2015

Conversational topic moderates visual attention to faces in autism spectrum disorder

Ashley Rae Brien
University of Vermont

CONVERSATIONAL TOPIC MODERATES VISUAL ATTENTION TO FACES IN AUTISM SPECTRUM DISORDER

A Thesis Presented

by

Ashley Brien

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
For the Degree of Master of Science
Specializing in Communication Sciences and Disorders

May, 2015

Defense Date: March 9, 2015
Thesis Examination Committee:

Tiffany Hutchins, Ph.D. Advisor
Robert Althoff, M.D., Ph.D. Chairperson
Patricia Prelock, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

Autism Spectrum Disorder (ASD) is often accompanied by atypical visual attention to faces. Previous studies have identified some predictors of atypical visual attention in ASD but very few have explored the role of conversational context. In this study, the fixation patterns of 19 typically developing (TD) children and 18 children with ASD were assessed during a SKYPED conversation where participants were asked to converse about mundane vs. emotion-laden topics. We hypothesized that 1) children with ASD would visually attend less to the eye region and more to the mouth region of the face compared to TD children and that 2) this effect would be exaggerated in the emotion-laden conversation. With regard to hypothesis 1, we found no difference between groups for either number of fixations or fixation time; however, children with ASD did evidence significantly more off-screen looking time compared to their TD peers. An additional analysis showed that compared to the TD group, the ASD group also had greater average fixation durations when looking at their speaking partner's face (both eyes and mouth) across conversational contexts. In support of hypothesis 2, eye tracking data (corrected for amount of time during conversation) revealed two interaction effects. Compared to the TD group, the ASD group showed 1) a decreased number of fixations to eyes and 2) an increased fixation time to mouths but only in the emotion-laden conversation. We also examined variables that predicted decreased number of eye fixations and increased mouth-looking in ASD in the emotion-laden conversation. Change scores (to be understood as the degree of visual attention shifting from the mundane to the emotion-laden condition) for the ASD group negatively correlated with age, perceptual reasoning skills, verbal ability, general IQ, theory of mind (ToM) competence, executive function (EF) subscales, and positively correlated with autism severity. Cognitive mechanisms at play and implications for theory and clinical practice are considered.

AKNOWLEDGEMENTS

I extend my deepest gratitude to my parents, grandparents, brother, and fiancé who have provided me with a tremendous amount of support and encouragement. Thank you for giving me the confidence to pursue this project and for reminding me of my capabilities in moments of doubt.

I would like to thank the Jean S. Garvin Fellowship for providing financial support for this project. Thank you for encouraging research aimed at gaining a further understanding of Vermont's children and youth with disabilities.

I would also like to thank the director of the University of Vermont's Statistical Counseling Clinic, Alan Howard. Thank you, Alan, for liberally providing me with statistical consultation and guiding me through this project's statistical analyses.

I would like to thank my committee: Dr. Patricia Prelock and Dr. Robert Althoff—thank you both for your guidance and for giving me focus when the project was too large for me to tackle in the given time period.

Finally, I would like to thank my advisor, Dr. Tiffany Hutchins, for her support and dedication to this project from the start. Thank you, Tiffany, for teaching me to be confident in my skills and providing me with feedback along the way. You have provided me with more support than I could have imagined and I am honored to have worked with you for the past two years.

TABLE OF CONTENTS

Page

LIST OF APPENDICES

Page

LIST OF FIGURES

Page

Chapter 1: Literature Review

Autism Spectrum Disorder (ASD) is often accompanied by atypical visual attention to social stimuli including attention to people's bodies and faces (Riby & Hancock, 2009), as well as marked impairment in eye contact during social interaction (Doherty-Sneddon, Riby, & Whittle, 2012; Lord, Risi, Lambrecht, Cook, Leventhal, & Di Lavore, 2000). Deficits in emotion and face recognition are also common (Baron-Cohen, Wheelwright, & Jolliffe, 1997). To better understand these atypicalities, much research has focused on how individuals with ASD and their typically developing (TD) peers process faces. Although mixed results in the literature have occurred, these discrepancies are partly accounted for by variation in task demands, methods of study, and participant characteristics (e.g., chronological age, language level, gender, ASD severity) (Harms, Martin, & Wallace, 2010).

Studies of ASD and TD between-group differences have been valuable for illuminating potential mechanisms that may be operating to produce atypical face processing in ASD (Baron-Cohen et al., 2000; Doherty-Sneddon et al., 2012). On the other hand, it is firmly established that ASD is not a monolithic disorder and what has rarely been addressed is what factors predict individual differences in face processing within this population. In a related vein, little is known about the contextual features that influence visual attention in ASD. This study aims to address these limitations in the literature by examining how conversational topic affects visual attention. This is accomplished by comparing visual attention during a mundane conversation in which children are asked to talk about "things that people do" and an emotion-laden conversation that is designed to place demands on children's 'theory of mind' and asks children to talk about "things that people feel." The study also aims to identify within-group participant characteristics that predict patterns of visual attention to faces in ASD and TD.

This literature review will 1) describe the theory of mind deficit characteristic of ASD, 2) describe findings from behavioral and eye tracking research in face processing in ASD, 3) discuss findings from brain imaging studies in face processing in ASD, 4) examine within-group differences in ASD for atypical visual fixation patterns, and 5) discuss the importance of context and its effects on the looking patterns of persons with ASD. It will ultimately be argued that an examination of between and within-group differences to compare performance between two conversational contexts is an important contribution for understanding visual attention to faces in ASD.

1.1 Theory of Mind Deficits of ASD

Theory of Mind (ToM; also commonly referred to as social cognition, perspective-taking, mentalizing, or mind-reading, Hutchins & Prelock, in press) may be defined as the ability to infer the mental states of others so as to know "that other people know, want, feel, or believe things" (Baron-Cohen, Leslie, & Frith, 1985, p. 38). Baron-Cohen (1995/2001) described ToM as a key component of mindreading and reported numerous findings in which children with ASD fail to employ a ToM, thus demonstrating "mindblindness" (p. 5) or the inability to read others' mental states. Baron-Cohen et al. (1985) suggested that ToM is a core cognitive deficit of ASD and found that those with ASD underperformed on false belief tasks compared to those with Down syndrome and TD controls despite the fact that the TD group and group with Down syndrome exhibited lower verbal and non-verbal MA than the group with ASD. The authors concluded that the inability to pass the false belief task was not due to other cognitive or performance factors and that participants with ASD differed from normal controls and those with severe intellectual disability precisely because they were unable to "appreciate the difference between their own and the doll's knowledge" (p. 43). While Baron-Cohen et al. (1985) found that individuals with ASD experience difficulty distinguishing their own from others' perspectives (and this was taken as evidence for an impaired ToM), some individuals with ASD routinely demonstrate an ability to employ various aspects of ToM (Baron-Cohen et al., 1985; Bowler, 1992; Happe, 1993; Ozonoff, Pennington, & Rogers, 1991) lending credibility to the notion that ToM is not uniformly affected in ASD. These findings gain importance in light of the fact that ToM is a complex and multifaceted construct. Indeed, deficits in ToM in ASD are not monolithic (Hutchins & Prelock, in press).

The ToM model has proven to be a particularly influential hypothesis for explaining the social deficits characteristic of ASD (Baron-Cohen, 1992, 1995/2001) and has gained considerable currency in the face processing and eye tracking literature as well. According to Hernandez and colleagues (2009), "face perception plays a critical role in the development of social interaction and understanding of the internal emotional state of others" (p. 1004) and the eye region of the face is particularly important for recognizing others' mental states. In a related vein, Riby and Doherty (2009) used eye tracking technology to study the ability of children with ASD to detect an examiner's gaze direction and found that young participants with ASD had difficulty detecting the target of the examiner's gaze. In short, gaze direction can indicate the mental states of others, is considered one measure of ToM (as it involves joint attention and intentionality; see Baron-Cohen, 1995/2001), and children with ASD tend to demonstrate difficulty in gaze monitoring (Mundy, Sigman, & Kasari, 1996). Similarly, Rutherford and Towns (2008) noted that participants with ASD spent less time looking at the eye region of faces when presented with complex emotions compared to simple emotions. This finding is consistent with the results of Baron-Cohen et al. (1997), who concluded that individuals with ASD demonstrate an ability to process simple emotions (e.g., happy, sad, angry), but tend to have trouble with the recognition of complex mental states (e.g., admire, interest, thoughtfulness) and implicates atypical visual attention in the ability to perform ToM tasks.

ToM and Executive Function. It is important to note that ToM dysfunction overlaps considerably with deficits in Executive Function (EF) (e.g., Hill, 2008; Pellicano, 2010). Ozonoff et al. (1991) have underscored the importance of EF, noting that persons with ASD are often deficient in EF and tend to be impulsive, have poor response inhibition, and lack future-oriented planning skills. Indeed, performance on EF tasks tends to predict performance on ToM tasks (Hughes, 1998), although the nature of the relationship between these constructs and their direction of influence remain a topic of debate (Pellicano, 2007, 2010). These considerations foreshadow the difficulty encountered when trying to disentangle effects and identify which constructs contribute to atypical patterns in core versus secondary ways (Happé, Roland, & Plomin, 2006; Pellicano, 2010).

Riby and Hancock (2009) used both static (e.g., picture) and dynamic (e.g., movie) stimuli to track the eye gaze of participants with ASD, William's syndrome (WS), and those who were typically developing. Compared to TD peers matched on age and non-verbal ability, participants with ASD showed reduced attention to the movie stimuli suggesting that they are less interested in or less able to process these types of complex stimuli. If this is the case, it is possible that the "complexity of the movie information in some way 'distracts' or 'overloads' the attention of participants with autism in a way that is not possible for static images" (p. 179). The potential cognitive overload associated with complex stimuli may also mean that if persons with ASD are not accessing social information from complex stimuli, they will have fewer opportunities to (naturally) learn social cues (Riby & Hancock, 2009).

Of course (and like ToM), EF is a complex construct that may be carved and construed in different ways. One approach to EF divides it into two categories: hot and cold (Castellanos, Sonuga-Barke, Milham, & Tannock, 2006). Cold cognition comes into play during problems that are decontextualized and abstract (e.g., the Stroop task). Hot cognition, however, is required for tasks with affective or motivational significance (Castellanos et al., 2006). Hot cognition connects with ToM insofar as the attribution of mental states involves the affective domain. As Zelazo, Qu, and Müller (2013) have argued "ToM is hot EF as expressed in the content domain of self and social understanding" (p. 86).

A different approach for understanding the connection between EF and ToM involves examination of two subcomponents of EF: working memory and inhibition (see Moses, Carlson, & Sabbagh, 2005/2013). Working memory has been described as "the maintenance of transient information over brief temporal intervals to direct future-oriented activity" (Welsh, 2002/2008, p. 144), whereas inhibition is the ability to suppress inappropriate and maladaptive responses and engage in those behaviors that are more appropriate. Moses et al., (2005/2013) described a number of studies that discuss the correlation between ToM and the working memory/inhibition components of EF. These authors propose that failure on false belief tasks may be due to an inability to remember one's own belief in light of the protagonist's belief (i.e., a working memory hypothesis) or the inability to suppress one's knowledge of facts evident in the task (i.e., an inhibition hypothesis). In summary, the ToM hypothesis of ASD has been linked to atypical face processing in ASD. While ToM and EF are associated in TD and ASD, the direction of effects is unclear and the nature of our understanding of ToM and EF connections is influenced by how we construe each: a topic that will be revisited later in this manuscript.

1.2 Face Processing in ASD

The ability to recognize faces and process both basic and complex emotions is an important skill because it provides information to guide successful social interactions (Ashwin, Wheelwright, & Baron-Cohen, 2006) including a person's mood, intentions, attentiveness, and identity (Bruce & Young, 1986). This selective review will describe some earlier studies that used static face stimuli. This will be followed by a discussion of more recent studies that used dynamic stimuli.

Some earlier studies using static stimuli. Baron-Cohen et al. (1997) tested the ability of adults with ASD to identify basic emotions and complex mental states using whole-face photographs and photographs depicting only the eye-region of the face. Results showed that participants with ASD were relatively good at detecting basic emotions but relatively impaired at recognizing the complex mental states in the whole-face condition. Further, the ASD participants were more impaired at identifying the complex mental states in the eyes-only condition. The ability to recognize others' emotions is essential to social functioning and the difficulty that participants with ASD demonstrated in the eyes-only task suggests that persons with ASD are less inclined or less able to make use of this information.

Other research has found impaired face recognition to be dependent on certain Areas of Interest (AOIs). Using eye tracking technology (discussed more fully below), Joseph and Tanaka (2003) noted that persons with ASD had an increased ability to recognize static faces when recognition was dependent upon the mouth region as opposed to the eye region of the face. The opposite effect was found for the TD group who relied more heavily on the eye region and less so on the mouth region for facial recognition. This evidence suggests that when deciphering emotional cues, persons with ASD tend to rely on information in the mouth region as opposed to the social cues provided by the eyes. Langdell (1978) concluded that if children with ASD do not view the eyes as social stimuli, then the eyes and the mouth must "rank equally as the most easily discriminable areas if the human face" (p. 265), lending a possible explanation for mouth-looking in ASD. This is also consistent with the contentions of Baron-Cohen et al. (1997) that persons with ASD are not gathering information from the eyes and are "not acquiring a language of the eyes in the same way [as their peers]" (p. 329).

Results are mixed for the visual allocation patterns observed in ASD when studies employ static stimuli. Snow and colleagues (2011) found that, like their TD peers, persons with ASD showed a strong preference for the eye region over other areas of the face. In a study examining adults with Asperger Syndrome (AS) and TD matched controls, Falkmer, Larsson, Bjällmark, and Falkmer (2010) determined which regions participants attended to during face recognition tasks. By using puzzle pieced photos as well as intact photos, Falkmer and colleagues found that participants with AS had more difficulty in the facial recognition task when the eye region in the puzzled-piece photo was distorted or not fully intact. These results demonstrate a strong reliance on the eye region in face recognition tasks for participants with AS and does not support previous research that participants with ASD show a strong tendency toward fixating on the mouth area in photographs. Static stimuli, however, may not provide adequate information about real-life to observers with ASD and may yield "decreased ecological validity" (McPartland, Webb, Keehn, & Dawson, 2011, p.153). Dynamic stimuli provide more naturalistic social information to which individuals with ASD may have difficulty attending as evidenced by their atypical fixation patterns (Klin, Jones, Schultz, & Volkmar, 2003).

Studies using dynamic stimuli. More recently, eye tracking technology has been employed to examine atypicalities in visual attention to dynamic social stimuli in ASD. These types of studies allow researchers to more fully explore the visual patterns associated with ASD, which may prove valuable for understanding how persons with ASD view and understand the social world (Freeth, Chapman, Ropar, & Mitchell, 2010; Jones & Klin, 2008; Riby & Hancock, 2008). Jones and Klin (2008) suggested that a child who does not focus on socially relevant stimuli (e.g., a person's face or eye gaze) may see a world in which physical events are more salient than social cues. In effect, this could alter the child's development to the extent that "the child's mind would specialize on physical contingencies rather than on social beings. Consequently, it would suggest that later in life [the child's] brain and…behavior will also show equally atypical specialization" (p. 69).

Shic, Bradshaw, Klin, Scassellati, and Chawarska (2011) examined the gaze patterns of toddlers with ASD compared to TD and developmentally delayed (DD) peers. They found that, while watching another child and adult interact, toddlers with ASD spent significantly more time than TD and DD peers attending to the background of the scene compared to the activity taking place between the adult and child. Moreover, while the toddlers with ASD and those who were TD and DD spent similar amounts of time attending to the *people* in the scene, closer examination revealed that children with ASD spent more time looking at the person's body compared to the person's head. In a similar study by von Hofsten, Uhlig, Adell, & Kochukhova (2009), findings revealed that compared to TD controls, young children with ASD spent less time fixating on the speaker's face when the speaker was engaged in a conversation with a listener. TD controls also looked significantly longer at the speaker's face than did participants with ASD.

Research also suggests that older children with ASD show a preference for non-face stimuli when viewing dynamic scenes. For example, Bird, Press, and Richardson (2011) found that children with ASD spent less time fixating on face stimuli compared to TD controls when viewing video stimuli. In addition to atypical looking patterns described above, dynamic stimuli studies often find that when the face is the focus of attention, participants with ASD show a preference for looking at the mouth region instead of the eyes (Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Norbury et al., 2009). To better understand the real world implications for looking patterns and social competence, Klin et al. (2002) used eye tracking technology to examine the visual allocation patterns of adolescents and young adults with ASD and TD participants while they watched emotionally charged videos that were more analogous to real-life situations compared to static photos. Researchers chose video clips from the film "Who's Afraid of Virginia Woolf?" because the "demanding social complexity in the movie mirror[ed] complicated social situations that individuals with autism may encounter in everyday settings" (p. 811). Results were consistent with the authors' predictions that individuals with ASD would focus more on the mouth, body, and object regions compared to the eye region (TD peers spent most of the time looking at the eye region). These results were replicated in Norbury et al., (2009) who examined the visual fixation patterns of adolescent ASD and TD participants. These researchers used less emotionally complex videos than those of Klin et al., (2002) but the video clips portrayed social interactions that prompted a variety of emotional responses. Results indicated that verbally able adolescents with ASD spent less time looking at the eyes and more time looking at the mouth compared to those participants with ASD that were language impaired or typically developing (Norbury et al., 2009). Similarly, Bird et al. (2011) found a significant preference for the eye region of the face in the TD control group, but not for participants with ASD.

Of interest to the present study are also the within-group findings of visual fixation patterns observed in ASD and the looking behaviors associated with different conversational contexts. Where Klin et al. (2002) found significant between group differences, further analyses revealed important within-group differences. Specifically, in the ASD group, longer mouth fixation time was associated with *higher* levels of social adaptation (as defined by scores on the Vineland Adaptive Behavior Scales, Expanded Edition) and *lower* levels of social impairment (as defined by scores on the Autism Diagnostic Observation Schedule). Moreover, longer object fixation time was associated with lower levels of social adaptation and higher levels of social impairment. Norbury et al. (2009) also looked at within-group differences and concluded that language ability and "autistic symptomology" (p. 839) were important predictors of functioning.

1.3 Visual Attention and Conversational Context

The current study aims to examine visual attention to a speaking partner when participants engage in different topics of conversation. Few studies have reported on the effects of conversational context on visual fixation patterns; however Chawarska, Macari, and Shic (2012) studied the visual fixation patterns of participants with ASD during socially relevant and non-social stimuli. The study consisted of four dynamic scenes (i.e., Dyadic Bid, Sandwich, Joint Attention, Moving Toys). In the Dyadic Bid scene the actress on the screen looked directly at the camera and engaged in child-directed speech (it is important to note that this was not an actual conversation in which the child was engaged, but involved the actress talking at the child through the camera); in the Sandwich scene, the actress looked down at the table, made a sandwich, and was not engaged in speech or eye contact with the camera; in the Joint Attention scene, the actress looked at the camera, said "uh-oh," and then turned her head to look at one of the toys in the scene; in the Moving Toys scene, the actress looked at the camera at the same time that a toy began moving and making noise, and then the actor looked at the toy on the opposite side of the screen than the moving toy. The authors found that, like the TD children, the children with ASD did not engage in atypical looking patterns due to the mere presence of the actress on the screen. The children with ASD did, however, show atypical looking patterns (particularly reduced attention to the scene, the actress's face, and the actress's mouth) when viewing the Dyadic Bid scene in which the actress was looking directly at the camera and using child-directed speech. Interestingly, the Joint Attention scene did not demonstrate atypical looking patterns to the same extent as the Dyadic Bid scene, likely due to the limited eye contact and speech production used in this scene. Chawarska and colleagues concluded that "limited attention to faces appeared context-dependent and was linked to the presence of explicit cues for dyadic engagement" (p. 909). These findings provide support for atypical visual fixation patterns associated with varying contexts, however more research is needed to build on the evidence presented by Chawarska et al. (2012).

1.4 Brain Imaging Studies in ASD

Recent brain-imaging studies have provided researchers with data regarding the neural abnormalities of persons with ASD. These studies, which have been used in conjunction with behavioral and eye tracking studies, have resulted in an increased knowledge about the underlying mechanisms involved in the facial emotion recognition deficits present in ASD (Harms et al., 2010). This research is in line with the amygdala hypothesis of autism, which proposes that the features of ASD are rooted in amygdala dysfunction and, indeed, several researchers have cited similarities in the performance of persons with ASD and those with amygdala dysfunction (Ashwin et al., 2006; Baron-Cohen et al., 2000; Spezio, Huang, Castelli, & Adolphs, 2007). Many of the abilities that are negatively affected are linked to lack of direct eye gaze (Grice et al., 2005), impaired judgment of trustworthiness of others (Adolphs, Sears, & Piven, 2001), and facial/emotional recognition (Baron-Cohen et al., 2000). Much of the research has focused on activation of the amygdala and the fusiform gyrus (FG).

Decreased activation of the amygdala is evident in neuroimaging studies monitoring the brain activity of persons with ASD while processing facial emotions (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, 2007; Pelphrey, Morris, McCarthy, & Labar, 2007) and mental states (Baron-Cohen et al., 2000). Ashwin et al. (2006) demonstrated that the impairments seen in participants with ASD in recognizing negative emotions (e.g., fear, anger, disgust) are similar to those facial emotion recognition deficits observed in individuals with considerable amygdala damage. Spezio et al. (2007) heightened these findings through a study of an individual with amygdala damage and her impairments in eye contact during live conversations. These researchers found that the impairments associated with extensive amygdala damage are similar to those observed in persons with ASD. It is suggested that amygdala dysfunction and those corresponding impairments seen in ASD negatively affects the ability to link social stimuli with social meaning (Adolphs et al., 2001).

In addition to the amygdala, activation of the fusiform gyrus is typically evident during both face perception and recognition of emotions (Baron-Cohen et al., 2000; Dawson, Webb, & McPartland, 2005; Grelotti, Gauthier, & Schultz, 2002; Grelotti et al., 2005; Harms et al., 2010; Jemel, Mottron, & Dawson, 2006; Schultz, 2005). In TD individuals, the fusiform face area (FFA) is activated by face perception, as well as other non-face objects for which the individual has object expertise (Jemel et al., 2006). Neuroimaging studies have found that persons with ASD show decreased activation of the FFA when processing faces. Interestingly, it has been reported that when individuals with ASD process faces, their inferior temporal gyrus is activated, which is typically the area of the brain that is used to process objects in TD individuals (Schultz et al., 2000).

Event-related potentials (ERP) in response to facial stimuli have provided additional insight into the ways people process faces. More specifically, primary visual information is believed to be processed at a low level using bottom-up (i.e., stimulus-driven) processing. Information is aggregated at the low-level and subsequently sent to higher levels where the organization of information is refined and details of information are elaborated. Top-down (i.e., knowledge-driven) processes are also triggered early in visual information processing whereby concept-level information contributes to category identification and object recognition. In short, bottom-up processing may occur as details of a visual stimulus are processed independently of the larger concept but the two are ultimately linked to form a complete visual representation of the object. Maekawa and colleagues (2011) found that individuals with ASD demonstrate abnormal top-down processing strategies, yet intact bottom-up processing abilities. They concluded that individuals with ASD use feature-based strategies, as opposed to holistic-based strategies, in FER tasks due to the abnormal processing abilities that occur early on in the process of visual sensory information. In summary then, information processing at the neuroanatomical (i.e., in particular the FFA and amygdala) has been implicated.

1.5 Summary and Statement of the Problem

Evidence is mixed regarding visual attention to faces in ASD. While some researchers did not find differences between the visual allocation patterns of participants with ASD and those who are TD (Snow et al., 2011; Falkmer et al., 2010), these studies reported on visual attention to static stimuli. Static stimuli do not engage participants in a way that mimics real life nor does it provide information in socially relevant contexts (McPartland et al., 2011). Dynamic studies, on the other hand, suggest that participants with ASD employ atypical looking patterns when viewing faces (Klin et al., 2002; Norbury et al., 2009). The mixed findings between stimuli suggest that the looking patterns (in those with ASD) toward static and dynamic stimuli differ. Even fewer studies examine how visual attention is affected by conversational contexts. The following study addresses these gaps in the literature by examining the within- and between-group differences in the visual attention of TD and ASD. This will be accomplished using a Skyped conversation where participants with ASD and TD are asked to engage in a conversation about “things that people do” (the mundane context) and “things that people feel” (the emotion-laden context). Based on the aforementioned research, we hypothesized that 1) children with ASD would visually attend less to the eye region and more to the mouth region of the face compared to TD children and that 2) this effect would be exaggerated in the emotion-laden conversation. We also sought to explore factors related to atypical visual allocation in ASD during conversation. This was accomplished by examining the correlations between factors of interest (e.g., ToM, EF) and change scores for any eye tracking data that showed a significant effect in the omnibus analysis.

Chapter 2: Method

2.1 Participants

Participants were 19 typically developing children (15 males, 4 females) ages 6 years, 3 months to 12 years, 11 months (*M* = 8 years, 8 months, *SD* =2.23) and 18 children (15 males, 3 female) ages 6 years, 1 month to 11 years, 9 months (*M* = 9 years, 3 months, *SD* = 1.55) diagnosed with ASD. All typically developing children were identified on the basis of parental report. More specifically, parents responded to a questionnaire designed to screen for a variety of conditions. Parents were asked to report whether their child had ever received a diagnosis or were ever concerned about the presence of a developmental delay (including ASD), learning impairment, speech and language impairment, and uncorrected visual or hearing impairment.

On the basis of parental report, four children had a diagnosis of autism, six had a diagnosis of Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS), and eight had a diagnosis of Asperger Syndrome (American Psychiatric Association, 2000). All children had been diagnosed by a psychologist or developmental pediatrician. Six children in the ASD group also had a concomitant diagnosis of Attention-Deficit Hyperactivity Disorder or Attention-Deficit Disorder. One child in the ASD group was functionally nonverbal (characterized by parental report as having "limited speech") but was able to attend to stimuli making collection of the eye tracking data possible. All 17 remaining children were verbal and could use language functionally and flexibly.

2.2 Measures

As described in this paper's introduction, several constructs have been implicated in atypical visual attention to faces in ASD. Measures for many of these constructs (ToM, EF, autism severity) were employed in this study to examine whether they predicted atypical visual attention. Data for general and subscale intelligence were also collected. These data were included in the predictor analyses but also used to evaluate whether our ASD and TD groups were distribution matched on general intelligence.

2.2.1 Behavior Rating Inventory of Executive Function. The *Behavior Rating Inventory of Executive Function* (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000) is a parent and/or teacher informant measure designed to assess executive function behaviors in individuals ages 5 through 18. The test is composed of 86 items divided into eight categories: Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. Scores are reported for each of the categories as well as an overall behavioral index, a metacognition index, and a global executive composite. The BRIEF has been evaluated for reliability (internal consistency, test-retest, and interrater) and validity (convergent and divergent) and has demonstrated adequate psychometric properties (Schraw, 2003).

2.2.2 The Gilliam Autism Rating Scale-2. The *Gilliam Autism Rating Scale, second edition* (GARS-2; Gilliam, 2006) is a norm-referenced rating scale based on the definitions of autism adopted by the Autism Society of America and the DSM-IV. The GARS-2 is typically used as a screening tool with children between the ages of 3 and 22 who show signs indicative of ASD. There are 42 items separated into three subscales: communication, social interaction, and stereotyped behaviors. This scale assesses behaviors using objective frequency-based ratings by individuals familiar with the individual and takes approximately 5-10 minutes to complete. The combined scores on these subscales yield an autism index (AI) score (with a mean of 100 and *SD* of 15), which provides a total score assessing the probability of autism and the degree of severity. Statistically significant validity and reliability (internal consistency, test-retest, and interrater reliability) were reported for each of the test domains and the AI (Lopez-Wagner, Hoffman, Sweeney, & Hodge, 2008).

2.2.3 The Theory of Mind (ToM) Task Battery. The *ToM Task Battery* (Hutchins, Prelock, & Chace, 2008) consists of 13 questions designed to tap a range of ToM tasks. The items range in complexity and are presented in the form of static visual stimuli. The initial task tests the ability to identify emotions associated with facial expressions, the second asks children to infer an emotion associated with a desire. The remaining tasks assess advanced abilities including belief-based emotion, reality-based emotion, second-order belief-based emotion, perception-based belief, and false beliefs. The *ToM Task Battery* has been evaluated for reliability (test-retest, internal consistency), which was adequate (Hutchins et al., 2008).

2.2.4 The Wechsler Abbreviated Scale of Intelligence. The *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler, 1999) is a brief norm-referenced IQ test for individuals between the ages of 6 and 89. The WASI is composed of four subtests: vocabulary, similarities, block design and matrix reasoning, yielding a verbal IQ score, a performance IQ score, and a full IQ score. The WASI has been evaluated for reliability (internal consistency, test-retest) and validity, and has demonstrated good psychometric properties (Lindskog & Smith, 2010).

2.3 Apparatus

Eye movements were recorded using the Mirametrix S2 Eye Tracker System to record X and Y coordinates of eye position. The screen-capture system promoted more natural behavior since it did not place restraints on participants such as a helmet, head-mounted sensor, or glasses. The system utilized a sampling rate of 60 Hz yielding an accuracy of .5-1 degree of the visual angle. Eye blinks and off screen gazes were identified by loss of corneal reflection and were excluded from the data. Participants were seated at a desk in front of a 22-inch computer monitor (1680 X 1050 pixels resolution) located approximately two feet away. The eye tracker was positioned just below the computer screen. Presentation of the stimuli was captured using Viewer software in Mirametrix and the resulting data was managed by conducting analyses of CSV (Comma Separated Values) files. Fixation calculation parameters for gaze were set at 20 pixels (for the maximum distance in pixels that a point may vary from the average fixation point and still be considered a fixation) and 2 samples (minimum number of samples to be considered a fixation) with 3 degrees of visual angle. Participants' gazes were calibrated quantitatively in the following manner:

> The user is required to look at these coordinates, in such a manner that the system associates to each of these a specific relative position of both the flint and pupil centers. Once these nine points are successfully recorded (about 15 seconds), the system is able to track the point-of-regard in every position of the screen, by means of computer vision techniques and trigonometric calculations. The mirametrix S2 device specifies that there will never be a drift over 0.3 degrees. Furthermore the device takes less than 16ms to reacquire the eyes image in case of need. Following the official device' specifications, its accuracy is in the range of 0.5-1 degrees of visual angle, meaning that with the user staying at 50 cm from the device, the error in the screen is going to be in the range of 0.44 cm to 0.87 cm approximately (Barral, 2013, p. 17).

2.4 Conversation Types

Two different conversational contexts were examined: a conversation about things that people do (the mundane conversation) and a conversation about things that people feel (the emotion-laden conversation). The primary investigator participated in a Skype conversation with the participant by asking the participant a variety of questions about mundane and emotion-laden topics (see Appendix 1 for a full list of the questions asked). For each question that the primary investigator asked, she attempted to elicit at least two responses from the child. Once a minimum of two responses were secured, the experimenter continued to the next question.

2.5 Procedure

Participants with ASD were recruited via informal contacts as well as notices to local support agencies for families with children with ASD. Additionally, six participants with ASD were recruited through an ad placed in the local newspaper. Participants with typical development were recruited via fliers and informal contacts. Parents received $25.00 compensation for participation in the study. Upon arrival to the laboratory, informed consent was obtained from a parent. The children were seated in front of the computer screen for the Skype conversation. The children were given the instruction "find your sweet spot" if they were looking at the computer screen but their eyes were not picked up by the eye tracker. Participants' eye gazes were calibrated using quantitative measures (described above) as well as visual confirmation (i.e., they were asked to look at particular objects on the screen and it was noted where their eyes were fixating on the computer screen). Visual confirmation checks were informal and used as necessary. Children were not repositioned if they were looking off the screen during thinking time or when answering questions so as to simulate a real conversation where it is not required to use constant eye contact. The primary investigator would ask the question "Tell me about things that people do for work." After the child gave at least two responses, he/she would be asked another question. If the child had difficulty giving two answers, the primary investigator would say, "Can you think of anything else that people do for work?" If the child answered with, "No" or "I don't know," a new questioned was asked. There were no time limits on how long a child's answer was or how long the child engaged in think time.

One random order for the stimuli was determined with the order in which the questions were asked. The order of conversational context was counterbalanced so that each order (do and feel) was as equally represented as possible. This study was part of a larger series of eye tracking studies using different sets of stimuli. Each set of stimuli was also presented in a counterbalanced order so that each order was as equally represented as possible.

After the gathering of the eye tracking data during the Skype conversation, the ToM Task Battery and the WASI were administered (in that order). During this time, parents were asked to complete the BRIEF and the GARS-2. The completion of all data collection procedures took between two to three hours.

2.6 Dependent Variables

Areas of interest (AOIs) were chosen based on the upper and lower regions of the face, as well as "other" AOIs. The upper region included the eyes and brows and extended to the temple area of the face. The lower region included the mouth. All "other" AOIs included all other areas of the screen, as well as off-screen looking time (see Figure 1). Data were collected for the number of fixations and the total proportional fixation time for each AOI in each condition (i.e., doing, feeling).

Research examining eye gaze to faces suggests that the nature of the viewing task influences spatial (e.g., fixation location) and temporal parameters (e.g., fixation duration) and some researchers have proposed independent mechanisms for control of these parameters (Mills, Hollingworth, Van der Sticgchel, Hoffman, & Dodd, 2011). Because we were interested in how conversational topic might influence spatio-temporal properties of eye gaze to faces, we examined the number of fixations and total fixation time to specific AOIs. Information about the number of fixations provides data regarding goal-directed visual search and reflects real-time orienting decisions based on current cognitive needs. By contrast, fixation duration is “crucially linked to information processing and can reflect ongoing visual processing during scene viewing” (Mills et al., 2011, p. 2). Greater fixation time is typically associated with increased cognitive demand, as when processing difficult words in a reading task (Pollatsek, Rayner, & Bolata, 1986) or adding noise during a speech intelligibility task (Buchan, Pare, & Munhall, 2008).

Chapter 3: Results

3.1 Participant Characteristics

Data for matching variables were submitted to a series of independent samples t-tests. No differences were found for the variables of child age or gender ($p > .45$ for each). Differences were observed for the WASI-2 verbal section ($p = .003$) and for the WASI-2 full scale ($p = .02$). It should be noted that these differences appear to be due to *over*-performance by the TD group which were typically 1 SD above the mean (WASI-2 verbal section $M = 116.00$; $SD = 16.42$; WASI-2 full scale $M = 112.84$; $SD = 14.90$). The average scores for the group with ASD were well within normal range for the verbal section ($M = 94.89$; $SD = 23.75$), as well as for the WASI-2 full scale ($M = 98.44$; $SD = 20.66$). All participants with ASD also completed the GARS-2 ($M = 87.11$; $SD = 16.23$), indicating that our sample generally fell within the range of mild autism severity and were considered high-functioning.

Boneferroni tests to correct for family-wise error rate were not conducted for these data. Given the large number of comparisons and the exploratory nature of this research, it is important to protect against the likelihood of multiple Type II errors (Hewes, 2003).

3.2 Hypothesis 1

Recall that hypothesis 1 was: Children with ASD will attend less to the eye region and more to the mouth region of the face compared to TD children in both conversational contexts. To address this hypothesis, data for mean number of fixations and mean fixation time for the three AOIs (i.e., eyes, mouth, and background) were first analyzed descriptively. These data are presented in Figures 2 and 3. Because time varied across groups and conditions, it was controlled in all subsequent analyses by dividing the values for the variables of interest by total time spent during interaction. Data were then analyzed inferentially using a series of 2 (group: ASD, TD) X 2 (condition: mundane, emotion-laden) mixed model Analyses of Variance (ANOVAs) with repeated measures on condition. No main effects for group were found for either number of fixations or fixation time data. Thus, contrary to our expectation, participants with ASD did not attend less to the eyes and more to the mouth region compared to TD peers in both the mundane and emotion-laden conversations, suggesting that looking patterns across conditions were similar across groups.

However, a main effect for group was observed for additional t-tests that we conducted. First, we found a main effect for total off-screen looking time (seconds), $F(1,33.51) = 8.19$, $p < .01$, such that participants with ASD spent significantly more time looking off-screen ($M = 244.1$; $SD = 54.4$) compared to their TD peers ($M = 26.4$; $SD = 53.0$) regardless of condition. These data are presented in Figure 4. We also observed a main effect for group for average fixation duration to eyes such that participants with ASD had significantly longer fixation durations in the doing condition ($M = .42$; $SD = .24$) compared to the TD group ($M = .28$; $SD = .07$), $t(35) = 2.39$, $p = .03$, and the ASD group had significantly longer fixation durations to the eyes in the feeling condition ($M = .41$; $SD = .27$) compared to the TD group ($M = .30$; $SD = .11$), $t(35) = 1.71$, $p = .11$. Finally, a main effect for average fixation duration to the mouth was observed, such that the ASD group had significantly longer fixation durations in the doing condition ($M = .38$; $SD = .29$) compared to the TD group ($M = .17$; $SD = .10$), $t(35) = 2.99$, $p = .01$, and the ASD group had significantly longer fixation durations to the mouth in the feeling condition ($M = .41$; $SD = .32$) compared to the TD group ($M = .17$; $SD = .11$), $t(35) = 3.13$, $p = .01$. No significant differences were found between groups and across conditions for average fixation duration to the background. These data are presented in Figures 5 and 6. In summary, number of fixations to eye and mouth and total fixation time to eye and mouth did not differ between groups, but the group with ASD looked off-screen more than the TD group. When looking at eye and mouth regions (but not the background), the group with ASD also had longer average fixation durations compared to the TD group.

3.3 Hypothesis 2

Hypothesis 2 was: Children with ASD will visually attend less to the eyes and more to the mouth during conversation and that this effect would be exaggerated in the emotion-laden condition. In short, conversational topic will moderate visual attention in ASD such that we should observe less eye looking and more mouth looking in the emotion-laden condition compared to the mundane condition.

The total looking time to eye, mouth, and background AOIs were submitted to three mixed model 2 (group: typically developing vs. ASD) X 2 (condition: doing, feeling) ANOVAs with repeated measures for condition. All analyses were corrected for time. An interaction was observed for total (cumulative) time looking at the mouth AOI, such that participants with ASD spent significantly more time (seconds), $F(1, 34.50) = 4.87$, $p < .05$), looking at the mouth in the feeling condition ($M = 35.46$; $SD = 5.00$) compared to TD peers ($M = 17.76$; $SD = 4.75$). Time spent looking at the mouth was not significantly different between groups in the doing condition. This interaction is represented in Figures 7 and 8.

The total number of fixations to eye, mouth, and background AOIs were submitted to three mixed model 2 (group: typically developing vs. ASD) X 2 (condition: doing, feeling) ANOVAs with repeated measures for condition. An interaction was found for number of fixations to the eye region of the face. Participants with ASD fixated significantly less on the eye region, $F(1, 34.43) = 5.74$, $p < .05$), in the feeling condition ($M = 105.81$; $SD = 15.47$) compared to TD peers ($M = 167.82$; $SD = 14.66$). Number of fixations to the eye AOI was not significant in the doing condition. This interaction is represented in Figures 8 and 9.

3.4 Research Question 1

Recall that research question 1 was: What factors predict visual allocation during conversation? To address this question, data for total time and number of fixations to eye, mouth, and background regions (i.e., outside eye and mouth AOIs) were correlated with measures of interest (i.e., ToM, EF, IQ, autism severity, age). Results indicated high collinearity among all variables presenting challenges to both statistical analysis and interpretation. To address this issue, change scores were calculated for dependent variables that showed a significant effect in our mixed model analyses reported above. That is, change scores were used to quantify the degree of shift in visual attention revealed in our two interaction effects. In the ASD group, attention away from the eye region (fewer fixations) in the emotion-laden condition was negatively correlated with age ($p = .04$), perceptual reasoning ($p = .04$), and EF subscales: working memory ($p = .003$) plan/organize ($p = .001$), and organization of materials ($p = .015$), and positively correlated with autism severity ($p = .04$). In the ASD group, increased attention (total amount of time) to the mouth region in the emotion-laden condition was negatively correlated with verbal IQ (p = .014), general IQ (p = .012), ToM competence (p = .035), and positively correlated with autism severity (p = .047). Increased mouth looking was also negatively correlated with EF subscales (inhibit p = .098, initiate p = .099, monitor p = .059), but these effects were marginal. In the TD group, no significant correlations were found for changes in number of fixations to the eye region from the doing to feeling condition (presumably a ceiling effect); however time spent attending to the mouth region decreased in the emotion-laden condition; a shift that was positively correlated with age (p = .04), verbal IQ (p = .05), matrix reasoning (p = .03), and general IQ (p = .01). These findings are represented in Figures 10 and 11.

Chapter 4: Discussion

Atypical looking patterns are often observed in children and adults with ASD and previous research has found mixed results regarding the visual fixation patterns of children with ASD with regard to both static and dynamic stimuli. Few studies have examined the effect of dynamic stimuli in the form of conversation on visual fixation patterns, and even fewer have explored different conversational contexts on these fixation strategies. The present study was designed to address this gap in the literature by analyzing the visual allocation strategies of TD children and children with ASD when engaged in conversations about mundane and emotion-laden topics.

4.1 Hypothesis 1

Recall that hypothesis 1 was: Children with ASD will attend less to the eye region and more to the mouth region of the face compared to TD children in both the mundane and emotion-laden conversational contexts. With regard to looking at the speaker's eyes and mouth, the visual allocation strategies of participants with ASD did not differ from those of their TD peers during both conversational contexts. This provides support for previous research that failed to find differences in the visual allocation patterns of those with ASD compared to TD and argues that visual attention in TD and ASD is more similar than dissimilar (Falkmer et al., 2010; Snow et al., 2011). On the other hand, children with ASD in this study also tended to look away from the speaker (in our case, "off-screen") across both conversational topics compared to their TD peers. This finding provides evidence that visual attention to a speaking partner is not equivalent across groups, and is consistent with evidence for inattention and lack of eye contact in ASD that are well documented (Charwarska et al., 2012; Joseph & Tanaka, 2003; Klin et al., 2002). These findings remind us that claims regarding the similarity and dissimilarity of visual attention in ASD and TD must be tethered to specific research questions and methods; findings that might first seem contradictory are not when they are understood in light of their specific operationalizations.

A main effect of group for average fixation duration to eyes and mouth revealed that participants with ASD tended to fixate longer on the speaker's eyes and mouth compared to their TD peers. This suggests more effortful and/or less efficient information processing presumably due to higher cognitive load in ASD across conditions. With regard to the number of fixations and total fixation time to eye and mouth regions, the descriptive data also reveal a portrait of more evenly distributed visual search and allocation to scenes with faces in ASD. Whereas the TD group seemed more selective and biased in visual search and information processing, the allocation of visual attention in ASD appears less strategic which may reflect underlying challenges in face processing, emotion reading, and ultimately, ToM.

4.2 Hypothesis 2

Hypothesis 2 predicted that the emotion-laden conversation (i.e., feeling condition) would be associated with more atypical visual allocation strategies in participants with ASD compared to the mundane conversation (i.e., doing condition). In support of this hypothesis, we found a group by condition interaction for two dependent variables: the number of fixations to eyes and the cumulative time spent looking to the mouth region. Each effect will be examined in turn.

Conversational topic moderates number of fixations to eyes. Participants with ASD showed significantly fewer fixations to the eyes in the emotion-laden (but not mundane) conversation compared to their TD peers. This is consistent with the findings from Baron-Cohen et al. (1997), who noted that as the complexity of emotions increased, persons with ASD demonstrated a decreased ability to decode information from the eyes. Of course, this interpretation raises a concern for children with ASD because eye information may be more relevant in conversations about emotions compared to mundane talk. It is precisely when these opportunities occur for learning the links between facial expression and underlying mental states (as well as the pragmatics of display rules) that children with ASD tend to neglect the regions of the visual field that help them decode the subtle social cues given in the eyes.

Conversational topic moderates looking time to mouth. Additionally, participants with ASD spent significantly longer fixating on the speaker's mouth in the emotion-laden (but not mundane) conversation compared to TD controls. One explanation for increased mouth looking has been suggested by Klin et al. (2002). They propose that children with ASD tend to look at the mouth more than TD peers because the mouth is moving during speech production and "by concentrating their efforts on something they understand, they might attain better understanding of social situations" (p. 814). Similar to the socially charged videos in Klin's study, our findings revealed that children with ASD spend more time looking at the mouth during emotional conversations, which are more socially charged than talk about mundane topics. While Klin's study focused solely on socially charged content, our study addressed both mundane and emotion-laden contexts.

Conversational topic moderates visual attention to faces. Taken together, our results for hypothesis 2 reveal that conversational context moderates the relationship between developmental status (ASD, TD) and visual allocation to the face of a speaking partner. These findings suggest that as the conversational task becomes more difficult, children with ASD shift visual attention from the eyes to the mouth. The emotion-laden conversation is likely more difficult because of the demands it places on ToM: a characteristic impairment of ASD. As a result, visual attention may be shifted to compensate for workload and conserve cognitive resources. Our findings align with those of Rutherford and Towns (2008), who found that when participants with ASD were asked to identify emotions, the complexity of the emotions moderated their visual attention to the person's face; as the complexity of the emotions increased, the visual allocation strategies shifted from the eye to the mouth area.

4.3 Research Question 1

Research question 1 was: "What factors predict visual allocation during conversation?" In our study, the TD sample shifted from fewer to more eye fixations and more to less mouth time in the emotion-laden conversation. Greater degrees of shifting from the mouth to the eyes were positively related to age, verbal ability, general IQ, and matrix reasoning in the TD sample. This suggests highly specialized and context-sensitive face viewing strategies with a critical developmental component. By contrast, the ASD sample shifted from more to fewer eye fixations and less to more mouth time in the emotion-laden conversation. The shift in TD children *toward* the eyes in the emotion condition appears to be driven by healthy developmental processes, but the shift in ASD *away* from the eyes and toward the mouth requires further scrutiny. When we solve problems, our prior knowledge usually helps us by efficiently guiding us toward solutions that have worked for us in the past. This seems to be at work in the TD sample, but it may or may not be at work in the ASD sample and it could be that in the mind of a person with ASD, the mouth (*not* the eyes) provides the most useful social information. However, one piece of evidence that argues against this possibility is that for the ASD group, decreasing (not increasing) age predicted more shift away from the eyes. In other words, as children with ASD in our study got older, they tended to *not* shift attention away from the eyes during the emotion-laden conversation to the degree that the younger children did. This could be a result of more social experience or general learning and maturation.

A weaker, more plausible form of this hypothesis might then be that under conditions of high cognitive load, the mouth provides the most salient or easily detectable face information for monitoring engagement and responses of a speaking partner. From this perspective, mouths could be salient for a variety of reasons. Perhaps they are salient due to their size and/or the fact that they move and produce audio-visual synchrony, which may aid in the disambiguation of speech sounds (Buchan et al., 2008; Klin et al., 2002, Jones & Klin, 2008). This would be consistent with a model wherein visual attention to low-level (largely bottom-up) features of visual input may compensate for limited cognitive resources due to load demands. Notably, Rutherford and Towns (2008) found similar results and concluded that it is “as if the more complex the target, the simpler the source of information must be” (p. 1379).

This simple principle of resource allocation is often illustrated by way of the ‘driving in a snow storm’ example. Driving in good conditions on a familiar route is not, relatively speaking, resource demanding: the act of driving almost becomes automatized and frees memory, attention, and planning cognitive reserves. By contrast, driving an unfamiliar route during a snowstorm is effortful and strains attention resources. Perhaps the emotion-laden condition caused participants with ASD to shift their visual attention in a way that the mundane condition did not. For individuals with ASD, engaging in emotion-laden conversations may be like driving in a snowstorm. This interpretation is consistent with the ToM hypothesis of ASD literature and aligns with previous research suggesting that social impairment and atypical visual attention are positively correlated in ASD (Elsabbagh et al., 2014; Speer, Cook, McMahon, & Clark, 2007).

Of course, when considering ToM effects in ASD, the role of EF cannot be overlooked. As previously described, Zelazo et al. (2013) have argued that hot EF and ToM are essentially the same construct and there is physiological evidence to support this assertion (e.g., Ashwin et al., 2007; Baron-Cohen et al., 2000). In the context of our interaction effects, one interpretation is that our two conversational contexts may differentially recruit hot and cold EF.

As indicated earlier in the Introduction section of this paper, how EF and its relation to ToM are constructed will shape one's interpretation of the results. Two sub-processes may be particularly relevant for this discussion: working memory and inhibition (see Moses et al., 2005/2013; Welsh, 2002/2008). Working memory and inhibition are often construed as independent, yet interrelated, processes and they may be important for understanding the attention shifting in the emotion-laden conversation that was observed for the ASD group in this study (see Figure 12). In our emotion-laden conversation, it may be that working memory contributes significantly to data for number of fixations (i.e., our eye-looking data). This may reflect a change in visual search strategy and implicates poor EF as a potential cause of the decreased number of fixations to the eye region. Perhaps children with ASD stop scanning the eye region because those stimuli are complex and difficult to read. That is, the cognitively demanding load that emotionally charged topics place on working memory may make rendering information from the eye region particularly difficult to interpret and the child with ASD may set off on a search for more easily accessible social information.

It may also be that inhibition contributes significantly to data for looking time (i.e., our mouth-looking data). In the emotion-laden condition, poor EF may primarily reflect deficits in inhibition, which may lead to increased mouth looking. More mouth looking may, in turn, implicate an impaired top-down attentional process (related to deficits in ToM) and intact bottom-up processing in ASD, which is consistent with some previous research (Klin et al., 2002; Neumann, Spezio, Piven, & Adolphs, 2006).

Indeed, in this study, different kinds of visual attention to different regions of a speaking partner’s face appear to involve distinct cognitive mechanisms. In the ASD group, greater degrees of shifting away from the eye region were negatively correlated with age, perceptual reasoning, and EF subscales: working memory, planning/organizing, and organization of materials, and positively correlated with autism severity. Furthermore, greater degrees of shifting toward the mouth region were negatively correlated with verbal IQ, general IQ, ToM competence, EF subscales: inhibit, initiate, and monitor, and positively correlated with autism severity. Our findings implicated different facets of visual attention (i.e., fixations to eyes and time to mouth). This is consistent with previous arguments that fixation and looking time are under the control of different cognitive operations (Mills et al., 2011). The fact that eye fixation shifting and mouth time shifting were predicted by different sets of variables in the TD and ASD samples further supports this assertion. As described above, one possibility is that eye fixation shifting is related to visual search (Mills et al., 2011; Theeuwes, 2012) and mouth time shifting is predicted by cognitive load (Lavie, Hirst, de Fockert, & Viding, 2004; Pollatsek et al., 1986).

4.4 Clinical Implications

Results of this study suggest that it is important to consider the stimuli and context in which intervention takes place. As previously discussed, there is mixed evidence for abnormal looking patterns when viewing static stimuli. Eye tracking technology reveals that there is a general consensus for abnormal visual allocation strategies during dynamic stimuli studies where participants with ASD are asked to view stimuli analogous to real-world settings. Since the present study was similar to a true conversational exchange, it is likely that when conversational exchanges are cognitively demanding and require the use of ToM, individuals with ASD allocate their visual attention to the speaking partner's mouth. For this reason, clinicians should carefully consider whether and when establishing sustained eye contact is a meaningful, important, or appropriate target of intervention. Clinicians should also note that when engaged in conversations that are cognitively demanding and require the use of ToM skills, eye contact may be especially rare in children with ASD. However, making eye contact is not a prerequisite for sharing attention. Interventionists might instead consider sharing attention to an object or activity by sitting beside a client with ASD to facilitate joint attention and shared meaning making.

Instructing students with ASD to "look at me" during therapy and conversational exchanges has the potential to overload cognitive resources particular during tasks that are already challenging because of their social nature. When cognitive resources are depleted, persons with ASD may experience difficulty accessing and taking meaning from the speaker and the speaker's message. According to Temple Grandin (Edelson & Grandin, 1996), "…if you start forcing eye contact, you are going to send the nervous system into sensory overload. [Those with ASD] will then shut down, and nothing is going to get through to them. They are 'mono channel,' and they can use only one sense at a time." This account rings true to John-Paul Bovee who is an adult who was diagnosed with ASD as a toddler. He states that "just because I am not making eye contact with you does not mean that I am not listening to you or paying attention to you. I can concentrate better not having to keep eye contact at the same time. I tell people, 'you have a choice. Do you want a conversation or do you want eye contact? You will not get both unless I am comfortable with you and do not have to concentrate so much on the eye contact" (Bovee, 1999 as cited in Stuart, 2000). In an interview about ASD, Grandin describes how forced eye contact can affect those with ASD:

> Donna Williams (1994) explained that forced eye contact caused her brain to shut down. She states when people spoke to her, "their words become a mumble jumble, their voices a pattern of sounds" (Painter 1992). She can use only one sensory channel at a time. If Donna is listening to somebody talk, she is unable to perceive a cat jumping up on her lap. If she attends to the cat, then speech perception is blocked. She realized a black thing was on her lap, but she did not recognize it as a cat until she stopped listening to her friend talk.
>
> She explained that if she listens to the intonation of speech, she can't hear the words. Only one aspect of incoming input can be attended to at a time. If she is distracted by the visual input of somebody looking in her face, she can't hear them. Other people with Autism have explained that they had a difficult time determining that speech was used for communication. Kins, a man with Autism, further explained that if somebody looked him in the eye, "My mind went blank and thoughts stop; it was like a twilight state" (Grandin, 2000).

These personal accounts describe the cognitive difficulty surrounding making eye contact. Not only is making eye contact uncomfortable for many people with ASD (Eye Contact Question, 2010), it also places a high demand on their EF system which has important implications for practice.

One idea worthy of future research is that when eye contact is a goal for meeting the expectations of listeners, and to make eye contact seem more natural even when it is not, it may be worthwhile to train eye gaze to face regions near the eyes but not directly to the eyes themselves. This might have the effect of increasing the impression of socially appropriate regulation of eye gaze, while not increasing cognitive load. This suggestion comes from the anecdotal testimony of persons with ASD; however, as far as we are aware, no studies have been conducted on the effectiveness of this approach.

Further support could be provided to children with ASD to reduce cognitive demand during ToM-laden activities. This could be done, for example, through the use of visual supports and repetition. Indeed, this is exactly what several well-established interventions do (e.g., Social Stories). Thus, reducing the cognitive load that is required of a person with ASD may remove some of the barriers associated with ToM tasks and provide access to increased learning.

4.5 Limitations and Directions for Future Research

Interpretation of the data obtained in this study is limited by the relatively small sample size. The literature suggests that a sample size of 25 is sufficient to identify effects in visual attention in ASD. Although our sample size was sufficient to detect many between- and within-group effects at the .05 level, a small number of effects (that are important to our current interpretation of the results) were only detected when alpha was .10. This suggests that the present study was slightly underpowered.

The ASD and TD groups were not matched on IQ. However, the ASD sample had a mean IQ in the normal range and the TD group had a mean IQ that was nearly one standard deviation above the mean. While matching the groups on IQ would result in more similar groups, it also has the potential to obscure effects associated with the diagnosis of ASD (Harms et al., 2010). In our study, analyses revealed that when the effects of verbal mental age were removed, no significant group by condition interactions persisted in the omnibus analyses. Therefore, it is difficult to control for IQ and verbal mental age and still find group differences, as IQ is "phenotypically linked with ASD" (Harms et al., 2010, p. 292).

Additionally, our data were only collected at one point in time. To address the limitations associated with cross-sectional design, future research should follow visual attention patterns over time using a longitudinal design. This would allow for examination of trends in visual attention over time, which may suggest a critical role of maturation.

Another limitation is that some of our participants had concomitant disorders. Specifically, parental report indicated that six participants with ASD also presented with ADHD or ADD. Despite any concomitant disorder, these participants were included in the study due to the small sample size. It should be noted that the present study was unable to assess the impact, if any, that the concomitant disorders or medications may have had on the study's results. ADHD and ADD have a particularly high comorbidity rate with ASD, which could be addressed through the use of additional screening measures or a more stringent set of inclusion criteria. Additionally, one participant with

ASD was functionally non-verbal, while all other participants were verbal and able to flexibly use language. This participant's data were analyzed with the others, as his verbal ability did not interfere with the eye tracking technology.

Additionally, making eye contact and looking other's in the eye is often reported to be difficult for those with ASD. This study did not determine whether children had received prior intervention on making eye contact and thus it is unknown whether any interventions had an effect on the present study.

Future research might also examine cognitive load across conditions by examining average pupil dilation. Research shows that pupil size tends to be positively associated with cognitive load (Hyönä, Tommola, & Alaja, 1995). Further, future research could utilize a dual task approach for TD children that measures eye looking versus mouth looking in noisy environments. This could assess whether TD individuals' visual allocation is similar to that of those with ASD when cognitive load has been increased. Buchan et al. (2008) did find evidence for increased mouth looking in TD children during a speech intelligibility task with extraneous background noise; however future research could compare similar findings with the increased mouth looking found in ASD during cognitively demanding tasks.

The present study may help to explain the conflicting results in eye tracking and ASD literature. Here we have identified several factors that influence the visual attention to eye and mouth regions of the face, including age, autism severity, perceptual reasoning, verbal and general IQ, and conversational context (and by extension, level of ToM and EF).

Chapter 5: References

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience, 13*(2), 232-240.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., & Bullmore, E. T. (2007). Differential activation of the amygdale and the 'social brain' during fearful face-processing in Asperger Syndrome. *Journal of Neurophysiology, 45*, 2–14.

Ashwin, C., Wheelwright, S., & Baron-Cohen, S. (2006). Finding a face in the crowd: Testing the anger superiority effect in autism. *Brain and Cognition, 61*, 78-95.

Baron-Cohen, S. (1992). The theory of mind hypothesis of autism: History and prospects of the idea. *The Psychologist, 5*, 9-12.

Baron-Cohen, S. (1995/2001). *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA: The MIT Press.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37-46.

Baron-Cohen, S., Ring, H., Bullmore, E., Wheelwright, S., Ashwin, C., & Williams, S. (2000). The amygdale theory of autism. *Neuroscience and Biobehavioral Reviews, 24*, 355-364.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4(3)*, 311-331.

Barral, O. (2013). Exploring the applicability of implicit relevance measures in varying reading speed for adaptive I.R. systems. Master's thesis, Universitat Politècnica de Catalunya/Helsignin Yliopsito, Spain/Finland.

Bird, G., Press, C., & Richardson, D.C. (2011). The role of alexithymia in reduced eye-fixation in autism spectrum conditions *Journal of Autism Developmental Disorders, 41,* 1556-1564.

Bowler, D. (1992). "Theory of mind" in Asperger's Syndrome. *The Journal of Child Psychology and Psychiatry, 33(5),* 877-893.

Buchan, J. N., Pare, M., & Munhall, K. G. (2008). The effect of varying talker identity and listening conditions on gaze behavior during audiovisual speech perception. *Brain Research*, *1242*, 162-171.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305-327.

Castellanos, X.F., Sonuga-Barke, E.J.S., Milham, M.P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *TRENDS in Cognitive Sciences, 10 (3),* 117-123.

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *The Journal of Child Psychology and Psychiatry, 58(3)*, 903-913.

Dawson, G., Webb, S., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology, 27(3),* 403-424.

Doherty-Sneddon, G., Riby, D., & Whittle, L. (2012). Gaze aversion as a cognitive load management strategy in autism spectrum disorder and Williams syndrome. *The Journal of Child Psychology and Psychiatry, 53(4)I,* 420-430.

Edelson, S. (Interviewer) & Grandin, T. (1996). *Interview with Dr. Temple Grandin* [Interview transcript]. Retrieved from Autism Research Institute Web site: http://www.autism.com/advocacy_grandin_interview

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., & Johnson, M. (2014). *What you see is what you get:* Contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience, 9(4)*, 538-543.

Eye Contact Question. (2010). Retrieved from http://www.wrongplanet.net/forums/viewtopic.php?t=141071http://www.wrongplanet.net/forums/viewtopic.php?t=141071

Falkmer, M., Larsson, M., Bjällmark, A., & Falkmer, T. (2010). The importance of the eye area in face identification abilities and visual search strategies in persons with Asperger syndrome. *Research in Autism Spectrum Disorders*, *4 (4)*, *724-730*.

Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? Evidence from eye tracking. *Journal of Autism and Developmental Disabilities, 40*, 534-547.

Gilliam, J. (2006). *Gilliam Autism Rating Scale manual* (2nd ed.). Upper Saddle River: Pearson Education.

Gioia, G., Isquith, P., Guy, S., & Kenworthy, L. (2000). *Behavior Rating of Executive Function.* Lutz, FL: Psychological Assessment Resources.

Grandin, T. (2000). *Visual Thinking, Sensory Problems & Communication Difficulties.* Retrieved from http://www.autism-help.org/story-sensory-communication.htm

Grelotti, D., Gauthier, I., & Schultz, R. (2002). Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Developmental Psychobiology, 40*, 213-225.

Grelotti, D., Klin, A., Gauthier, I., Skudlarski, P., Cohen, D., Gore, J., Volkmar, F., Schultz, R. (2005). fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not t faces in a boy with autism. *Neuropsychologia, 43,* 373-385.

Grice, S., Halit, H., Farroni, T., Baron-Cohen, S., Bolton, P., & Johnson, M. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex, 41(3),* 342-353.

Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition, 48,* 101-119.

Happé, F., Roland, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9(10),* 1218-1220.

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology review*, *20*(3), 290-322.

Hernandez, N., Metzger, A., Magne, R., Bonnet-Brilhault, F., Roux, S., Barthelemy, C., & Martineau, J. (2009). Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning. *Neuropsychologia, 47*, 1004–12.

Hewes, D. (2013). Methods as tools: A response to O'Keefe*. Human Communication*

Research, 29(3), 448-454.

Hill, E. L. (2008). Executive functioning in autism spectrum disorder: Where it fits in the causal model. In E. McGregor, M. Nunez, K. Cebula, & J.C Gomez (Eds.), *Autism: An integrated view from neurocognitive, clinical, and intervention research* (pp. 145-165). Malden, MA: Blackwell Publishing.

Hughes, C. (1998). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology, 34(6),* 1326-1339.

Hutchins, T. L., & Prelock, P. A. (in press). Beyond the theory of mind hypothesis: Using a causal model to understand the nature and treatment of multiple deficits in autism spectrum disorder. *Handbook of communication sciences and disorders.*

Hutchins, T., Prelock, P., & Chace, W. (2008). Test-retest reliability of a theory of mind task battery for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 23(4), 195-206.

Jemel, B., Mottron, L., & Dawson, M. (2006). Impaired face processing in autism: Fact or artifact? *Journal of Autism and Developmental Disorders, 36(1),* 91-106.

Jones, W., & Klin, A. (2008). Altered salience in autism: Developmental insights, consequences, and questions. In E. McGregor, M. Nunez, K. Cebula, & J.C Gomez (Eds.), *Autism: An integrated view from neurocognitive, clinical, and intervention research* (pp. 62-82). Malden, MA: Blackwell Publishing.

Joseph, R., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *The Journal of Child Psychology and Psychiatry, 44(4)*, 529-542.

Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). "Defining and quantifying the social phenotype in autism": Dr. Klin and colleagues reply. *American Journal of*

Psychiatry, 160(7), 1359-1360.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*, 809–16.

Langdell, T. (1978). Recognition of faces: An approach to the study of autism. *The Journal of Child Psychology and Psychiatry, 19*, 255-268.

Lavie, N., Hirst, A., de Fockert, J.W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133(*3), 339-354.

Lindskog, C., & Smith, J. (2010). [Review of the Weschler Abbreviated Scale of Intelligence]. In *The fourteenth mental measurements yearbook.* Available from http://www.unl.edu/buros/.

Lopez-Waner, M., Hoffman, C., Sweeney, D., & Hodge, D. (2008). Sleep problems of parents of typically developing children and parents of children with autism. *The Journal of Genetic Psychology*, 169(3), 245-259.

Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., & Di Lavore, P.C. (2000). The Autism Diagnostic Observation Schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *The Journal of Autism and Developmental Disorders, 30 (3)*, 205-223.

Maekawa, T., Tobimatsu, S., Inada, N., Oribe, N., Onitsuka, T., Kanba, T., & Kamio, Y. (2011). Top-down and bottom-up visual information processing or non-social stimuli in high-functioning autism spectrum disorder. *Research in Autism*

Spectrum Disorders, 5, 201-209.

McPartland, J., Webb, S., Keehn, B. & Dawson, G. (2011). Patterns of visual attention to faces and objects in autism spectrum disorder. *Journal of Autism and Developmental Disabilities, 41,* 148-157.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11(8),* 1-15.

Moses, L.J., Carlson, S.M., & Sabbagh, M.A. (2005/2013). On the specificity of the relation between executive function and children's theories of mind. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp.131-145). New York, NY: Psychology Press.

Mundy, P., Sigmna, M., & Kasari, C. (1996). The theory of mind and joint-attention deficits in autism. In E. Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D.J. (Eds.), *Understanding other minds: Perspectives from autism* (pp. 181-203). Oxford: Oxford Medical Publications.

Neumann, D., Spezio, M., Piven, J., & Adolphs, R. (2006) Looking you in the mouth: Abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience, 1,* 194-202.

Norbury, C., Brock, J., Cragg, L., Einav, S., Griffiths, H, & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *The Journal of Child Psychology and Psychiatry, 50(7),* 834-

842.

Ozonoff, S., Pennington, B., & Rogers, S. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *The Journal of Child Psychology and Psychiatry, 32(7),* 1081-1105.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43(4)*, 974-990.

Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology, 46(2),* 530-544.

Pelphrey, K., Morris, J., McCarthy, G., & Labar, K. (2007). Perception of dynamic changes in facial affect and identity in autism. *Social Cognitive and Affective Neuroscience, 2(2)*, 140-149.

Pollatsek, A., Rayner, D., & Bolata, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Percept Psychophysiology, 40,* 123-130.

Riby, D.M., & Doherty, M.J. (2009). Tracking eye movements proves informative for the study of gaze direction detection in autism. *Research in Autism Spectrum Disorders, 3,* 723-733.

Riby, D.M., & Hancock, P.J. (2008). Viewing it differently: Social scene perception in Williams syndrome and autism. *Neuropsychologia, 46*, 2855–2860.

Riby, D., & Hancock, P.J. (2009). Looking at movies and cartoons: Eye tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability*

Research, 53(2), 169-181.

Rutherford, M.D., & Towns, A.M. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38,* 1371-1381.

Schultz, R.T. (2005). Developmental deficits in social perception in autism: The role of the amygdale and fusiform face area. *International Journal of Developmental Neuroscience, 23*, 125-141.

Schultz, R., Gauthier, I., Klin, A., Fulbright, R., Anderson, A., Volkmar, F., Skudlarski, P., Lacadie, C., Cohen, D., & Gore, J. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archives of General Psychiatry, 57(4),* 331-340.

Schraw, G. (2003). Review of the Behavior Rating Inventory of Executive Function. In *The fifteenth mental measurements yearbook.* Available from http://www.unl.edu/buros/.

Shic, F., Bradshaw, J., Klin, A., Scassellati, B., & Chawarska, K. (2011). Limited activity monitoring in toddlers with autism spectrum disorder. *Brain Research, 1380,* 246-254.

Snow, J., Ingeholm, J., Levy, I., Caravella, R., Case, L., Wallace, G., & Martin, A. (2011). Impaired visual scanning and memory for faces in high-functioning autism spectrum disorders: It's not just the eyes. *Journal of the International Neuropsychological Society, 17,* 1021-1029.

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism, 11*(3), 265-

277.

Spezio, M., Huang, P-O., Castelli, F., & Adolphs, R. (2007). Amygdale damage impairs eye contact during conversations with real people. *The Journal of Neuroscience, 27(25)*, 3994-3997.

Stuart, R. (2000). Should we insist on eye contact with people who have autism spectrum disorders. *The Reporter, 5(3)*, 7-12. Retrieved from http://www.iidc.indiana.edu/?pageId=472

Theeuwes, J. (2012). Automatic control of visual selection. In M.D. Dodd, & J. Flowers, (Eds.), *The influence of attention, learning, and motivation on visual search* (pp. 23-62). New York, NY: Springer.

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders, 3(2),* 556-569.

Wechsler, D. (1999). *Wechsler Abbreviated Scales of Intelligence* (WASI). TX: Psychological Corporation: San Antonio.

Welsh, M.C., (2002/2008). Developmental and clinical variations in executive functions. In Molfese, D.L. & Molfese, V.J. (Eds.), *Developmental variations in learning: Applications to social, executive function, language, and reading skills* (pp.139-185). New York, NY: Psychology Press.

Zelazo, P., Qu, L., & Müller, U. (2013). Hot and cool aspects of executive function: Relations in early development. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp.71-93). New York, NY: Psychology Press.

Appendix 1: Skype Conversation Questions

Hi ________. So I want to have a conversation with you. First, I want to ask you about what people do. Can you tell me…

What kinds of things do people do people do for work?
What kinds of things do people do at home?
What kinds of things do people eat?
What kinds of things do people do for fun?

Now I'd like to ask you questions about what YOU do.

Ya know, I have lots of things that I do in the morning. What do you do when you get up in the morning?
And sometimes I have things I like to do with my friends. Tell me, what do you do at school at recess time?
And sometimes I have a dinner routine. What do you do at dinner time?
And you know what else? I usually have a night time routine. Tell me, what do you do before you go to bed at night?

Now I'd like to ask you about feelings. Tell me…

What kinds of things make people scared?
What kinds of things make people happy?
What kinds of things make people sad?
What kinds of things make people mad?

Now I'd like to talk about YOUR feelings.

Ya know, sometimes I get scared…like when my dog got lost and I couldn't find her and I didn't know if she was ok. I was scared. What kinds of things scare you?
And sometimes I get mad like when someone cuts in front of me when I am in line. What kinds of things make you mad?
And sometimes I feel sad like when I can't spend time with a friend because I am too busy. What kinds of things make you sad?
And ya know what I like to do? I like to watch my favorite movies. These things make me happy. What kinds of things make you happy?

Figure 1: AOIs for Skype Conversation

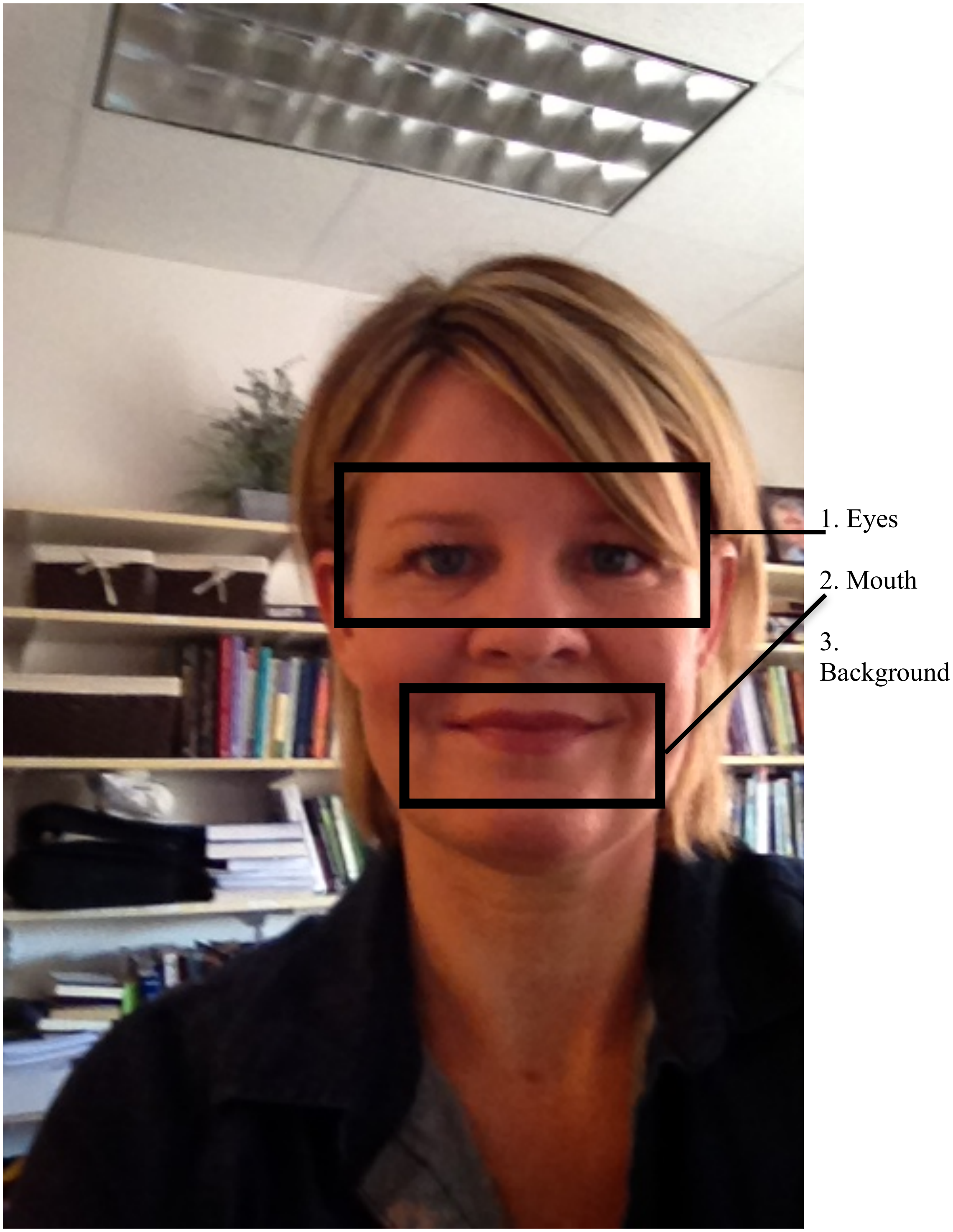
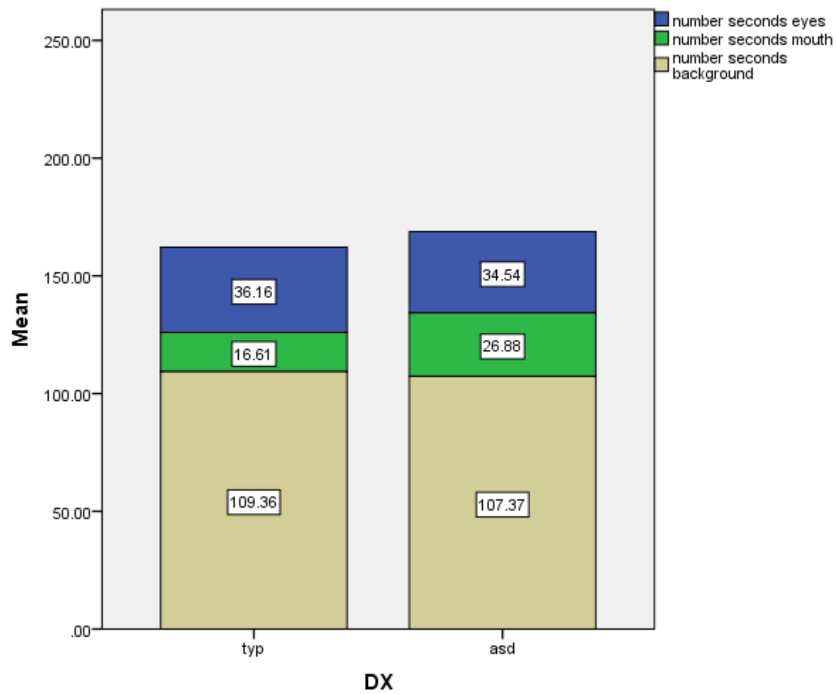


Figure 2: Mean Looking Time (seconds). Vertical axis represents mean number of seconds

Doing Condition



Feeling Condition

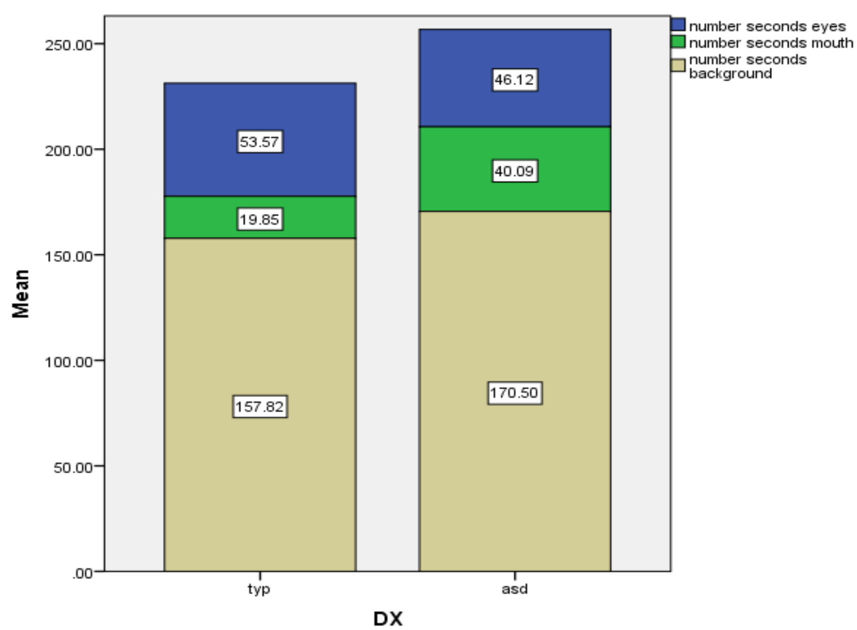
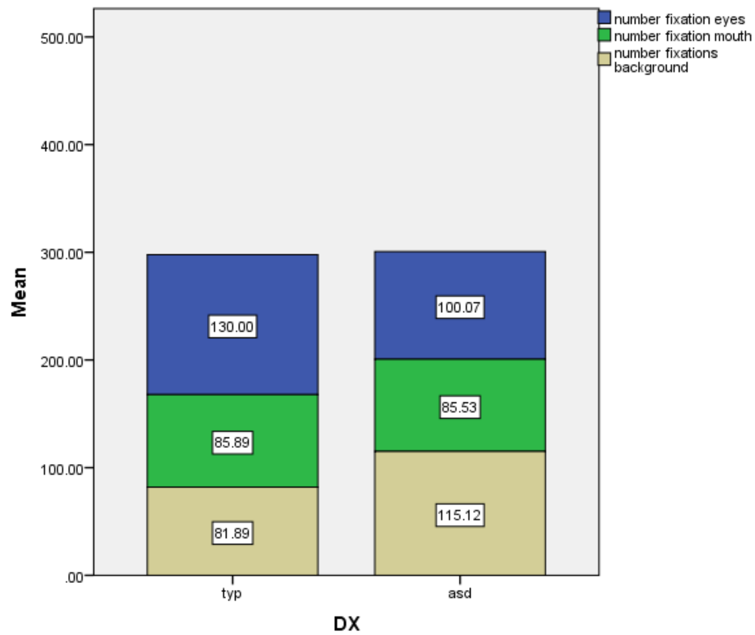


Figure 3: Mean Number of Fixations. Vertical axis represents mean number of fixations

Doing Condition



Feeling Condition

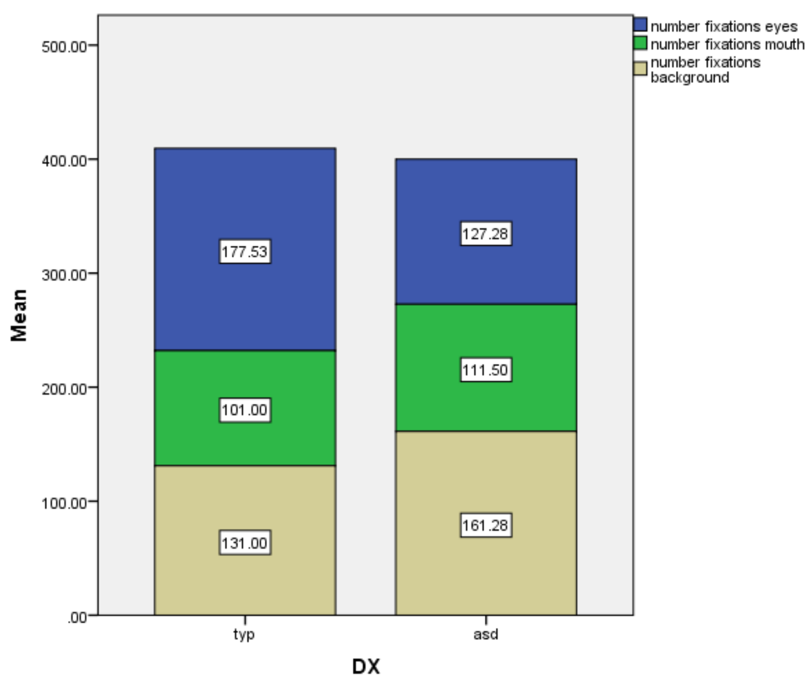


Figure 4: Main Effect for Group for Total Off-screen Looking Time (seconds). Error bars represent the standard error

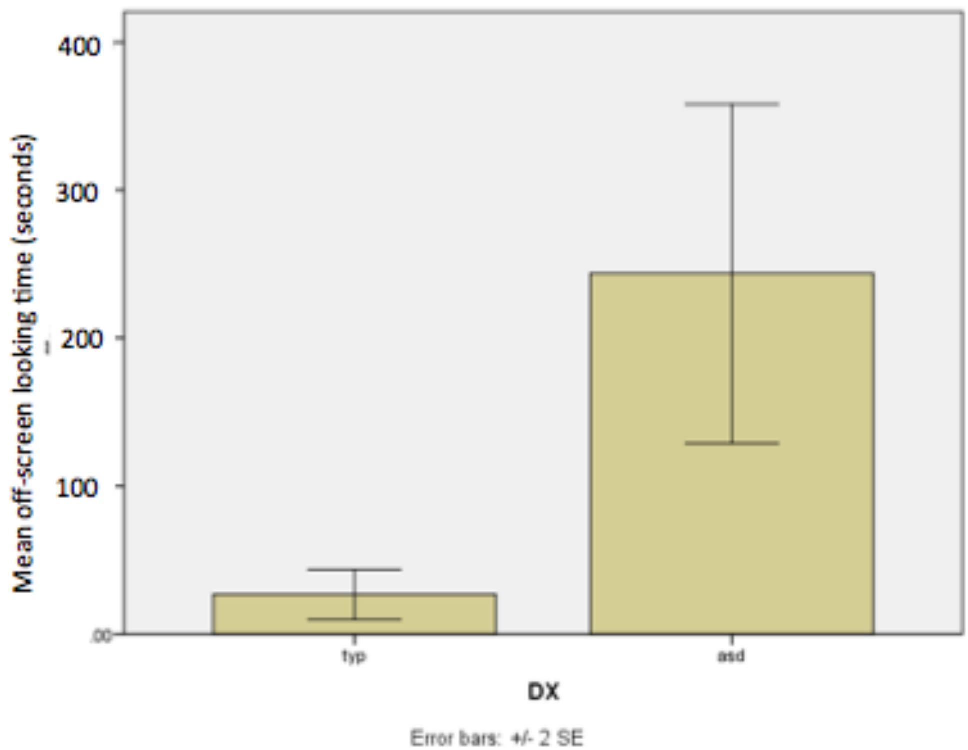


Figure 5: Average Fixation Duration in the Doing Condition. Vertical axis represents mean number of seconds

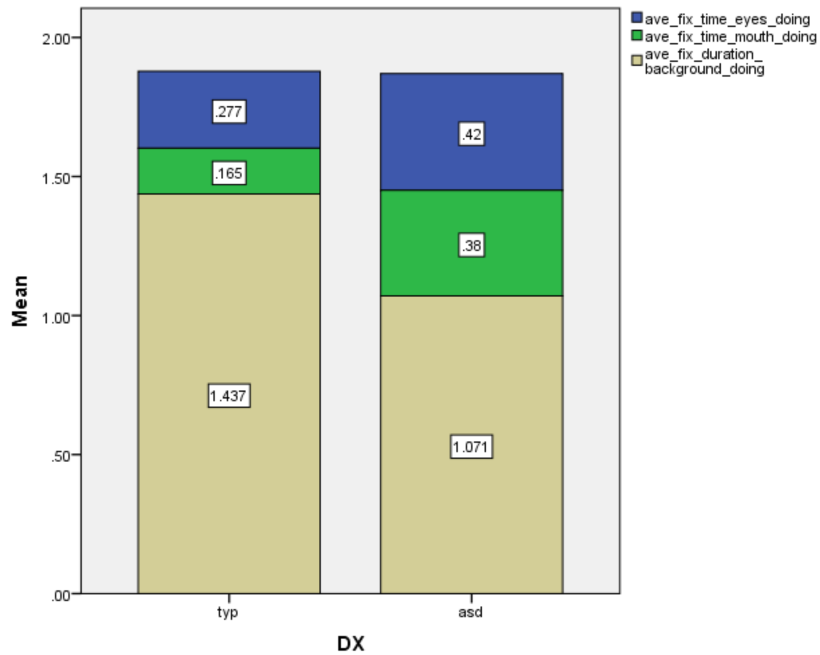


Figure 6: Average Fixation Duration in the Feeling Condition. Vertical axis represents mean number of seconds

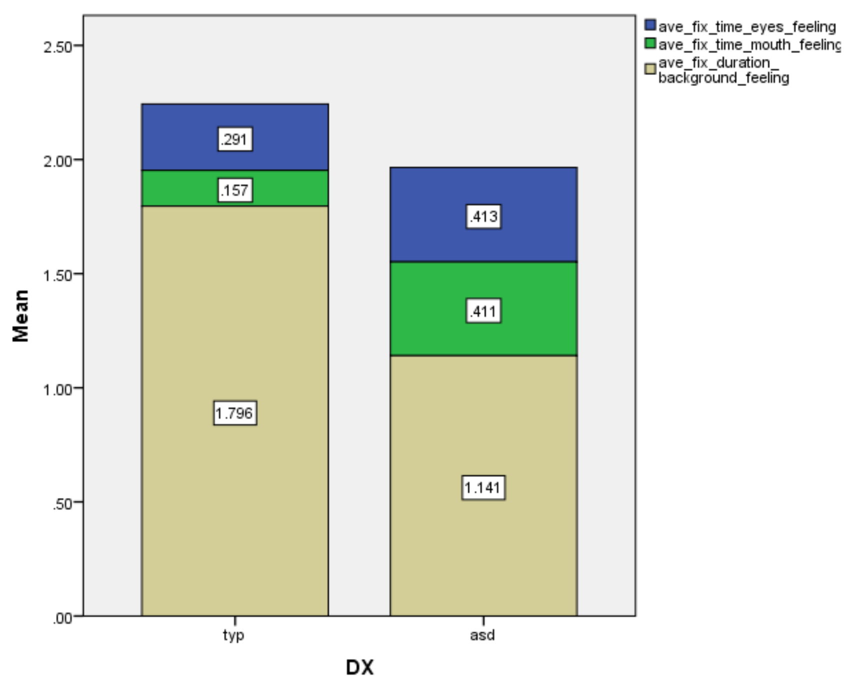


Figure 7: Total Time Looking at the Mouth Region of the Face (seconds)

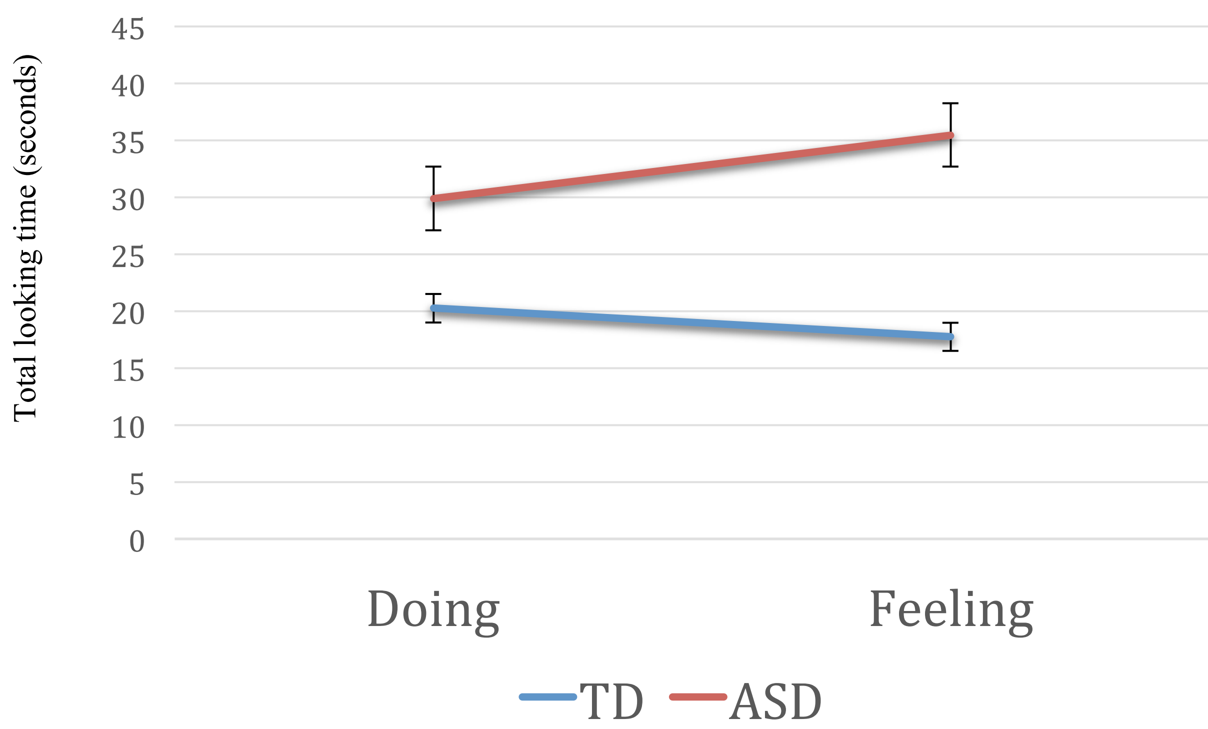


Figure 8: Looking Patterns of Participants with ASD in Feeling Condition Compared to TD peers

Figure 9: Number of Fixations to Eye Region of the Face

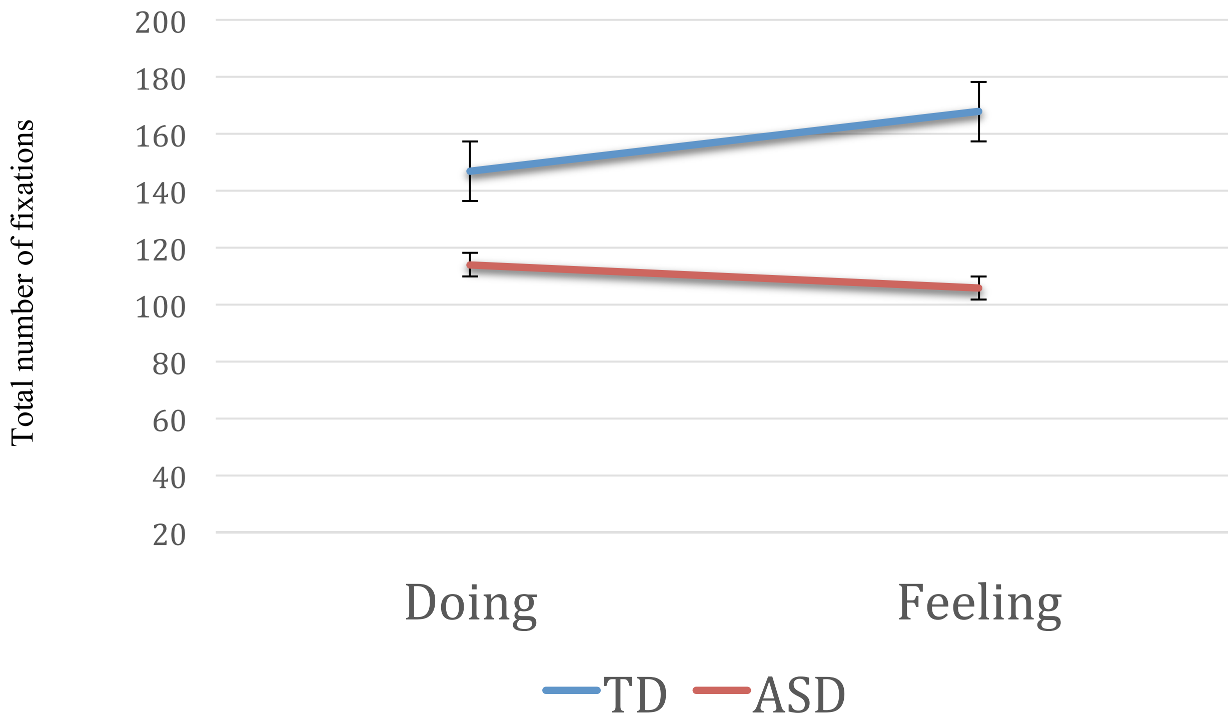


Figure 10: Change Score Correlations for Eye Number of Fixations

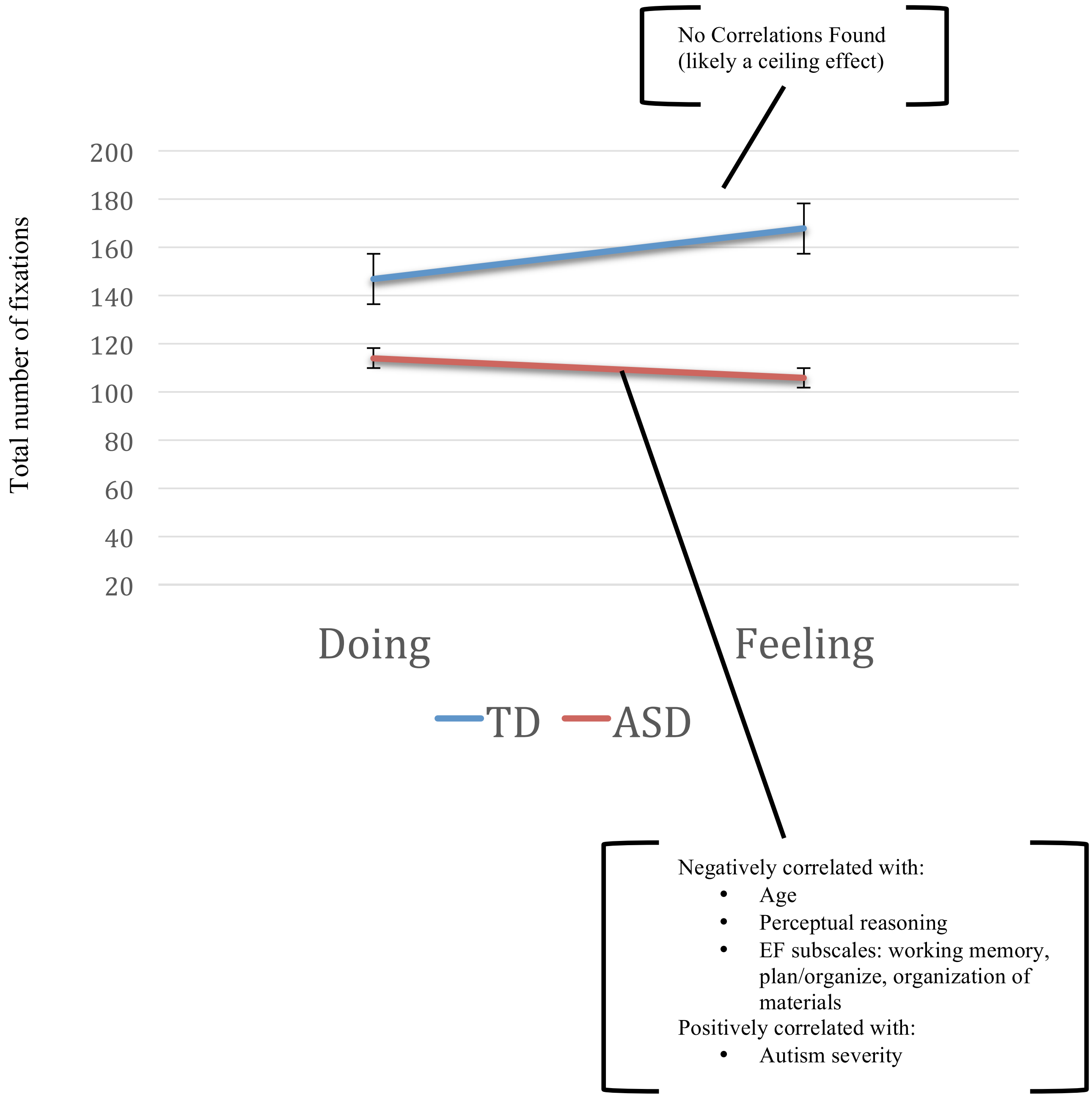


Figure 11: Change Score Correlations for Mouth Time

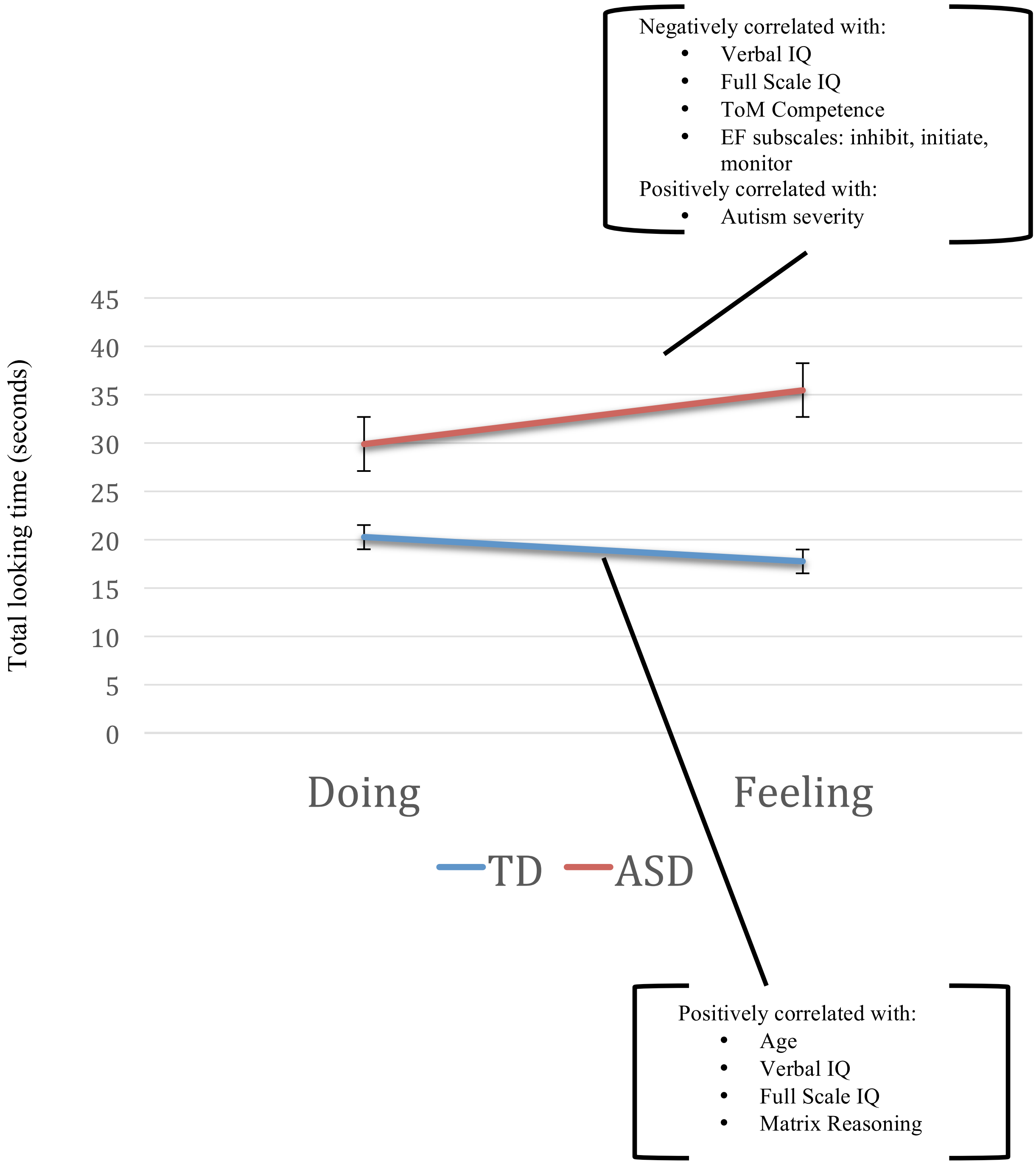
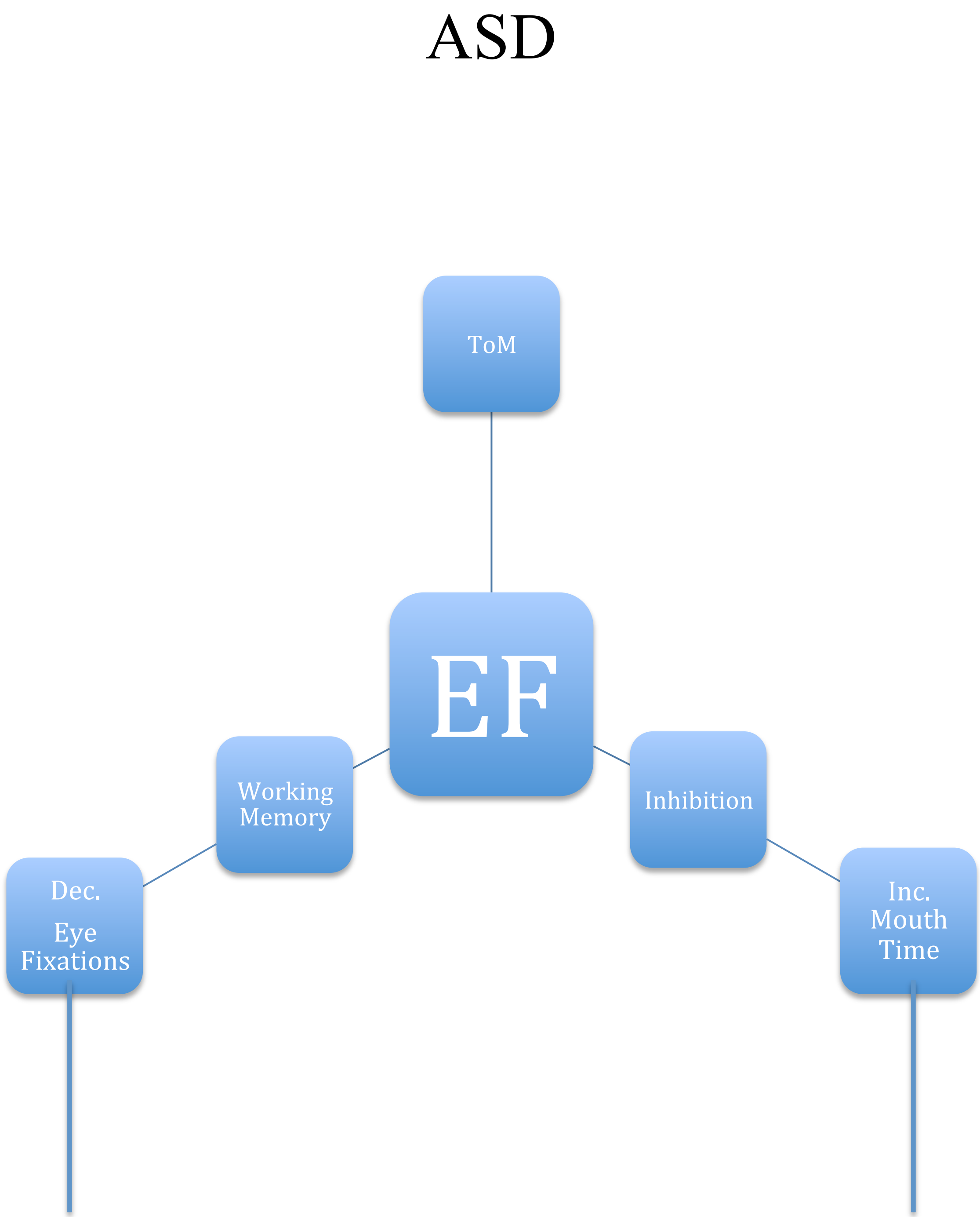


Figure 12: Cognitive Mechanisms Involved in Visual Attention in ASD



- Higher ASD Severity
- Younger
- Poorer Perceptual Reasoning
- Poorer EF subscales: WM, plan/organize, organization of materials

- Higher ASD severity
- Poorer Verbal IQ
- Poorer Full Scale IQ
- Poorer ToM Competence
- Poorer EF subscales: inhibit, initiate, and monitor