Examining Social Attention and Verbal Exchange in Children with ASD around a High-Interest Object during Real-Life Interaction

by

Zhizun (John) Yang

B.A., Simon Fraser University, 2016

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
Educational Psychology Program
Faculty of Education

SIMON FRASER UNIVERSITY
Spring 2021

Declaration of Committee

Name:	**Zhizun (John) Yang**
Degree:	**Master of Arts**
Thesis title:	**Examining Social Attention and Verbal Exchange in Children with ASD around a High-Interest Object during Real-Life Interaction**
Committee:	**Chair:** Robert Williamson Assistant Professor, Education
	Elina Birmingham Supervisor Associate Professor, Education
	Maureen Hoskyn Committee Member Associate Professor, Education
	Lucy LeMare Examiner Professor, Education

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Abstract

Previous research has demonstrated that social attention is reduced in school-aged children with Autism Spectrum Disorder (ASD) when compared with typical developing (TD) children. However, the majority of studies on this topic focused on computer-based stimuli. How school-aged children with ASD attend to social stimuli during real-life interactions is not well understood. The current exploratory study aims to investigate, under the social motivation and CI-distractor framework, how social attention and verbal exchange change in the presence of objects that are of high interest to children with ASD. Nonparametric analysis revealed that the ASD group spent significantly less time viewing the experimenter's full face than the TD group, and they spent more time speaking than children in the TD group. These provide support for the CI-distractor hypothesis, although future research is needed to confirm the reported pattern of results using an experimental design.

Keywords: social attention; verbal exchange; real-life interaction; high interest object; Autism

Dedication

This paper is dedicated to my mother, Xueping Xie, who has been supportive without the need to say a word, and I have no other big word to say other than saying thank you for all these years.

Acknowledgements

I would like to specifically extend my gratitude to my senior supervisor Dr. Elina Birmingham, for being supportive and understanding throughout my MA thesis project. This thesis can't be done without your guidance and patience.

I would also like to give my thanks to Dr. Maureen Hoskyn for helping out with my thesis and my coursework during my graduate degree.

Thank you also to Joshua Wiebe, and Vida Nacho for volunteering their time in this project.

Lastly, this is for those who believe in me when I am still struggling with my life and emotion. Without your company, from near and far, I would not have made it through this crazy year.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ASD	Autistic Spectrum Disorder
TD	Typically Developing

Chapter 1. Introduction

The goals of the present study were threefold. First, this study aimed to understand how the social attention of children with ASD differs from that of typically developing (TD) children when interacting with an adult around a high-interest object. This study also explored group differences in social attention when dyads interacted around a low-interest object. Second, this study aimed to uncover how the length of verbal exchanges differs between children with ASD and TD children in the presence of a high-interest object. Third, this study examined how social attention differs as function of speech state (participant speaking vs. adult speaking), and to determine if group differences in speech state account for observed group differences in social attention. Below, the relevant literature, theoretical frameworks and alternative hypotheses are described.

1.1. Literature Review

Social attention is generally referred to as overt visual attention allocated to social stimuli such as human face or body (Ames, et al., 2010; Salley and Colombo, 2016). Paying attention to social stimuli is a critical skill for everyday life, it informs us our social interactions and surroundings. A wealth of research has already shown that individuals preferentially attend to social stimuli (e.g., face) when they are presented (Birmingham et al., 2008; Gliga & Csibra, 2007; Vuileumier, 2002; Wilson et al., 2010). It is thought that this attentional bias towards faces enables individuals to develop a range of social cognitive and emotional skills (e.g., Grelotti et al., 2002; Schultz 2005). This fundamental ability has been demonstrated to be impaired in the population with Autistic Spectrum Disorder in numerous studies (e.g., Klin et al., 2002, Ozonoff et al., 2010; Shi et al., 2015). For example, Klin et al. (2002), in an early eye-tracking study that presented participants with ASD a video clip, found that compared to controls, children with ASD fixated more on the object areas and less on the actors face and demonstrated an atypical pattern of attention such as focusing more on the actors mouth region than the eye regions. Ozonoff et al. (2010) examined gaze to faces and social smiles in toddler with ASD and found that between 12 to 18 months of age, there was a significant decrease in gaze to faces and social smile compared to their TD controls. Over the past decade, a number of other researchers have detailed similar results detecting reduced social attention in children with ASD compared to their TD counterparts (Bhat et al., 2010; Moore et al., 2012; Riby et al., 2012). Further, this atypical attentional pattern was also found across different age groups in individuals with ASD (for review see, Chita-Tegmark 2016; Guillon et al. 2014). However, what leads to such reduced attention in individuals with ASD is still a controversial topic.

1.1.1. CI-distractor Hypothesis

One possible account for the attenuated social attention is that children with ASD have an attentional bias towards non-social stimuli, especially stimuli with high saliency such as circumscribed interest (CI) related objects. Circumscribed interest refers to specific restricted, repetitive pattern of behavioral interests and activities that are part of the diagnostic criteria of ASD (American Psychiatric Association [APA], 2013). Such behaviours are usually manifested as intense interests in certain objects or topics, and often at the exclusion of other activities (see Klin et al., 2007) Common topics that belong to CI categories include Japanese animation (Pokémon), vehicles, electronics, dinosaurs, and transportation systems. (South et al., 2005, Sasson et al., 2008). According to the *CI-distractor hypothesis* (Sasson et al., 2008), when CI-related objects are present together with social stimuli, they will likely become distractors to social communication and interaction for children with ASD due to their high interest value.

To examine social attention behaviors in individuals with ASD in the presence of CI-related stimuli, Sasson and Touchstone (2014) presented young children with ASD and their age-matched TD controls with computer-based picture of either non-CI items (e.g., clothes) or CI-related items (e.g., train), alongside pictures of human faces as social stimuli. They found that children with ASD spent less time looking at the social stimuli than their TD counterparts when a CI-related object was simultaneously presented, but this reduction in attention to social stimuli in the ASD group was not evident in the presence of *non-CI* objects. That is, the reduction in social attention in ASD relative to TD children was contextually dependent on the presence of a CI object that competed for attention. Complementary evidence comes from research on neural activity in response to CI-related objects. Grelotti et al. (2005) examined fMRI activation in the amygdala, a region thought to be important for estimating the salience of relevance of percepts, of a child with ASD who had a fascination with Digimon characters. The child with ASD showed heightened neural activation in the amygdala while processing Digimon characters relative to unfamiliar faces. In contrast, the TD comparison child showed heightened amygdala activation to unfamiliar faces relative to Digimon. The authors concluded that CI-related objects may have a particularly high reward value for individuals with ASD, and much more so than human faces. The implication of these results and those of Sasson and colleagues (Sasson et al., 2005, 2011; Sasson & Touchstone, 2014) is that CI-related stimuli may disproportionately capture the attention of children with ASD early in development, detracting attention away from social input that is necessary for specialization of face processing and social cognition brain networks needed for typical development.

1.1.2. Social Motivation Account in ASD

Alternatively, other researchers posit that social motivation is more globally reduced in ASD and that this reduced motivation is not contingent on the presence of non-social (particularly CI) stimuli competing for attention. According to the social motivation account (see Chevallier et al., 2012, p.1), social attention deficits that characterize individuals with ASD are a result of pervasive and early deficits in social motivation, defined as “a set of psychological dispositions and biological mechanisms biasing the individual to orient to the social world (social orienting), seek out and take pleasure in social interactions (social reward), and to work to foster and maintain social bonds (social maintaining).” Evidence to support this framework would include findings that reduced social attention in ASD persists across contexts. For example, in contrast to Sasson and Touchstone (2014), Harrison and Slane (2020) found that children with ASD showed reduced attention to faces relative to TD individuals, regardless of the type of object distractor (an object belonging to a general CI category, object belonging to a CI category personalized to the participant’s interests, or a control object) that was presented alongside the face. Indeed, a surprising finding was that only the TD participants’ attention to faces was reduced by the presence of a CI object (either general or personalized) relative to control objects. These findings suggest that reduced social attention in ASD is a more global phenomenon that is not contextually dependent, in line with the social motivation account.

1.1.3. The issue of ecological validity

What may account for the contradictory evidence for *CI-distractor* and *social motivation* accounts? In addition to the above conflicting evidence between Sasson & Touchstone (2014) and Harrison & Slane (2020), the field is full of examples of contradictory evidence for both reduced social attention in ASD and its context-dependency (see Guillon et al., 2014, for a thorough review). Complicating things further, some researchers have found no social attention impairments in ASD, regardless of context (Guillon et al., 2014). Some researchers have suggested that the reason for these mixed results may be due to the use of different stimuli in different studies. For example, Noris et al. (2012) pointed out that many researchers prefer to use static social stimuli because of their repeatability in a controlled lab setting. However, static stimuli have not been consistent in producing group differences in social attention between children with ASD and TD children (Chevallier et al., 2015). Chevallier and her colleagues compared the respective effects of static, dynamic, and interactive social stimuli (a video clip that shows a social interaction) on eliciting group differences in visual exploration task. She found that only the video that depicted social interactions elicited significant group differences (ASD vs TD) in social attention.

Other researchers using video stimuli have not produced the expected results in terms of eliciting group differences in social attention. For example, Chawarska et al. (2012) compared the social attention of children with and without ASD while viewing video clips that featured 4 different social contexts. The 4 different contexts included: 1) an actress saying “uh-oh” and trying to look at toys placed at each corner of the camera; 2) the actress making a sandwich; 3) the actress moving one of the toys; and 4) the actress trying to “interact” with the participants by biding for their attention. They found that only the dyadic attention bid condition elicited group differences in social attention between the ASD and TD groups. Although all stimuli used in the research were embedded in videos, the findings of this study indicated that exposure to videos alone did not result in differences in social attention between children with and without ASD. The significant result found only for the dyadic bid condition may implicate the importance of actual naturalistic social interaction in studying social attention. As Chita-Tegmark et al. (2016) pointed out, computer-based stimuli as a representation of social behaviours is not the same as social behaviours during an interaction with a real person. The over-reliance on this type of stimuli may hinder us from understanding social attention in a realistic setting. Therefore, conducting studies with higher ecological validity can minimize the gap between children's social behaviours in real life and those shown in a laboratory-controlled environment (Guillon et al., 2014; Klin et al., 2003).

1.1.4. Examining the CI-distractor and Social motivation Hypotheses in Real-Life Settings

With respect to the *CI-distractor hypothesis*, thus far, few researchers have attempted to study how the presence of non-social objects may impact the social attention of children with ASD in a naturalistic setting (Noris et al., 2012). Noris et al. conducted a face-to-face study of social attention and found that visual attention to faces in children with ASD was shorter compared to TD controls, when participants were able to play with non-social objects during a non-verbal social interaction. However, the objects used in the study were not designed by the researchers to examine the effect of CI on social attention, therefore it was unknown how rewarding the non-social object was to both of the diagnostic groups, nor was it clear whether it has any relationship with participants' social attention.

Another study conducted by Nadig et al. (2010) examined the effect of CI as a conversation topic on eye-gaze and conversation contingency during a face-to-face interaction with an experimenter. Contradictory to the view that CI distracts attention from social stimuli, Nadig et al. found that time spent looking at the face of their conversational partner increased for both children with ASD and TD children when speaking about the CI topic relative to a non-CI topic. Furthermore, no group differences in social attention were found in either topic condition. In addition, Nadig et al. (2010) found that when talking about a CI-related topic, children with ASD showed more one-sided non-contingent utterances than the TD controls. The authors reasoned that for children with ASD, utterances related to their CI-topic were so well-practiced that they became effortless, and therefore simultaneously freed up cognitive resources for social attention. In contrast, the authors speculated that for TD children, increased looking to the partner's face's reflected increased engagement with the partner during the conversation about a high interest topic. Critically, however, that Nadig et al. (2010) only employed CI as a conversational topic rather than presenting real objects during the social interaction. Thus, the present study extends Nadig et al.'s research by adding physical CI-related objects which the participants can interact with during their conversation with the experimenter.

Although the body of literature investigating the social motivation hypothesis has grown in recent years, to the best of my knowledge, no study has directly examined how social attention is affected by high-interest object in children with ASD in a real-life setting based on this framework. A recent review of the literature on the social motivation hypothesis in ASD found that the majority of the studies on this topic employed computer-based social stimuli such as pictures of faces or videos clips of social interactions (for review see, Bottini, 2018). As for the non-social stimuli, only 37% of the studies used non-social objects that are considered CI-related (e.g., pictures of cars or train) for individuals with ASD. Bottini's review indicated that findings were mixed and suggested social stimuli that are more naturalistic may assist in detecting differences not found in studies that only employed computer-based stimuli. Given this, the current research may add to the literature by examining social attention in children with ASD alongside with high interest object during a face-to-face interaction.

1.1.5. The present study

Hypotheses about the social attention of children with ASD during a real-life interaction with high vs. low interest objects

The first goal of the current study was to investigate the social attention behaviours of children with and without ASD while interacting with an experimenter in a face-to-face setting. Child-adult dyads interacted during two conditions: the first condition was an interaction around a preferred-object (high interest) chosen by the participants; the second condition was an interaction around a non-preferred-object (low interest) chosen by the experimenter. Additionally, this study aimed to discover how social attention in both groups would change over the course of conversation in the preferred-object condition.

Three alternative hypotheses emerge with respect to how social attention of children with ASD might differ from TD children during a real-life interaction around a high-interest objects vs. low interest object.

1. The *CI-distractor hypothesis* posits that high-interest objects function as distractors for individuals with ASD. Thus, relative to TD children, children with ASD should show reduced social attention when engaging in a conversation in the presence of a high-interest object. However, there should be no group difference in social attention when engaging in a conversation in the presence of a low interest object.

2. Alternatively, the *social motivation hypothesis* posits that individuals with ASD are less motivated to attend to social stimuli in general, and that this effect is not dependent on the presence of high-interest objects. Thus, social attention should be reduced among individuals with ASD relative to TD controls, regardless of whether they are interacting around a high or low interest object.

3. A third hypothesis, consistent with Nadig et al. (2010), is that there will be no group differences in social attention in either the high or low-interest object condition, but that social attention will increase for both groups while conversing about a high-interest object relative to a low-interest object.

Verbal exchanges and social attention

The second aim of this research was to explore how verbal exchange differs in conversations around high vs. low interest objects and whether conversational phase during social interaction has an impact on social attention in children with ASD relative to TD controls. Research on this topic is rare because the nature of the topic requires naturalistic settings (Hutchins & Brien, 2016; Freeth et al.,2013; Freeth & Bugembe, 2019). Based on the methodology developed by Tager-Flusberg and Anderson (1991), Nadig et al. (2010) investigated the contingency of verbal exchange and social attention around a CI-related topic. As mentioned previously, the authors found that children with ASD produced more one-sided speech, elaboration, and atypical utterances than the TD controls when the topic was of high interest. A study by Freeth and Bugembe (2019) explored the effect of conversational phases on social attention. They found that both TD adults and adults with ASD, spent more time looking at their partner's face while listening, compared to when participants were speaking. Although there is no established theory to guide the formation of a hypothesis, these studies suggest that the availability of cognitive resources is a possible explanation for their findings. In particular, Nadig et al. (2010) explained that because the speech of children with ASD about CI-related topic was well practiced, it freed up cognitive resources for other social behaviors such as producing longer utterances and increased social attention during the interaction. Likewise, Freeth and Bugembe (2019) suggested that their findings are consistent with past research showing that averting the eyes while speaking decreases cognitive load (Doherty-Sneddon & Phelps, 2005). It should be noted that this study did not attempt to examine verbal exchange directly. Because the content of the conversation was not analyzed, there was no reliable way to confirm whether the participants were actively listening to and comprehending the experimenter's speech, for example by providing contingent responses or responses indicating comprehension of the conversational partner's utterances. Thus, it is possible that interlocuters were passively listening to their partner's speech and were paying attention elsewhere during the interaction (Wang & Holland, 2014). Rather, the present study examined length of utterances for both the participants and the experimenter as a proxy for participants' inclination to speak during an interaction. According to Nadig et al. (2010), conversations become one-sided for children with ASD when the conversation topic is related to the CI. Therefore, in the present research, I hypothesized that, relative to the TD group, the ASD group would speak longer in high-interest condition.

I also examined if the participant's social attention behaviours differed when the experimenter was speaking vs. when the participant was speaking. As mentioned, Freeth and Bugemebe (2019) found that children with ASD tend to pay more attention to faces when they are listening rather than speaking to their interlocutors. Whereas Freeth and Bugemebe's study was designed specifically to elicit both speaking and listening phases (participants were asked a series of questions to which they had to listen and then respond), the present study examined a natural conversation that unfolded around an object of interest, with little imposed structure. Thus, while the present study cannot firmly distinguish between the conversational phases of speaking vs. listening, one might expect that participants were in a listening state when the experimenter was speaking. Thus, the tentative hypothesis explored was that when the experimenter was speaking, participants' attention to the experimenter's face would be higher than when the participant was speaking.

Chapter 2. Methods

2.1. Study design

In the current exploratory study, analyses of visual social attention, verbal exchange and social attention when either participants or experimenter is speaking were conducted. This is a mixed design study that involved both between and within-subjects designs for all three analyses. For social attention, the analysis examines how the presence of two different types of toys (preferred vs. non-preferred condition) would have an impact on social attention between two groups of school-age children (ASD vs. TD). An attempt was also made to examine how time would affect each group of participants' social attention behaviors. For verbal exchange, the analysis focuses on how the presence of toys (preferred vs. non-preferred condition) would have an effect on the length of utterance, defined as combined duration of speech during the entire analyzed video, from each interlocuter (participant speaking vs. experimenter speaking). Time was considered as a within-group variable to examine how it may have an effect on length of utterance. The third analysis examined social attention of children with and without ASD, anchored specifically to when the participant vs. experimenter was speaking. The within-subject variable was speaking state (participant speaking vs. experimenter speaking) in this analysis and how speaking states influenced social attention in each diagnostic group were investigated.

It should be emphasized that the manipulation of the preferred and non-preferred conditions was not set up ideally (see details in limitation in discussion), thus the current study did not intend to compare the two conditions directly. Instead, analyses on preferred condition and non-preferred conditions were run independently.

2.2. Participants

21 children with ASD (16 males and 5 females) and 32 typically developing (TD) children (20 males and 12 females) between the ages of 6 and 13 participated in the study as part of a summer camp at Simon Fraser University. The children were recruited through community advertisements, and the Autism and Developmental Disorder Lab's database. Participant groups were age and IQ matched (see Table 2-1

for participant information, final sample after exclusions). None of the participants included were reported to have visual or neurological impairments. The parent or guardian provided informed consent for participation. All procedures used were approved by the Simon Fraser University Research Ethics Board and were in accordance with the World Medical Association 2013 Declaration of Helsinki.

Diagnosis for participants with ASD was confirmed by review of a British Columbia (BC) clinical diagnostic report along with the Ministry of Child and Family Development ASD funding eligibility report. In the province of BC, substantial government funding is allocated to children with an ASD diagnosis, and there are rigorous standardized diagnostic practices that are required for a diagnosis. This includes diagnosis by a trained clinician who uses the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) and their clinical judgement to determine the diagnosis. To assess current level of ASD symptoms, parents completed the Autism Quotient (AQ) parent-report questionnaire (Baron-Cohen et al, 2001). As expected, AQ scores for children in the ASD group were significantly higher than those for the children in the TD group (see Table 2-1).

Table 2-1: Participant characteristics

		ASD	TD
AGE	**Mean**	10.01	9.50
	SD	.42	.30
	Range	7.14 – 12.77	6.79 – 12.52
IQ	**Mean**	104.00	107.33
	SD	3.57	2.38
	Range	74 – 138	87 – 135
AQ	**Mean***	31.50	15.07
	SD	5.49	6.03
	Range	19.0 – 39.0	6.0 – 31.0
Gender (M	**M:F**	15:5	19:11

Note: Age is in years; IQ was estimated through WASI-II, Full Scale-2; * indicates a significant ($p < .05$) group difference was found

One participant from the ASD group was excluded due to low IQ (IQ < 50). The experimenter had difficulties engaging in conversation for meaningful data. Another two participants from TD group were excluded because the experimenter accidentally displayed the preferred and non-preferred objects simultaneously (they were meant to be used in separate conditions). The final sample after these exclusions consisted of 20 ASD (15 males and 5 females) participants and 30 TD participants (19 males and 11 females). Independent t-tests were run on both age and full-scale IQ [Wechsler Abbreviated Scale of Intelligence 2nd edition; (WASI-II)] to determine whether there are any group differences for the final sample. No significant differences in age, $t(48) = -1.02$, p = .311; or in IQ, $t(48) = .81$, p = .423, were found.

2.3. Procedure, Apparatus and Materials

Nine (9) objects were placed in a display measuring approximately 60 cm (H) x 60 cm (W), see Figure 2-1. Five objects (i.e., Pokémon, dinosaur, train, spaceship, car) on the array were selected based on categories that were shown to be associated with circumscribed interests for individuals with ASD (Sasson et al., 2008). The other four objects (i.e., yarn, flower, plastic cup, doll) came from categories that are not commonly known to be preferred in this population. The objects were placed approximately 15 cm apart and remained in the same spatial locations of the display for every participant.

To measure participants' visual attention, SMI ETG 2.6 eye tracking glasses were used. The device was worn as a pair of glasses during the experiment and was equipped with two small cameras on the bottom rim which were used to capture users' eye movements and simultaneously map their gaze on to a scene video (head camera capturing the participant's point of view). A microphone on the edge of the rim recorded the participants' conversation with the experimenter (ETG manual, SMI). In addition to the eye tracking camera, a Logitech HD webcam was positioned in front and to the left of the participant and was aimed at the participant's partial profile as well as the experimenter's profile. Due to unreliable eye tracking data for the majority of participants, we coded only the head camera footage (see Coding Visual Attention below).

Before the study started, the experimenter explained the purposes of the study to the participant, by saying "We do lot of research here at SFU and we need to know which toys kids like the most. We're going to show you some toys and we want you to tell us which one you like the most." The experimenter then assisted the participants in putting on the eye-tracking glasses, and then explained what they were for, by saying "These goggles have a camera at the front and will record what you see. This will help our study because we will be able to see things from your point of view." Participants were seated directly across the table from the experimenter, at a distance of approximately 70cm (see Figure 2-2). A 5-point calibration was then performed on the eye tracker. After calibration, the experimenter led participants to the other side of the room and stood approximately 70cm away from the toy display. Masking tape was placed on the floor to mark the spot where the participant should stand. To ensure that participants did not pay attention to the stimuli beforehand, the array was covered by a cardboard. When the participant was ready, the experimenter removed the cardboard and instructed the participant that they could select one item they like the most by either pointing towards it or saying the name of the object out loud. After the experimenter removed the participant's preferred toy from the display, she then always choose the blue plastic cup on the shelf (except in two cases in which the participants chose the cup, the experimenter had to choose the ball of yarn) to be used in the second (non-preferred) condition.

The participants were then told that they would have a conversation about why each of them chose their selected items. Prior to the Preferred Condition, the experimenter hid her selected object such that the object selected by the participant was the only stimulus of focus. Participants were free to pick up or touch their chosen toy (see Figure 2-2). The experimenter then proceeded to ask the participant, "why do you like this toy"? If needed, the experimenter used prompts (e.g., "do you have one of these at home? How do you play with it?"). An attempt was made to engage the participant in a natural back-and-forth conversation about the toy for about 1 minute. Thus, the conversational content naturally varied between participants, as did the number of prompts required to elicit speech from each participant.

In the second (Non-preferred) phase, the experimenter swapped the objects so that experimenter's selected item would become the conversation topic. The experimenter held her chosen object in this condition. In this case, the conversation resumed when experimenter said, "I chose this because…" After the experimenter finished her explanation, she allowed the participant the chance to ask to follow up questions or provide comment; if this did not happen, the experiment ended. In all cases, the non-preferred condition was significantly shorter than the preferred condition (see Results) because most children did not ask to follow up questions after the experimenter finished her explanation for choosing the cup.

Because the preferred and non-preferred conditions were not adequately matched, nor was their order counterbalanced, the conditions were analyzed separately (as opposed to including preferred/non-preferred as an independent variable).

Figure 2-1: Camera view from the eye-tracker during object selection from the display of 9 items.

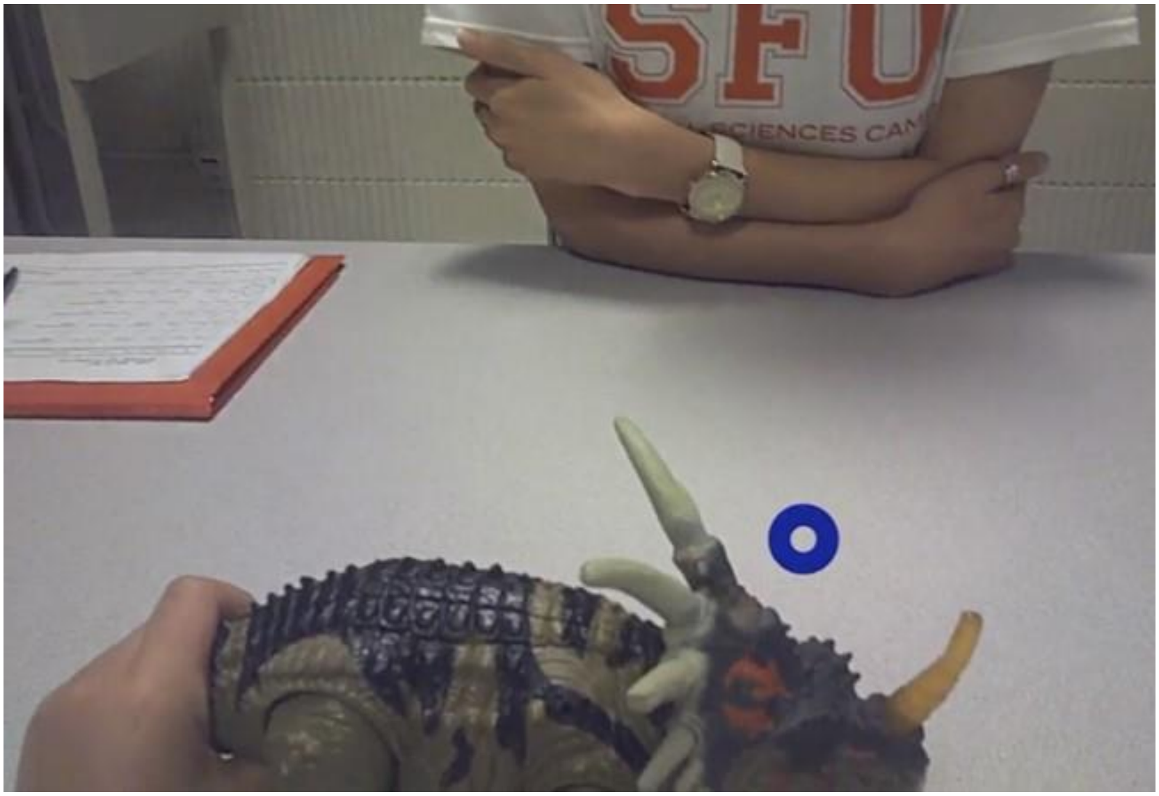

Figure 2-2: Camera view of the eye-tracker when participant was interacting with the experimenter.

2.4. Coding and data preparation

Fixation data were computed by SMI software: BeGaze. 2.0. However, for a large proportion (over 50%) of participants in both groups, the eye tracker produced unreliable data such that fixations were not detectable or were located outside of the head camera footage, and therefore not codable. In some cases, participants squinted during the experiment leading to data loss because the pupil was not detectable. In other cases, participants found the glasses uncomfortable to wear (e.g., too big for the younger participants; too hot due to the internal electronics heating up over time) and data loss occurred when they intentionally or unintentionally tried to adjust the glasses during the experiment. Due to unreliable eye tracking data for over half of the sample, I decided to code the head camera footage from the eye-tracking device to approximate participants' visual attention (see details in *Coding Visual Attention*, below).

Given there was no time limit for the experiment, each participant had different lengths of conversations. For the preferred condition, the mean duration of the videos across all participants was 177.48s, with the shortest video lasting 95.91s. As mentioned previously, conversations in the non-preferred condition were much shorter than in the preferred condition, lasting on average 35.62s on average, with the shortest video lasting 19.60s. To obtain consistency of conversation length across participants, I therefore decided to extract the first 60s of the videos from the preferred-toy condition and the first 20s of the videos from the non-preferred condition for further coding and analysis. With an attempt to examine the two conditions, and to look at the effects of time on visual attention behavior for each group, I further split the 60s clips form the preferred condition into three 20s segments. The start of the video segment in both preferred and non-preferred condition was defined by the video frame in which experimenter hid the selected object irrelevant to the current phase of conversation topic and it became invisible in the head camera view.

2.4.1. Coding Visual Attention

The primary goal of this study was to examine how participants allocated their visual attention towards experimenter's face during conversations about preferred and nonpreferred objects. While unfortunately the eye tracking data were not reliable, participants' point-of-view head camera footage provided useful information about where the participants were directing their attention. Previous findings suggest saccades are predictive of head movements (Land, 1987), and that they are highly correlated with each other in response to external cues (Khan et al., 2012). It was reasonable to expect that head movements would align with that of saccades, and therefore this coding method could be used as a proxy for visual attention. Head camera footage was coded into 4 discrete duration-based visual attention categories: 1) *Full Face*, when the experimenter's face and all the major facial features (i.e., eyes, nose, mouth) were clearly visible with in the frame; 2) *Face out of view*, where either the experimenter's face was completely out of frame (e.g. because the participant was looking down or the participant was holding and looking at their preferred object, totally blocking their view of the experimenter's face); 3) *Partial face*, only part of the experimenter's face was visible in the frame because of the camera angle, or because the child held the preferred object in a particular way. I also coded a small proportion of frames (9%, mean duration 1.4s) as "*Other situations",* where the experimenter's face was fully within the frame but her facial features were blocked from view by objects that were not the experimental stimuli (e.g., experimenter's hair or experimenter's hand covered her face). For the non-preferred condition, the same variables were coded, but since the participant did not have the opportunity to hold the non-preferred object, *Face out of view* was only coded when the experimenter's face was completely out of frame, and *Partial face* was coded when the head camera angle was positioned so that only part of the experimenter's face was visible in the frame. All categories were coded as length of looking time measured in second.

Video annotation software Chronoviz V. 2.0.2 was used for behavioral coding (Fouse et al., 2011). Video timelines for each participant were annotated with each of the above 4 visual attention categories, and the software was used to time stamp the start and end points of each event. The time stamps for all annotations were exported to Microsoft Excel, the duration (in seconds) of each event was extracted, and a total duration value for each variable (full face, face out of view, face partial, other) was computed for each group of participants.

Two researchers (JY and VN) coded the video footage. VN, who was blind to group membership and the aims of the experiment coded a random 60% of all data, while JY coded all data. To measure the extent to which the two coders agreed with each other, inter-rater reliability (IRR) was computed through the use of single measure Intraclass correlation coefficient (ICC; see Table 2-2) with two-way mixed model and absolute agreement. Inter-rater reliability was generally high.

Table 2-2: Intraclass correlation coefficients with 95% CI

	Preferred			**Non-preferred**
	0s-20s	*20s-40s*	*40s-60s*	*0s-20s*
Full face	.739	.999	.999	.729
Face out of view	.851	.991	.995	.828
Partial face	.975	.987	.962	.942
Experimenter speaking	.947	.965	.895	.865
Participant speaking	.964	.985	.956	.986

Note: All ICC scores obtained $p < .001$

2.4.2. Coding speech states

There were two categories in speech coding: *experimenter speaking and participant speaking*. Unlike with the visual attention data, the two categories were not mutually exclusive (i.e., they could overlap in time if the two people were speaking simultaneously). I did not additionally code for silence, or code for the content or the interpretability of the utterances. All categories were coded as duration of speaking time measured in second.

2.4.3. Coding visual attention in relation to speech state

Freeth et al., (2019) suggested that social attention is influenced by conversational phase (i.e., if a participant was speaking or listening) during a face-to-face interaction. In this study, I was interested in whether social attention varied as a function of speech state (experimenter speaking vs. participant speaking). It should be noted that in Freeth and colleague's experiment, conversation was strictly structured in a question and answer format. In their study, participants' listening state could be assumed as they needed to answer each question to proceed, whereas the conversation in the present study was closer to a real life conversation, making it difficult to know whether participants were truly listening to the experimenter's speech.

To examine visual attention behaviours during each speech state, I extracted every visual attention coding event, but only when either the participant or experimenter was speaking, in Microsoft Excel. Data for this specific analysis was calculated in percentage because total speaking time varied from one participant to another. To compare visual attention data in relation to verbal exchange across participants, the duration of each event was first summed up to compute the total duration for each visual attention category (collapsed over the 60s video) for each participant. Then the total duration of each social attention categories was then divided separately by the total duration of each speech states (experimenter vs. participants) to create proportion values. Proportions were used to account for differences in length of speech utterances between participants, which would have conflated the social attention results if raw durations were used instead of proportions. The resulting data were then exported to SPSS for further analysis. I did not additionally code for incidents in which both participants and experimenter stayed silent or spoke simultaneously as they were not of current research interest. The same procedure was conducted for the non-preferred object condition, but as mentioned before, only 20s of footage was analyzed

2.5. Testing for violations of normality, handling outliers

2.5.1. Visual attention data

Preferred condition

Analysis was constrained to the variables *Face Full, Face Out of View, Face Partial,* excluding the “other situation” category as it was short in duration (average duration is less than 0.1s for both group in the 60s long video) and was not of interest to the current research question. Therefore, for the preferred condition, for each diagnostic group (ASD, TD) it had three 20s time segments (i.e., 0-20s, 20-40s, 40-60s) x 3 visual attention variables (i.e., full face, partial face, and face out of view), generating a total of 18 dependent variables.

The data for each variable were analyzed to identify extreme values from the dataset. Eight (8) extreme values were detected by SPSS, and all of these values came from ASD group. Removing data points may directly bias the results if they are valid and relevant data points (Tukey, 1962, p. 18). Therefore, I chose to winsorize the extreme values by replacing each with the next highest or lowest value that was not suspected to be an extreme value.

A Kolmogorov-Smirnov test was run to test the assumption of normality required for parametric statistical testing. Results indicated that all 9 dependent variables for the ASD group violated the assumption of normality, whereas in the TD condition, 4 out of 9 violated the assumption. Data transformations were performed as an attempt to correct the normality violations. I first performed a logarithmic transformation in which only positive non-zero values were allowed. Because the distributions contained several zero values, to run the transformation, all data need to be added by 1 in order to eliminate the zero values. After running the logarithmic transformation, 6 out of the 18 transformed dependent variable distributions still violated the normality assumption. Similar failures to normalize the data were obtained with square and square root transformations (16/18 and 4/18 respectively violated normality assumption after transformations were performed).

One possibility of these failed attempts can be attributed to the uncommonly large amount of zero values in the dataset (e.g., 32% of participants failed to look at the experimenter's face at all resulting in zero values for Full Face; 4.7% participants looked at her often leading to zero values for Face Out of View), which inevitably skewed the distribution to one way or the other. Because of the skewed nature of the data and lack of success in transforming the data, I chose to use nonparametric methods to test the study hypotheses. This was a mixed design study in which diagnostic group (ASD vs. TD) was the between-subject factor, and time (0-20, 20-40, 40-60) was the within-subject variable. To the best of my knowledge, there is no adequate non-parametric equivalent test for a mixed design ANOVA. Hence, I decided to use Mann-Whitney U test for the between-subject variable (group membership), while a Friedman's test was run separately for the effect of within-subject variable (time).

Non-preferred condition

Results from a Kolmogorov-Simonov test showed 2/3 of the dependent variables in the ASD group and 2/3 in the TD group violated the assumption of normality in parametric tests. Data for the partial face variable in the ASD group had the most heavily skewed distribution with skewness of 1.88 (*SE* = .512) and Kurtosis of 4.14 (*SE* = .992). Three extreme values were detected and winsorization was performed. Data transformation was also attempted to normalize the data distributions. Logarithmic, square and square root transformations were applied separately, but all failed to improve the distributions (6/10 variables still violated normality after log transformation, and 3/10 and 6/10 variables still violated normality after square and square root transformation respectively). Similar to preferred condition, these data did not meet the assumptions for parametric testing, therefore Mann-Whitney's U test was run for the between-group effect (diagnostic groups). Given that the non-preferred condition had only one 20s segment, no within-subject variable was in this analysis.

2.5.2. Speech data

Preferred condition

A Kolmogorov-Smirnov test revealed that one of 1/6 variables in the ASD group and 4/6 in TD group did not satisfy the normality assumption of parametric testing. The most heavily skewed distribution was the 20-40s segment where participants were speaking in TD group, with skewness of 1.62 (*SE* = .427) and Kurtosis of 2.64 (*SE* = .833). Data transformations that attempted to normalize data were again not successful (5/12 DVs still violated normality after log transformation, and 2/12 and 4/12 for square and square root transformation respectively). For this analysis, diagnostic group (ASD vs. TD) remained as the between-group factor and time (0-20, 20-40, 40-60) as the within-group variable. Again, because variable distributions did not satisfy the normality assumption, Mann-Whitney U test was used for the between-subject variable (group membership), while a Friedman's test was run separately for the effect of the within-subject variable (time).

Non-Preferred condition

Results of the Kolmogorov-Simonov test revealed that two variables in the ASD group and one variable in the TD group violated the assumption of normality. Data on participant speaking in TD group has the most heavily skewed distribution with skewness of 3.04 (*SE* = .427) and Kurtosis of 9.319 (*SE* = .833). Three extreme values were detected and winsorization was performed. Data transformation was also attempted to normalize the data distributions. Logarithmic, square and square root transformations were applied separately, but these transformations were not successful (2/4 of the variables were still violated normality after log transformation. square and square root transformation). No within-subjects variable was used in this analysis. Mann-Whitney's U test was run for the between-group effect (diagnostic group).

2.5.3. Visual attention in relation to speech state data

This analysis focuses on visual attention data restricted to only when either the experimenter or participants were speaking. For this specific analysis, the between-group variable was group. The within-group variable was speaking state (experimenter speaking vs. participant speaking).

Preferred object condition

Similar to the data of only visual attention, descriptive statistics indicated that the nature of the data did not meet the normality assumption for parametric test. Specifically, Kolmogorov-Smirnov test revealed that 5/6 dependent variable in ASD group and 1/6 in TD group did not satisfy the normality assumption of parametric test. Data on *full face in* *relation to participant speaking* variable in the ASD group had the most heavily skewed distribution with skewness of 1.70 (SE = .512) and Kurtosis of 1.85 (SE = .992). Thirteen (13) extreme values in the ASD group and five (5) in the TD group were identified. Winsorization was performed. Logarithmic, square and square root transformations were not adequate for proportional data. Logit transformation was also considered, but this transformation cannot be performed with zero values in the dataset, therefore data transformation was not successful. Again, Whitney U test for the between-subject variable (group membership) was used, and Friedman's test was run for the speech states (experimenter vs. participants) as a within-group variable. It should be noted that since time was not the primary interest in this analysis, the three 20s segments were collapsed together in this specific analysis.

Non-Preferred condition

This section involved the experimenter explaining the reason of her choice of object, and the conversation was relatively one-sided as the experimenter was the only one talking for most of the condition. Because the social attention data were coded as percentage (proportional to the speaking state durations), no actual data can be input into the calculation of proportions if the participant did not speak at all (cannot divided by zero). Indeed, 30% of the data for ASD and 56.7% for TD was missing in the participant speaking category. Therefore, it was decided that the data were too scarce to further run any useful analytic test

Chapter 3. Results

3.1. Visual attention

3.1.1. Preferred object condition

The total duration of time during which participants visually attended to discrete foci (i.e., full face, face out of view, and partial face) was analyzed for between-group differences (ASD vs. TD) at different time intervals (0-20, 20-40, 40-60s). Mann-Whitney U tests (with alpha of .016 after Bonferroni correction) indicated that the ASD group spent significantly less time than the TD group attending to the experimenter's full face at each time interval (0-20s, $U = 169.0$, $z = -2.647$, $p = .008$; 20-40s, $U = 180.0$, $z = -2.415$, $p = .016$; and for 40-60s, $U = 134.5$, $z = -3.301$, $p = .001$). No significant between-group differences in time attending was detected when the experimenter's face was out of view in the 0-20s video segment, $U = 189.0$, $z = -2.20$, $p = 0.028$; however, relative to the TD group, the ASD group spent significantly more time looking away from the experimenter's face during the 20-40s and 40-60s segments, $U = 127.0$, $z = -3.431$, $p = .001$ and $U = 164.0$, $z = -2.697$, $p = .007$ respectively. No significant between-group difference was detected when attending to the experimenter's partial face during the first 20s segment, $U = 246.5$, $z = -1.061$, $p = .289$; however, the ASD group spent significantly less time than the TD group attending to the experimenter's partial-face than the TD group during the 20-40s and 40-60s segments, $U = 163.5$, $z = -2.707$, $p = .007$ and $U = 153.5$, $z = -2.907$, $p = .004$ respectively. Figure 3-1 illustrates these results.

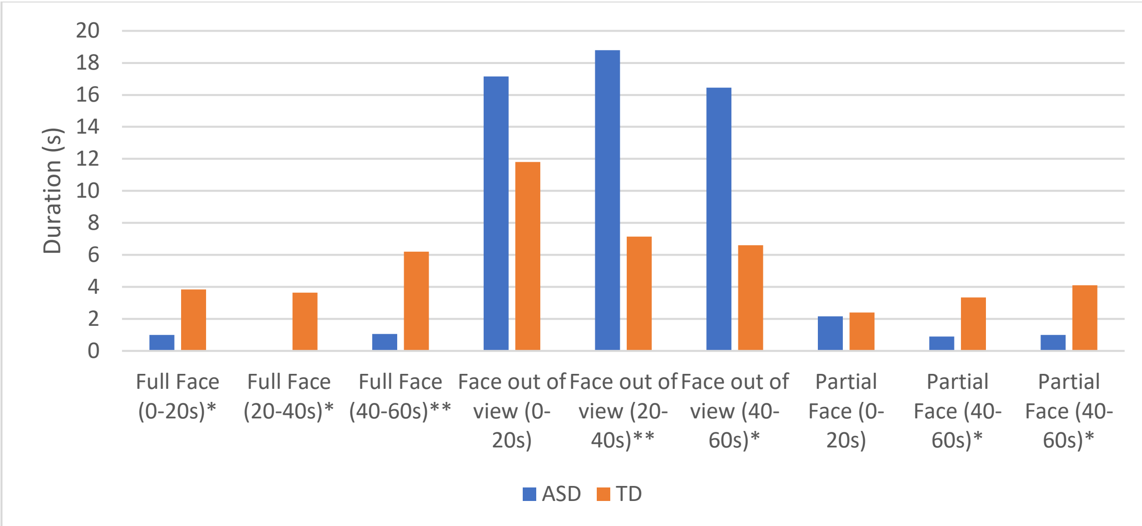

Figure 3-1: Social attention in preferred condition (between-group)
Note: Bars represent median durations (s) in each social attention coding category (Full Face, Face out of view, Partial face) in the preferred condition. Blue bars represent ASD, orange bars represent TD. Asterisks indicate the significance of the between-group effect; * $p < .05$; ** $p < .01$.

The effect of time interval on visual attention to different foci as a within-group variable was analyzed using Friedman's test separately for each of the participant groups. In the ASD group, time interval had no significant effect on the total time spent attending to the experimenter's full face, χ^2 (2) = .406, p = .816; face out of view, χ^2 (2) = 4.914, p = .086; and partial face, χ^2 (2) = 2.225, p = .329). However, within the TD group, an effect of time interval was detected in the full-face category, χ^2 (2) = 8.420, p = 0.015, and in the out-of-view category, χ^2 (2) = 11.916, p = 0.003. No significant within-TD group effect was detected for the partial face, χ^2 (2) = 3.345, p = .188. A follow-up Dunn-Bonferroni post hoc test detected the TD group spent more time in the full-face category at 40-60s than at 0-20s (p = .008), and conversely, less time in the face out-of-view category at 40-60s than at 0-20s (p = .001). Thus, on average, the TD participants increased their attention to the experimenter's face over time, whereas the attention of the participants with ASD was directed equally over time to the experimenter's full face, partial face and away from the experimenter's face. Figure 3-2 and 3-3 illustrate these results in bar graph.

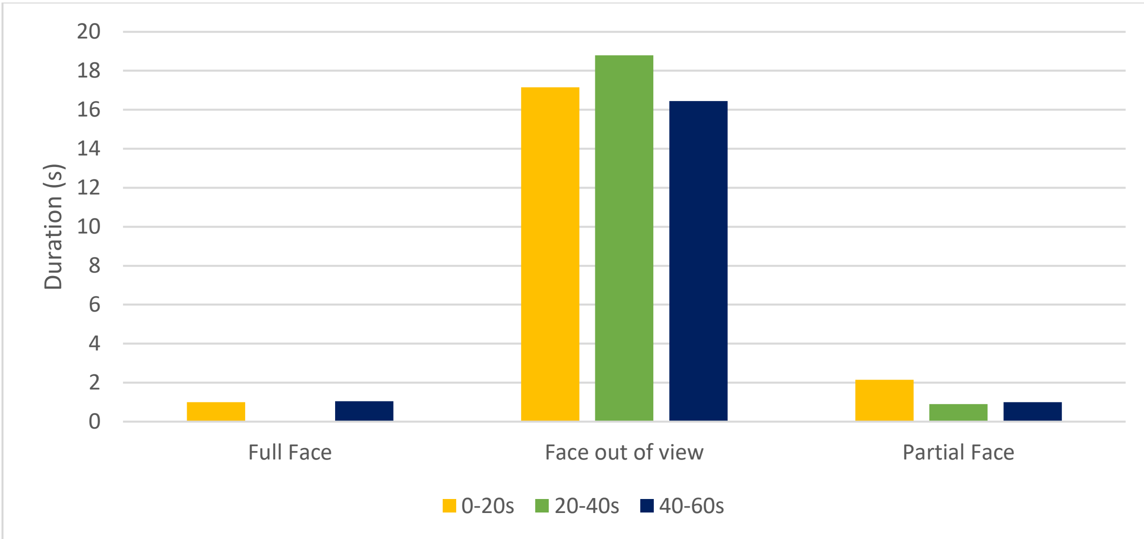

Figure 3-2: Social attention in preferred condition, within-group (ASD)
Note: Bars represent median durations (s) in each social attention coding category (Full Face, Face out of view, Partial face) for the ASD group in the preferred condition. Data are plotted as a function of the within-group variable of time interval (blue bars represent time interval 0-20s, orange bars represent 20-40s, grey bars represent 40-60s). No signfiicant effect of time interval was observed;

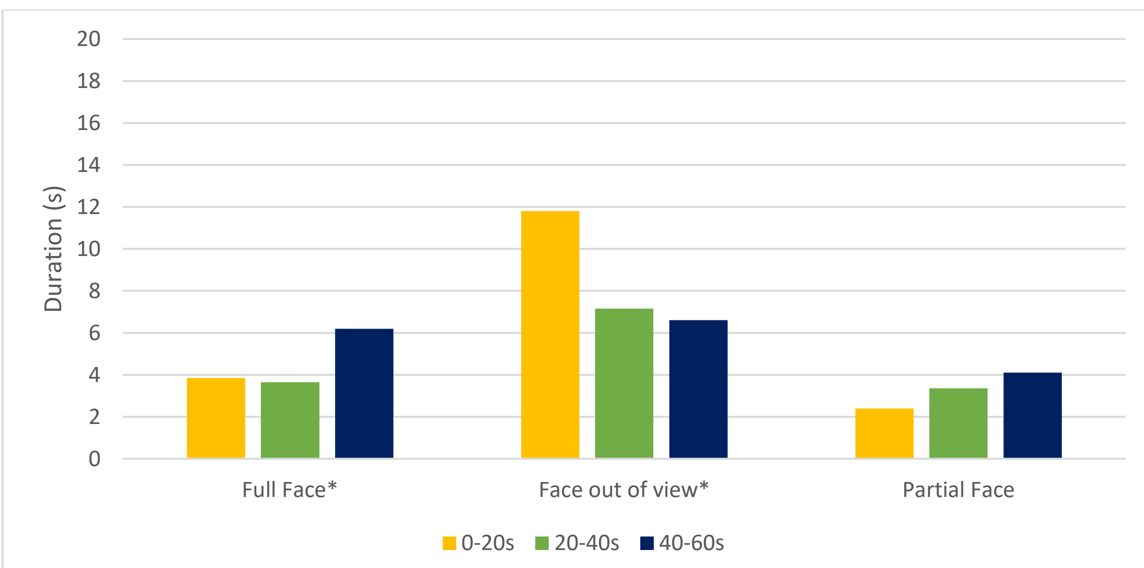

Figure 3-3: Social attention in preferred condition, within-group (TD)
Note: Bars represent median durations (s) in each social attention coding category (Full Face, Face out of view, Partial face) in the non-preferred condition. Blue bars represent ASD, orange bars represent TD. No between-group effects were observed.

3.1.2. Non-preferred object condition

ASD-TD group variation in the total time spent attending to each of the three foci of visual attention was analyzed. Mann-Whitney U tests (with alpha of .016 after Bonferroni correction) did not reveal any between-group differences for each of the three coding categories (full face, U = 288.5, z = -.232, p = .817; face out of view, U = 231.0, z = -1.370, p = .171; and for partial face, U = 232.0, z = -1.351, p = .177). As such, ASD children did not differ from TD children in the time spent visually attending to the experimenter's face in the non-preferred objection condition (see figure 3-4).

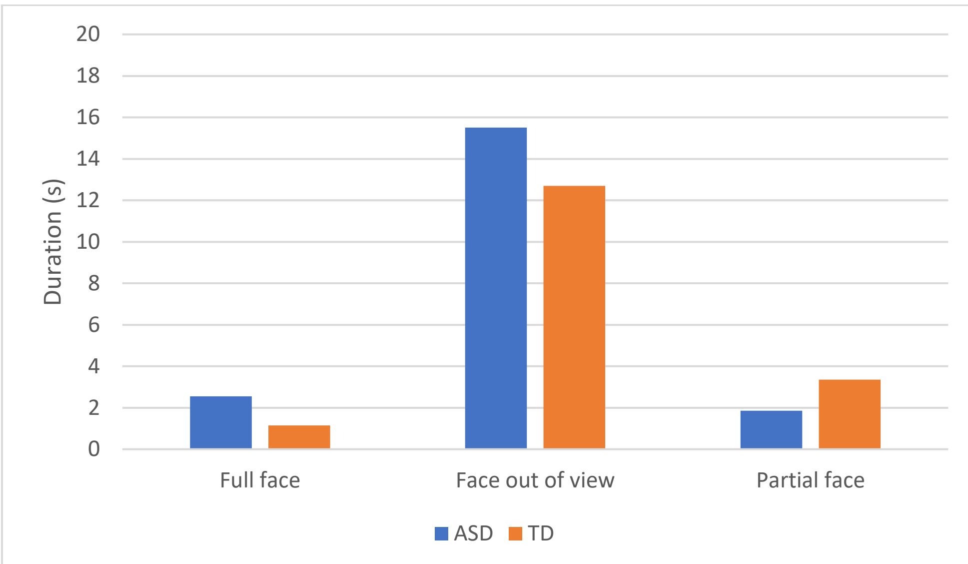

Figure 3-4: Social attention in non-preferred condition (between-group)
Note: Bars represent median durations (s) in each social attention coding category (Full Face, Face out of view, Partial face) in the non-preferred condition. Blue bars represent ASD, orange bars represent TD. No between-group effects were observed.

3.2. Speech states

3.2.1. Preferred object condition

The total time during which the experimenter or participant was speaking was analyzed for group differences (ASD vs. TD) at different time intervals (0-20, 20-40, 40-60s). For **experimenter speaking**, Mann-Whitney U tests (with alpha of .016 after Bonferroni correction) indicated that initially (0-20s) there was no significant group difference between ASD group and TD group (0-20s, U = 289.0, z = -.218, p = .828). However, it was shown that the experimenter spent less time speaking to the ASD group than to the TD group for both the 20-40s and 40-60s video segments (20-40s, U = 158.0, z = -2.813, p = .005; and for 40-60s, U = 150.0, z = -2.971, p = .003). In the **participant speaking** category, no significant between-group difference was detected for the first 40s of the videos (0-20s, U = 288.0, z = -.238, p = .812; and for 20-40s, U = 228.5, z = -1.416, p = .157). However, in the 40-60s video segments, the ASD group spent significantly more time speaking than the TD group, U = 171.5, z = -2.545, p = .011. Taken together, participants with ASD spent more time speaking about their preferred toy than their TD counterparts, but only from 40-60s. Conversely, the experimenter spent less time speaking to the ASD group than to the TD group, not initially (0-20s), but during the 20-40s and 40-60s segments. These results are depicted in Figure 3-5. I next examined within-group effects of time to further understand these patterns.

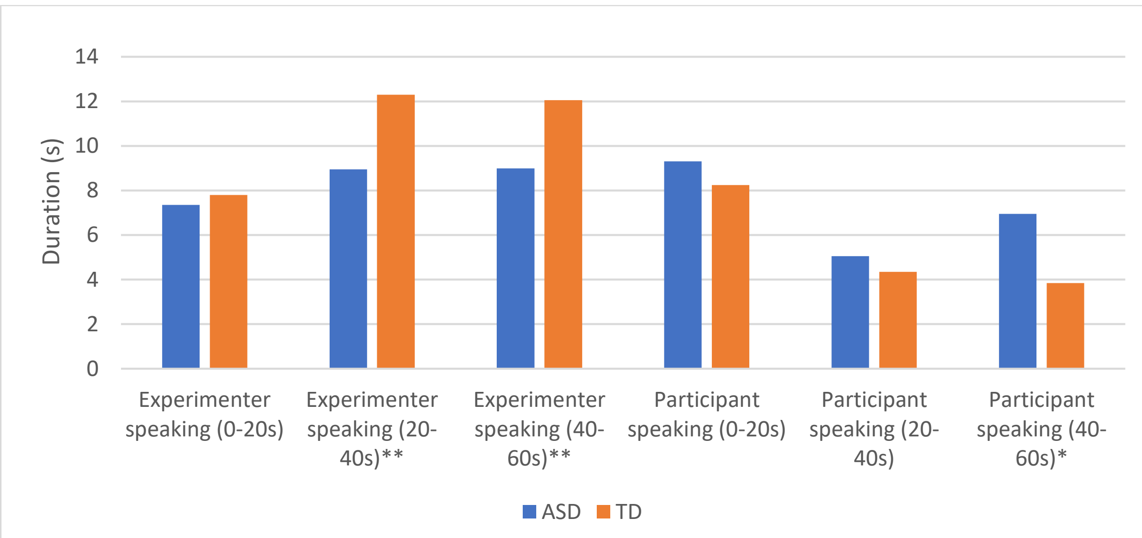

Figure 3-5: Length of utterance in preferred condition (between-group)
Note: Bars represent median durations (s) in each speaking category (Experimenter speaking, Participant Speaking) in the preferred condition. Blue bars represent ASD, orange bars represent TD. Asterisks indicate the signficance of the between-group effect, * p < .05; ** p < .01.

The effect of time interval on the amount of time during which the experimenter or participant was speaking as within-group variables was analyzed using Friedman's test separately for each of the participant groups. In the ASD group, there was a significant effect of time interval on **experimenter speaking** time, χ^2 (2) = 9.333, p = .009. Dune's pairwise post-hoc tests were run to compare each of the time intervals and revealed that the experimenter spoke more during the 20-40s segment than during 0-20s (p = .011), and that the experimenter spent more time speaking during the 40-60s segment than during 0-20s segment (p = .007). We did not detect differences between 20-40s and 40-60s segments (p = .874). However, for the **participant speaking** category, there were no significant differences in speaking time due to time interval detected, χ^2 (2) = 2.70, *p* = .259. In other words, within the ASD group, the time during which the experimenter was speaking increased during the interval of 0 to 40 seconds, then the time spent speaking became stable to 60 seconds. The amount of speaking time among the ASD participants was stable over time (see Figure 3-6).

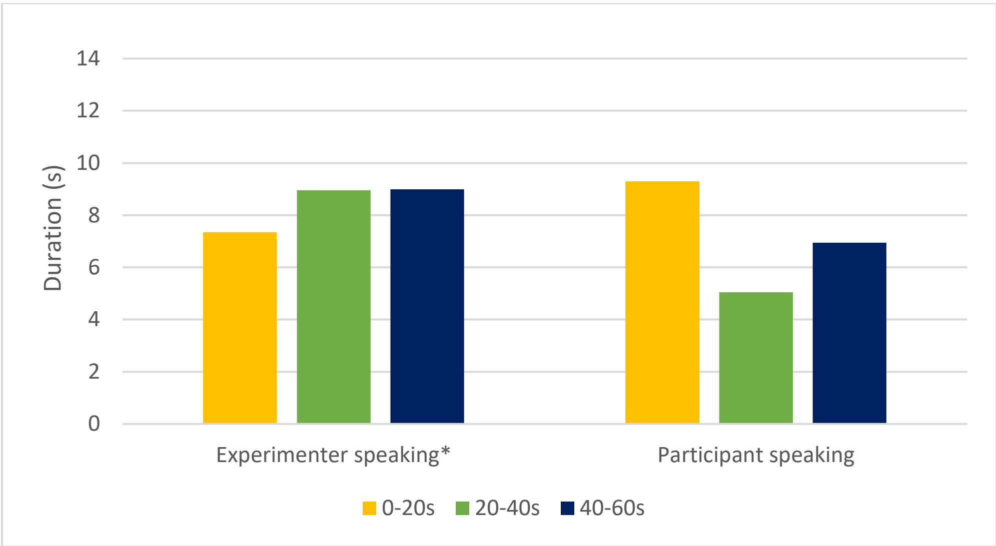

Figure 3-6: Length of utterance in preferred condition, within-group effect of time (ASD)

Note: Bars represent median durations (s) in each speaking category (Experimenter speaking, Participant Speaking) in the preferred condition, for the ASD group. Data are plotted as a function of the within-group variable of time interval (blue bars 0-20s, orange bars represent 20-40s, grey bars indicate 40-60s). Asterisks indicate the significance of the within-group effect, * p < .05; ** p < .01.

As for the TD group, an effect of time interval was detected in the **experimenter speaking** category, χ^2 (2) = 22.20, *p* < .001, Dune's pairwise post-hoc tests revealed that the experimenter spoke more during the 20-40s segment than during the 0-20s (p < .001), and that the experimenter spent more time speaking during the 40-60s segment than during the 0-20s segment (p < .001). The results did not detect differences between 20-40s and 40-60s segments (p = .699). In the **participant speaking** category, there was also a significant effect of time, χ^2 (2) = 15.610, *p* < .001. Dune's pairwise post-hoc tests found that TD participant spoke more during 0-20s than during the 20-40s segments (p < .001), and participants spent more time speaking during 0-20s segments than during 40-60s segment (p < .001). The test did not detect differences between 20-

40s and 40-60s segments (p = .846). In summary, the duration of time the experimenter spent speaking to the TD group increased during the interval 0 to 40 seconds and became stable to 60 seconds, whereas the participants in TD group spent less time speaking about their selected toy from 0 to 40 seconds and remained stable until the end (40-60s). Figure 3-8 illustrates the results.

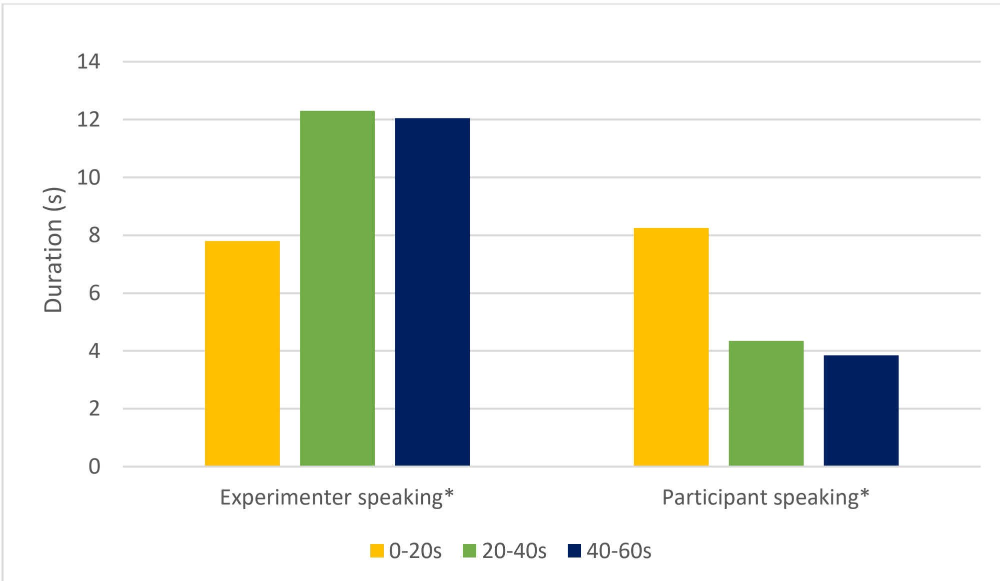

Figure 3-7: Length of utterances in preferred condition, within-group (TD)
Note: Bars represent median durations (s) in each speaking category (Experimenter speaking, Participant Speaking) in the preferred condition, for the TD group. Data are plotted as a function of the within-group variable of time interval (blue bars 0-20s, orange bars represent 20-40s, grey bars indicate 40-60s). Asterisks indicate the significance of the within-group effect, * p < .05; ** p < .01.

3.2.2. Non-preferred object condition

The total duration spent in experimenter speaking and participant speaking categories during the 0-20 sec interval in the non-preferred object condition were analyzed for ASD-TD group differences. Mann-Whitney U tests (with alpha of .016 after Bonferroni correction) did not reveal any between-group differences in the time during which the experimenter was speaking U = 298.5, z = -.030, p = .976. However, the results detected that the ASD group spent more time speaking than the TD group, U = 146.0, z = -3.192, p = .001. In fact, the median duration of participant speaking for the TD group was zero; the speech time for both participant groups was very low in this condition. Therefore, in the non-preferred object condition, speech time for the experimenter did not differ between the two groups, but the ASD group on average, spoke longer than the TD group (who, at the group level, did not speak at all) during this 20s interval (see Figure 3-8).

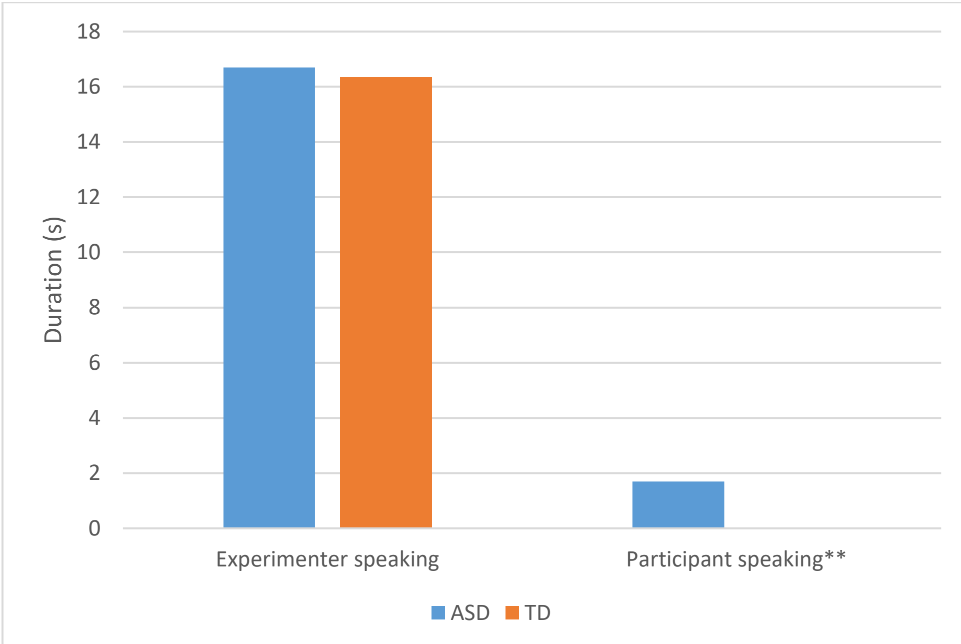

Figure 3-8: Length of utterances in non-preferred condition (between-group)
Note: Bars represent median durations (s) in each speaking category (Experimenter speaking, Participant Speaking) in the non-preferred condition. Blue bars represent ASD, orange bars represent TD. Asterisks indicate the significance of the between-group effect, * p < .05; ** p < .01.

3.3. Visual attention in relation to speech states

3.3.1. Preferred-object condition

The proportion of time that participants attended to each of the three foci of visual attention (i.e., full face, partial face, and face out of view) during each of the two different speaking states (experimenter speaking vs. participant speaking) was analyzed for group differences (ASD vs. TD). Mann-Whitney U tests (with alpha of .016 after Bonferroni correction) indicated that on average, the ASD group spent a significantly lower proportion of time than the TD group attending to the experimenter's full face during both speaking states (experimenter speaking, $U = 164.0$, $z = -2.70$, $p = .007$, and participant speaking, $U = 171.0$, $z = -2.57$, $p = .010$). Conversely, the ASD group spent a higher proportion of time than the TD group looking away from the experimenter's face during both speaking states (experimenter speaking, $U = 144.0$, $z = -3.09$, $p = .002$, and participant speaking, $U = 149.0$, $z = -2.99$, $p = .003$). Finally, the ASD group also spent a significantly lower proportion of time than the TD group attending to the experimenter's partial face in experimenter speaking state, $U = 168.0$, $z = -2.62$, $p = .009$; however, no significant between-group differences in visual attention to the experimenter's partial-face were found in the participant speaking state, $U = 216.0$, $z = -1.66$, $p = .096$ (see Figure 3-9). These results are consistent what was found in the analysis of between-group effects on social attention over the entire session, not anchored to speaking states (see Figure 3-1), and suggest that the social attention patterns found in that analysis cannot be accounted for by group differences in the amount of time spent in the two speaking states.

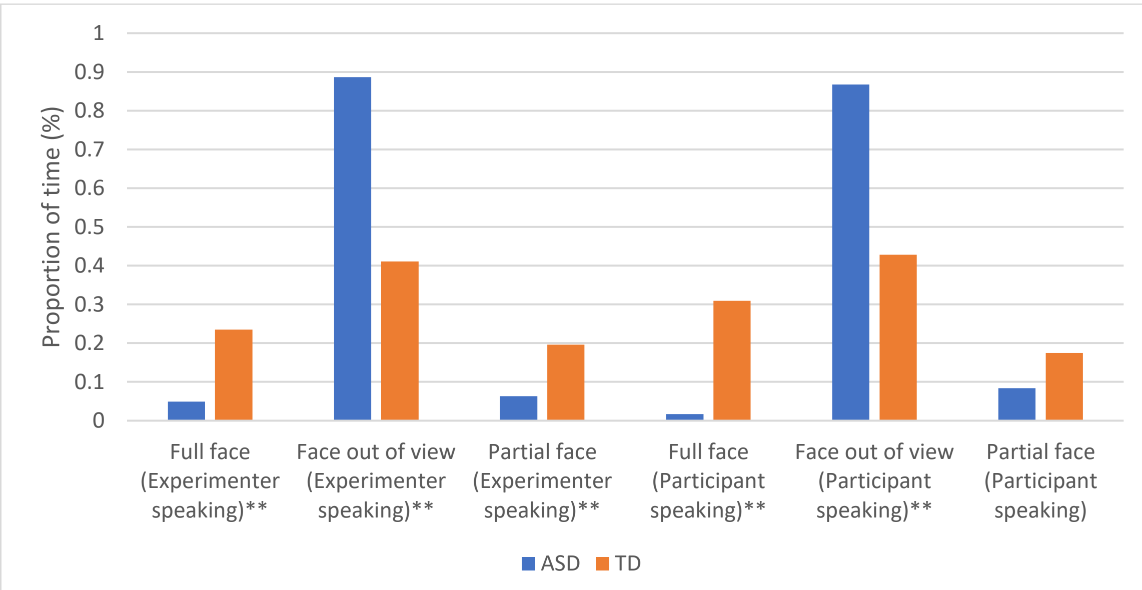

Figure 3-9: Visual attention in relation to speaking states (between-group)
Note: Bars represent median proportion of time spent in each social attention coding category (Full Face, Face out of view, Partial face) when either the Experimenter or Participant was speaking in the preferred condition. Blue bars represent ASD, orange bars represent TD. Asterisks indicate the significance of the between-group effect, * $p < .05$; ** $p < .01$.

The influence of different speaking states (experimenter speaking vs. participant speaking), as a within-group variable, on visual attention was analyzed using Friedman's test separately for each of the participant groups. In the ASD group, speaking state had no significant effect on proportion of time spent on any of the three foci of visual attention (full face, $\chi^2(2) = 96.0$, $p = .356$; out of view, $\chi^2(2) = 73.0$, $p = .586$; and partial face, $\chi^2(2) = 110.0$, $p = .286$). Similar results were found in the TD group, where no effect of speaking state was detected for proportion of time spent on any of the three visual attention foci (full face, $\chi^2(2) = 214.0$, $p = .328$; out of view, $\chi^2(2) = 194.0$, $p = .611$;

and partial face, χ^2 (2) = 231.0, p = .975). As such, speaking state was not associated with the proportion of time spent attending to different foci, for either diagnostic group. See Figures 3-10 and 3-11.

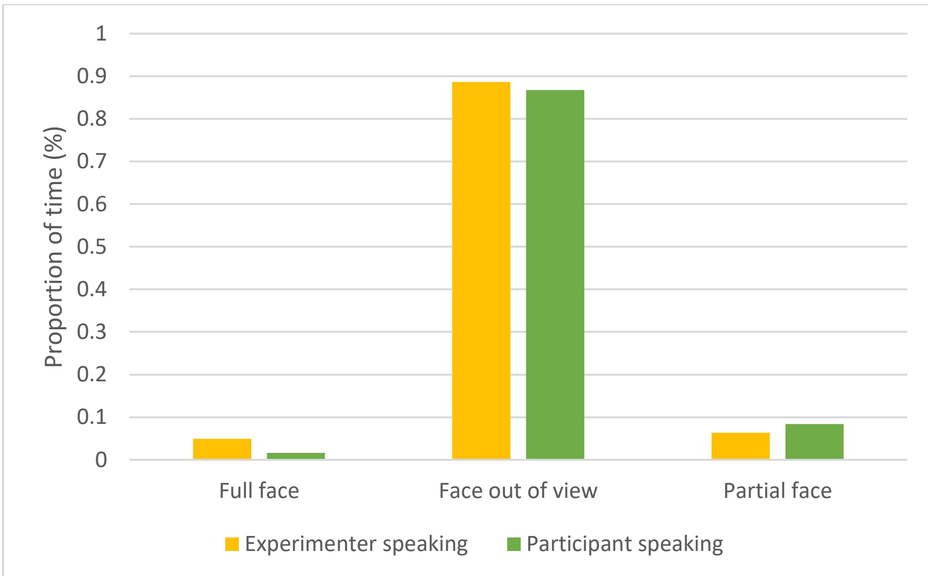

Figure 3-10: Visual attention in relation to speaking states, within-group (ASD)
Note: Bars represent median proportion of time spent in each social attention coding category (Full Face, Face out of view, Partial face) as a function of whether the Experimenter or Participant was speaking, in the preferred condition, for the ASD group. Blue bars represent Experimenter speaking, orange bars represent Participant speaking.

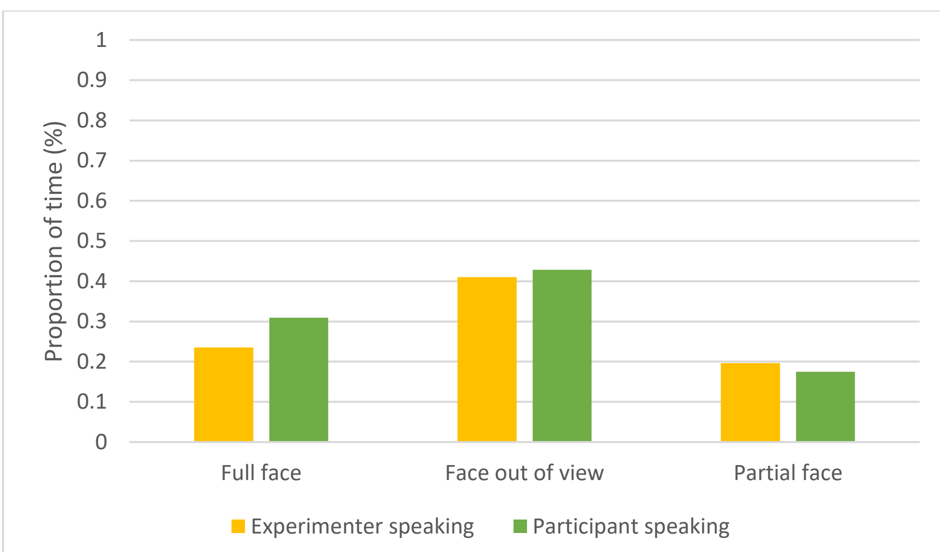

Figure 3-11: Visual attention in relation to speaking states, within-group (TD)
Note: Bars represent median proportion of time spent in each social attention coding category (Full Face, Face out of view, Partial face) as a function of whether the Experimenter or Participant was speaking, in the preferred condition, for the TD group. Blue bars represent Experimenter speking, orange bars represent Participant speaking.

3.3.2. Non-preferred object condition

The proportion of time spent attending to each of the three foci of visual attention (i.e., full face, partial face, and face out of view) spent during **experimenter speaking** category was analyzed for group differences (ASD vs. TD) in non-preferred object condition. Mann-Whitney U test revealed that there was no between-group difference for any of the three categories of visual attention during the experimenter speaking state (full face, $U = 245.0$, $z = -1.107$, $p = .268$; face out of view, $U = 205.0$, $z = -1.887$, $p = .059$; and partial face, $U = 222.0$, $z = -1.549$, $p = .121$). Thus, similar to the analysis of social attention during the whole session, not anchored to speaking states (see Figure 3-4), participants in the ASD group did not differ from the TD group in their pattern of visual attention behaviors when the experimenter was speaking about their personal preference for a toy.

In the **participant speaking** state, the participants rarely spoke regardless of their group membership. In particular, 30% participants in the ASD group and 53.3% in the TD group did not speak at all. No statistical analysis could be run due the scarcity of data in this condition.

Chapter 4. Discussion

The current thesis investigated how social attention and verbal exchange behaviors differ in children with ASD, relative to TD controls, when they were engaging a face-to-face social interaction around a high interest object. Research on this topic in naturalistic settings is scarce. This study was conducted to be exploratory in nature to gain insight and guidance for future studies on the same topic.

4.1. Social attention

The primary goal of the current research was to investigate social attention of children with ASD compared to typically developing peers during a face-to-face social interaction (with an experimenter) around high- and low-interest objects. In particular, this study examined three hypotheses derived from the literature. 1) The *CI-distractor hypothesis* (Sasson et al., 2008; Sasson and Touchstone, 2014) predicted that relative to TD children, children with ASD would show reduced social attention when discussing a high-interest object, but there would be no ASD-TD group differences in social attention when discussing a low interest object. 2) The *social motivation hypothesis* (Chevallier et al., 2012), predicted that children with ASD would show reduced social attention relative to TD children regardless of the interest-level of objects they were interacting with. 3) Based on research by Nadig et al. (2010), the third hypothesis was that no group differences would be found in either high or low interest object condition, but social attention would be increased for both groups in the high-interest condition compared to the low-interest condition.

The current study revealed that, when interacting around a high interest object, participants with ASD spent significantly less time looking at experimenter's full face during the entire 60s interaction window, and significantly less time in the "partial face" category (where only part of the experimenter's face was visible) relative to TD children during the 20-40s and 40-60s intervals. Not surprisingly, children with ASD spent significantly more time in the "face out of view" category than their TD counterparts, suggesting that children with ASD spent more time than TD children looking away from the experimenter. In contrast, no ASD-TD group difference in social attention was detected when participants were interacting around a low interest-object. These findings are in line with recent studies that showed reduced social attention (as evidenced by eye tracking) in children with ASD relative to TD children in the presence of a high-interest object, but no group differences in social attention in the presence of low-interest stimuli (i.e., Ewing et al., 2013; Parish-Morris et al., 2013; Sasson and Touchstone, 2014). Thus, of the three hypotheses presented, the present study largely supports the CI-distractor hypothesis.

Importantly, the findings of the present study may provide some clarity on two seemingly contradictory findings from past literature: while a conversation topic about a CI promotes social attention in children with ASD (Nadig et al., 2010), an interaction around a physical high interest object on the contrary seems to reduce their social attention, at least as found in studies in which participants passively viewed static images (e.g., Sasson & Touchstone, 2014). The present study provides support for the CI distractor hypothesis even during a face-to-face conversation, and suggests that the physical presence of a high interest object leads to different effects on social attention in children in ASD, relative to conversation about a CI topic without the physical object present. This may have an implication for research on intervention utilizing CI as a tool to improve social communication. A growing body of intervention studies have found CI to produce improvements social skills and behavior of children with ASD (for review see, Harrop et al., 2017). For example, Boyd (2007) conducted an intervention study and found that three pre-schooler participants with ASD engaged in more social interactions (measured by behavioral coding) when they were presented with CI-related object, compared to a less preferred item. In contrast, other studies were less conclusive about CI as a beneficial teaching tool in intervention. For example, one study embedded personalized CI (e.g., toy car) within intervention procedures, with an attempt to improve participants' initiating joint attention ([IJA]; Kryzak & Jones, 2015). The result showed that, among the three participants aged from 3 to 8 years old, two were responsive to the intervention, however the third participant did not master IJA until the CI was removed from the activities. The authors argued that the CI-related object was too intense for the child and became an interference to the teaching procedure. Among these studies in Harrop et al.'s (2017) review, the format of CI presentation ranged from physical objects (e.g., Kryzak & Jones, 2015), conversation topics (e.g., Lepper et al., 2017), and activities (e.g., Koegel et al., 2012). If different formats of the CI presentation may elicit different social behaviors as implicated by current study and studies mentioned above, future research may be beneficial in directly comparing how different formats of CI presentation (e.g., conversation topic, physical object, or activities) would have an impact on interventions targeting social interaction in ASD.

This study further examined how social attention behaviors unfold over time during the social interaction. Interestingly, I found significant within-group effects of time on social attention in the TD group, but not in the ASD group. Specifically, participants in TD group gradually increased their time spent on looking at the experimenter’s face and reduced their time looking elsewhere, whereas ASD group did not change their social attention behaviors during the course of the interaction. One possible explanation is that the participants in TD group were becoming more comfortable talking with the experimenter over time and therefore paid more attention to the experimenter’s face, whereas the children with ASD did not become more comfortable in the unfamiliar social interaction. This finding aligns with previous research that studied interpersonal coordination, a concept that depicts the process of social partners aligning the form and timing of their behaviours (Bernieri & Rosenthal, 1991), for example, through dyads naturally coordinating their body movements, speech, gaze, and facial expressions. Interpersonal coordination has been linked with feelings of rapport and connectedness between social partners (Chartrand & Bargh, 1999), and is considered a measurable reflection of social emotional reciprocity. Zampella et al. (2020) found that, relative to children with ASD and their conversational partners (participant's mother or an unfamiliar RA), typically developing children developed stronger interpersonal facial affect coordination (e.g., smile coordination) with their conversational partners (participant’s mother or an unfamiliar RA) during the course of a conversation. While there was no significant effect of partner familiarity, the group differences in affect coordination were numerically larger for interactions with an unfamiliar RA, than they were for mother-child interactions. The authors argued that the typically developing participants were more motivated to coordinate with the RA’s smiles to potentially make the interaction more comfortable, whereas children with ASD were less likely to be motivated to match the unfamiliar RA’s smiles over time. In the present study, while I did not examine interpersonal coordination, the increase in social attention over time found for TD children may reflect their motivation to make the interaction more comfortable. In contrast, the children with ASD may have been less motivated to engage the experimenter to improve the comfort level of the interaction and to develop social affiliation with the experimenter. Alternatively, consistent with the CI-distractor hypothesis, perhaps the preferred toy was sufficiently distracting for children with ASD throughout the social interaction to detract from the development of social-emotional reciprocity through increasing social attention over time. Future research should examine these possibilities.

Another interesting factor to be considered for future research is the familiarity of the conversation partner. Participants in this study only interacted with an unfamiliar social partner (experimenter). However, Harrison and Slane (2019) found that both participants with and without ASD preferred to attend to social stimuli that were familiar to them over other non-familiar stimuli (e.g., non-familiar face or objects). It was also suggested that familiar social stimuli or contexts may facilitate effective social communication in children with ASD (Zampella et al., 2020). Therefore, one might predict that group (ASD-TD) differences in social attention would become smaller if children were interacting with a familiar social partner.

4.2. Verbal exchange

The secondary focus of this study was to explore the length of verbal exchange during the social interaction and whether verbal exchange influences or is influenced by, social attention. Based on previous research (Nadig et al., 2010), I hypothesized that the ASD group would speak longer about their high-interest toy compared to the TD group. Overall, the results showed that participants with ASD spoke more compared to their TD counterparts from 40 to 60 seconds of the video analyzed. At first glance, this aligned with the prediction made by Nadig et al. (2010), that children with ASD would be more likely to produce one-sided and over-informative speech when they are talking about their CI-related interests. Participants with ASD may speak longer about a high interest object because they are more knowledgeable in the topic being discussed, therefore they are more likely to spend more time elaborating on the questions asked. Another possibility is that they initiated other information of their CI-related topic that the experimenter did not ask for, thereby lengthening their speech duration (Capps et al., 1998). Caution must be exercised for such an interpretation. Firstly, the current study did not analyze the content of the conversation, and therefore, it is not possible to determine the underlying reason for group differences in speech durations. Future research should incorporate methods in analyzing the content of the conversation. For example, Tager-

Flushberg et al. (1991) developed a method that measures the contingency of utterances in individuals with ASD. This method focuses on how contingent the individuals' response or initiations of speech are during a conversation. An utterance is considered contingent if it gives appropriate amount of information that is relevant to the prior comment made by the conversation partner. Another method that targets the quality of conversation is conversation analysis. This method allows researchers to test their hypothesis by qualitatively analyzing transcribed conversation and identifying emerging themes or recurring patterns in conversation (Dickerson & Robins, 2017). By applying these methods, future research could examine and interpret the relationship between content/quality of the conversation and social attention in a naturalistic setting. Furthermore, it must be noted that there was no significant group difference in duration of participant speaking for the first 40s of the conversation. My analysis of time as a within-group variable suggests that TD children actually decreased their speaking time over the course of the 60s conversation, whereas children with ASD did not show any changes over time. Taken together, the group difference in participant speech duration could be driven by children in the TD group speaking less over the course of the conversation with the experimenter, as opposed to the interpretation that children with ASD produced more speech related to their preferred item.

Interestingly, the time-related patterns found for speech durations concur with my results for social attention. In particular, typically developing children increased their social attention over time but spoke less over the course of the conversation, whereas children with ASD did not change their social attention or speech throughout the interaction. This again may suggest that children with ASD may perceive social environment differently than their TD counterparts and therefore did not act accordingly to the adapt to the social context (Zampella et al., 2020). In addition, the gradual decrease in speaking time in the TD group may also suggest that they are listening more over time, to the experimenter (who spoke more over time, to both groups of children), which in turn, may increase their social attention (Freeth and Bugembe, 2019). To examine this further, I investigated participants' social attention in relation to their speaking states (discussed in more detail in section 4.3)

In the non-preferred object condition, it was found that participants in the ASD group spent more time talking than their TD counterparts. The finding that children with ASD spoke more in both the preferred and non-preferred object conditions was unexpected and suggests that the increased speech time in ASD is not related to the topic (high interest vs. low interest). However, it is important to take into account that the conversation in non-preferred condition was very one-sided because the experimenter inadvertently occupied most of the time in this session to explain her choice of object. Indeed, 44% of all participants had limited to no opportunities to speak and just passively listened to the experimenter's speech in this condition. Future research should use an improved experimental design to include a better manipulation that ensures the nature of the conversation and social interaction is similar between conditions.

4.3. Visual attention in relation to speech state

As mentioned earlier, this study was also interested in participants' social visual attention behavior while they were in different conversational phases (participant speaking vs. experimenter speaking) because previous research showed that when individuals with ASD were speaking, their social attention was reduced, compared to when they are listening to their interlocuter (Freeth and Bugembe, 2019). Given that, in the present study, the children with ASD spoke for longer than the TD children, one might be concerned that differences in speech state may account for the social attention findings of the present study. In the preferred object condition, the current results indicated that in *both* speaking states (participant speaking vs. experimenter speaking), children with ASD looked less at the experimenter's full face and spent more time in the face-out-of-view category than the TD controls. This suggests that the between-group differences in social attention between the ASD and TD group were not driven by differences in time spent in the two speaking states. Specifically, I would argue that the reduced social attention observed in ASD group was not likely a result of their increased speaking time during the course of the interaction compared to the TD group.

This study also examined how social visual attention differed as a function of speech state within each group of the participants. Previous research would predict that children increase their attention to the experimenter's face when the experimenter is speaking, versus when the participant is speaking (Freeth et al., 2019). This pattern was not revealed for either the ASD or TD group. this would suggest that speaking states, at least during a semi-structured conversation, is not related to social visual attention, contradicting Freeth et al.'s (2019) findings. However, caution should be made in this interpretation. It should be noted that the two speaking states in this study (participant vs. experimenter speaking) differed from Freeth et al.'s (speaking vs. listening state of the participants) in an important aspect. In Freeth et al.'s study, listening and speaking states were very clear and controlled. Participants were aware that the experimenter was going to ask four questions sequentially and wait for them to answer (participant speaking state), and then followed by next phase in which participants need to ask the experimenter the same questions and listen to their answer (participant listening state). However, in current study, although less structured, the format of conversation was more similar to real life back-and-forth conversation. It is possible that participants' social attention would behave differently when the conversation format is changed. More research on for both formats can be conducted to systematically compare the two types of conversations

4.4. Limitations

This study had some noteworthy limitations. Although the objects in this study were selected from common CI categories (South et al., 2005), this study did not confirm personal CI for each participant, nor did this study assess how interested participants were in the object they had selected. However, in a recent study that compared the effect of personalized CI and common CI on eliciting group differences in social attention between children with ASD and typical developed children, Harrison et al., (2019) did not find the personalized CI to be any more powerful in attracting social attention than objects belonging to the general CI categories. However, future research using this paradigm should always run manipulation check on participants' subjective evaluation of their preferred toys.

Importantly, even though the ASD-TD group differences in social attention were significant in the preferred condition but not significant in in the non-preferred object condition, the design of the study does not permit strong support for the CI-distractor hypothesis. Indeed, there are several possible interpretations of the data. Crucially, the preferred and non-preferred conditions were not manipulated in a way that allowed a direct comparison between the two conditions. In particular, the order of the preferred and non-preferred conditions was fixed (preferred first, non-preferred second), not counterbalanced. Without counterbalancing the order of the two conditions, it is impossible to disentangle order effects from experimental effects. As mentioned previously, children in TD group may have been more likely to build rapport with the experimenter over the course of conversation in preferred condition; thus it is not clear if this would lead to the participants showing different social attention behavior in the non-preferred condition that would be otherwise not present, because of the order of the conditions. Another limitation is that the two conditions were not matched in terms of the nature of the conversational exchange or physical interaction with the object. For instance, the prompts given by the experimenter that led up to beginning of the conversation differed between conditions, and the preferred condition conversation was centered on the participant's rationale for choosing his/her preferred toy (which he/she held while speaking), whereas the non-preferred condition conversation was focused on the experimenter's reasons for choosing her item (which the experimenter held while speaking). Not only was the nature of the social interaction and the conversational exchange different between the two conditions, but this also resulted in unequal speech time between the experimenter and participant in the non-preferred condition. This may account for why the participants in both groups barely spoke in the non-preferred condition, in that most participants passively listened to the experimenter's explanation for her choice in object. Thus, unlike in the preferred condition, in the non-preferred condition children had little opportunity to act as communicators in the conversation – they were largely recipients. It may be that the passive nature of the interaction eliminated the opportunity to uncover the ASD-related atypicality in social communication needed to produce a group difference in social attention. In sum, examining the experimental effect of a high- vs. low-interest object on the social attention of children with ASD was not possible with the current study design. Future studies should develop a better design for the comparison condition to minimize confounding factors.

The present study intended to analyze participants' eye movements from the eye-tracking apparatus. Unfortunately, data produced from the eye-tracker was not reliable, and the study had to rely on coding point-of-view camera footage recorded from the eye-tracking glasses instead of analyzing participants' actual eye movements. Although previous research has found that head movements are highly correlated with eye-movements to external cues (Khan et al., 2009), more objective evidence is needed that participants were indeed fixating the experimenter's face (vs. other items in camera view). In addition, the coding method adopted in this study divided social attention behavior into three categories: the participants were either assumed to be looking at the face (full face), away from it (face out of view), or somewhere in between (partial face). Without pinpointing the eye-gaze of the participants, it was not possible to compare how much time they spent looking on the actual objects (preferred or non-preferred) to the time spent viewing the social stimuli. An eye-tracker with increased performance would provide a more detailed analysis of visual attention behaviours for participants of each group.

4.5. Conclusion

This exploratory study adds to the literature by gaining some insight in social attention and verbal exchange in children with ASD in a naturalistic environment in the presence of high-interest objects. The main finding of current research was that school-age children with ASD spent less time looking at social stimuli than typically developing controls when they were having a social interaction around a high-interest object. However, such group differences were not detected when participants were interacting around a low-interest object. These findings provide support for the CI-distractor hypothesis proposed by Sasson and colleagues (2008) and challenge the social motivation accounts. The data support the notion that atypical social attention in ASD may be a context-dependent phenomenon rather than a global deficit. However, these conclusions are tentative, given the limitations of the study design and dependant measures, and thus, the underlying mechanisms of how high-interest objects influence social interaction and speech in children with ASD are still far from clear. Future research may further our knowledge of this topic.

References

American Psychiatric Association. (2013). Anxiety Disorders. In *Diagnostic and statistical manual of mental disorders* (5th ed.).

Bernieri, F. J., & Rosenthal, R. (1991). *Interpersonal coordination: Behavior matching and interactional synchrony.* In R. S. Feldman & B. Rimé (Eds.), *Studies in emotion & social interaction. Fundamentals of nonverbal behavior* (p. 401–432). Cambridge University Press.

Bhat, A.N, Galloway, J.C, & Landa, R.J. (2010). Social and non-social visual attention patterns and associative learning in infants at risk for autism. *Journal of Child Psychology and Psychiatry*, *51*(9), 989–997.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real-world scenes: The roles of action, competition and social content. *The Quarterly Journal of Experimental Psychology*, *61*(7), 986–998.

Bottini, Summer. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders*, *45*, 9–26.

Boyd, B.A., Conroy, M.A., Mancil, G.R. *et al.* Effects of Circumscribed Interests on the Social Behaviors of Children with Autism Spectrum Disorders. *J Autism Dev Disord* **37,** 1550–1561 (2007).

Capps, Lisa, Kehres, Jennifer, & Sigman, Marian. (1998). Conversational Abilities Among Children with Autism and Children with Developmental Delays. *Autism : the International Journal of Research and Practice*, *2*(4), 325–344.

Chartrand, Tanya L, & Bargh, John A. (1999). The Chameleon Effect. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

Chawarska, Katarzyna, Macari, Suzanne, & Shic, Frederick. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry*, *53*(8), 903–913.

Chevallier, Coralie, Kohls, Gregor, Troiani, Vanessa, Brodkin, Edward S, & Schultz, Robert T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239.

Chevallier, Coralie, Parish-Morris, Julia, McVey, Alana, Rump, Keiran M, Sasson, Noah J, Herrington, John D, & Schultz, Robert T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, *8*(5), 620–628.

Chita-Tegmark, Meia. (2016). Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Research in Developmental Disabilities*, *48*, 79–93.

Dickerson, P., & Robins, B. (2017). Conversation analysis with children with an autism spectrum disorder and limited verbal ability. In M. O'Reilly, J. N. Lester, & T. Muskett (Eds.), *A practical guide to social interaction research in autism spectrum disorders.* (pp. 167–191). Palgrave Macmillan/Springer Nature.

Doherty-Sneddon, G, & Phelps, F G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & Cognition*, *33*(4), 727–733.

Ewing, Louise, Pellicano, Elizabeth, & Rhodes, Gillian. (2013). Using effort to measure reward value of faces in children with autism. *PloS One*, *8*(11), e79493–e79493

Freeth, Megan, & Bugembe, Patricia. (2019). Social partner gaze direction and conversational phase; factors affecting social attention during face-to-face conversations in autistic adults? *Autism : the International Journal of Research and Practice*, *23*(2), 503–513.

Freeth, Megan, Foulsham, Tom, & Kingstone, Alan. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PloS One*, *8*(1), e53286–e53286.

Fouse, A., Weibel, N., Hutchins, E., & Hollan, J. D. (2011, May). *ChronoViz: A system for supporting navigation of time-coded data.* Paper presented at CHI 2011, SIGCHI Conference on Human Factors in Computing Systems: Interactivity Track, Vancouver Canada.

Gliga, Teodora, & Csibra, Gergely. (2007). Seeing the face through the eyes: a developmental perspective on face expertise. *Progress in Brain Research*, *164*, 323–339.

Grelotti, David J, Klin, Ami J, Gauthier, Isabel, Skudlarski, Pawel, Cohen, Donald J, Gore, John C, Volkmar, Fred R, & Schultz, Robert T. (2005). fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia*, *43*(3), 373–385.

Guillon Q, Hadjikhani N, Baduel S, et al. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews*, *42*, 279–297.

Harrison, Ashley J, & Slane, Mylissa M. (2020). Examining How Types of Object Distractors Distinctly Compete for Facial Attention in Autism Spectrum Disorder Using Eye Tracking. *Journal of Autism and Developmental Disorders*, *50*(3), 924–934.

Hutchins, Tiffany L, & Brien, Ashley. (2016). Conversational topic moderates social attention in autism spectrum disorder: Talking about emotions is like driving in a snowstorm. *Research in Autism Spectrum Disorders*, *26*, 99–110.

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.

Klin, Ami, Jones, Warren, Schultz, Robert, Volkmar, Fred. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions. Biological Sciences*, *358*(1430), 345–360.

Klin, Ami, Jones, Warren, Schultz, Robert, Volkmar, Fred, & Cohen, Donald. (2002). Visual Fixation Patterns During Viewing of Naturalistic Social Situations as Predictors of Social Competence in Individuals With Autism. *Archives of General Psychiatry*, *59*(9), 809–816.

Klin, Ami, Danovitch, Judith H, Merz, Amanda B, & Volkmar, Fred R. (2007). Circumscribed Interests in Higher Functioning Individuals with Autism Spectrum Disorders: An Exploratory Study. *Research and Practice for Persons with Severe Disabilities*, *32*(2), 89–100.

Koegel, R., Fredeen, R., Kim, S., Danial, J., Rubinstein, D., & Koegel, L. (2012). Using Perseverative Interests to Improve Interactions Between Adolescents with Autism and their Typical Peers in School Settings. *Journal of positive behavior interventions*, *14*(3), 133–141.

Kryzak, L.A., Jones, E.A. The Effect of Prompts within Embedded Circumscribed Interests to Teach Initiating Joint Attention in Children with Autism Spectrum Disorders. *Journal of Developmental and Physical Disabilities*, *27*(3), 265–284.

Land, Michael F. (1992). Predictable eye-head coordination during driving. *Nature (London)*, *359*(6393), 318–320.

Lepper, Tracy L, Devine, Bailey, & Petursdottir, Anna Ingeborg. (2017). Application of a lag contingency to reduce perseveration on circumscribed interests. *Developmental Neurorehabilitation*, *20*(5), 313–316.

Moore, David J, Heavey, Lisa, & Reidy, John. (2012). Attentional Processing of Faces in ASD: A Dot-Probe Study. *Journal of Autism and Developmental Disorders*, *42*(10), 2038–2045.

Nadig, Aparna, Lee, Iris, Singh, Leher, Bosshart, Kyle, & Ozonoff, Sally. (2010). How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism. *Neuropsychologia*, *48*(9), 2730–2739.

Noris, Basilio, Nadel, Jacqueline, Barker, Mandy, Hadjikhani, Nouchine, & Billard, Aude. (2012). Investigating gaze of children with ASD in naturalistic settings. *PloS One*, *7*(9), e44144–e44144

Ozonoff, Sally, Iosif, Ana-Maria, Baguio, Fam, B.S, Cook, Ian C., Hill, Monique Moore, M.S, Hutman, Ted, Rogers, Sally J., Rozga, Agata, Sangha, Sarabjit, B.S, Sigman, Marian, Steinfeld, Mary Beth, M.D, & Young, Gregory S. (2010). A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(3), 256–266.e2.

Parish-Morris, Julia, Chevallier, Coralie, Tonge, Natasha, Letzen, Janelle, Pandey, Juhi, & Schultz, Robert T. (2013). Visual attention to dynamic faces and objects is linked to face processing skills: a combined study of children with autism and controls. *Frontiers in Psychology*, *4*, 185–185.

Riby, Deborah M, Riby, Deborah M, Brown, Philippa H, Brown, Philippa H, Jones, Nicola, Jones, Nicola, Hanley, Mary, & Hanley, Mary. (2012). Brief Report: Faces Cause Less Distraction in Autism. *Journal of Autism and Developmental Disorders*, *42*(4), 634–639.

Salley, Brenda, & Colombo, John. (2016). Conceptualizing Social Attention in Developmental Research. *Social Development (Oxford, England)*, *25*(4), 687–703

Sasson, Noah J, Elison, Jed T, Turner-Brown, Lauren M, Dichter, Gabriel S., & Bodfish, James W. (2011). Brief Report: Circumscribed Attention in Young Children with Autism. *Journal of Autism and Developmental Disorders*, *41*(2), 242–247.

Sasson, Noah J, & Touchstone, Emily W. (2014). Visual Attention to Competing Social and Object Images by Preschool Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *44*(3), 584–592.

Sasson, Noah J, Turner-Brown, Lauren M, Holtzclaw, Tia N, Lam, Kristen S.L, & Bodfish, James W. (2008). Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. *Autism Research*, *1*(1), 31–42.

Schultz, Robert T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience*, *23*(2-3), 125–141.

Shi, Lijuan, Zhou, Yuanyue, Ou, Jianjun, Gong, Jingbo, Wang, Suhong, Cui, Xilong, Lyu, Hailong, Zhao, Jingping, & Luo, Xuerong. (2015). Different visual preference patterns in response to simple and complex dynamic social stimuli in preschool-aged children with autism spectrum disorders. *PloS One*, *10*(3), e0122280–e0122280.

South, Mikle, Ozonoff, Sally, & McMahon, William M. (2005). Repetitive Behavior Profiles in Asperger Syndrome and High-Functioning Autism. *Journal of Autism and Developmental Disorders*, *35*(2), 145–158.

Tager-Flusberg, Helen, & Anderson, Marcia. (1991). The Development of Contingent Discourse Ability in Autistic Children. *Journal of Child Psychology and Psychiatry*, *32*(7), 1123–1134.

Vuilleumier, Patrik. (2002). Facial expression and selective attention. *Current Opinion in Psychiatry*, *15*(3), 291–300.

Wang, Y., & Holland, S. K. (2014). Comparison of functional network connectivity for passive-listening and active-response narrative comprehension in adolescents. *Brain Connectivity*, *4*(4), 273–285.

Wilson, C. E, Brock, J, & Palermo, R. (2010). Attention to social stimuli and facial identity recognition skills in autism spectrum disorder. *Journal of Intellectual Disability Research*, *54*(12), 1104–1115.

Zampella, Casey J, Bennetto, Loisa, & Herrington, John D. (2020). Computer Vision Analysis of Reduced Interpersonal Affect Coordination in Youth With Autism Spectrum Disorder. *Autism Research*, *13*(12), 2133–2142.