BC0C82>.
- Martins, M. P., & Harris, S. L. (2006). Teaching children with autism to respond to joint attention initiations. *Child & Family Behavior Therapy*, 28(1), 51–68. [https://doi.org/10.1300/J019v28n01\\_04](https://doi.org/10.1300/J019v28n01_04)
- Murias, M., Major, S., Davlantis, K., Franz, L., Harris, A., Rardin, B., Sabatos-DeVito, M., & Dawson, G. (2018). Validation of eye-tracking measures of social attention as a potential



- biomarker for autism clinical trials. *Autism Research*, 11(1), 166–174. <https://doi.org/10.1002/aur.1894>
- Murza, K. A., Schwartz, J. B., Hahs-Vaughn, D. L., & Nye, C. (2016). Joint attention interventions for children with autism spectrum disorder: A systematic review and meta-analysis. *International Journal of Language & Communication Disorders*, 51(3), 236–251. <https://doi.org/10.1111/1460-6984.12212>
- Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M., & Rogers, S. J. (2008). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism*, 12(5), 457–472. <https://doi.org/10.1177/1362361308096402>
- Plesa Skwerer, D., Jordan, S. E., Brukilacchio, B. H., & Tager-Flusberg, H. (2016). Comparing methods for assessing receptive language skills in minimally verbal children and adolescents with autism spectrum disorders. *Autism*, 20(5), 591–604. <https://doi.org/10.1177/1362361315600146>
- Smith, Y., Yoshida, C., & Fausey, H., & M, C. (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. <https://doi.org/10.1080/15248372.2014.933430>
- Strid, K., Heimann, M., & Tjus, T. (2013). Pretend play, deferred imitation and parent-child interaction in speaking and non-speaking children with autism. *Scandinavian Journal of Psychology*, 54(1), 26–32. <https://doi.org/10.1111/sjop.12003>
- Sun, L., Griep, C. D., & Yoshida, H. (2022). Shared multimodal input through social coordination: Infants with monolingual and bilingual learning experiences. *Frontiers in Psychology*, 13, 745904. <https://doi.org/10.3389/fpsyg.2022.745904>
- Sun, L., & Yoshida, H. (2022). Why the parent's gaze is so powerful in organizing the infant's gaze: The relationship between parental referential cues and infant object looking. *Infancy*, 27(4), 780–808. <https://doi.org/10.1111/inf.12475>
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57(6), 1454–1463. <https://doi.org/10.2307/1130423>
- van Beers, R. J. (2007). The sources of variability in saccadic eye movements. *The Journal of Neuroscience*, 27(33), 8757–8770. <https://doi.org/10.1523/JNEUROSCI.2311-07.2007>
- Vivanti, G. (2020). Ask the editor: What is the most appropriate way to talk about individuals with a diagnosis of autism? *Journal of Autism and Developmental Disorders*, 50(2), 691–693. <https://doi.org/10.1007/s10803-019-04280-x>
- Wang, S., Jiang, M., Duchesne, X. M., Laugeson, E. A., Kennedy, D. P., Adolphs, R., & Zhao, Q. (2015). Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. *Neuron*, 88(3), 604–616. <https://doi.org/10.1016/j.neuron.2015.09.042>
- 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), 14–14. <https://doi.org/10.1111/desc.12667>
- Williams, K. (2007). *Expressive vocabulary test* (2nd ed. (EVT-2)). AGS.
- Yoshida, H., Cirino, P., Mire, S., & Burling, J. (2019). Parents' gesture adaptations to children with autism spectrum disorders. *Journal of Child Language*, 47(1), 1–20. <https://doi.org/10.1017/S0305000919000497>
- Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy*, 13(3), 229–248. <https://doi.org/10.1080/15250000802004437>
- Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244–262. <https://doi.org/10.1016/j.cognition.2012.06.016>
- 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), 79659–79659. <https://doi.org/10.1371/journal.pone.0079659>
- Yu, C., & Smith, L. B. (2017). Multiple sensory-motor pathways lead to coordinated visual attention. *Cognitive Science*, 41(S1), 5–31. <https://doi.org/10.1111/cogs.12366>
- 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), 12735. <https://doi.org/10.1111/desc.12735>
- 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. *Scientific Reports*, 11(1), 3578. <https://doi.org/10.1038/s41598-021-81102-0>
- Yurkovic-Harding, J., Lisandrelli, G., Shaffer, R., Dominick, K., Pedapati, E., Erickson, C., Yu, C., & Kennedy, D. (2022). Children with ASD establish joint attention during free-flowing toy play without face looks. *Current Biology*, 32(12), 2739–2746.e4. <https://doi.org/10.1016/j.cub.2022.04.044>