Temporal Synchrony in the Processing of Social Stimuli: Bimodal Integration in Autism Spectrum Disorder

Beatch, Jacqueline

http://hdl.handle.net/11023/1859

Downloaded from PRISM Repository, University of Calgary
UNIVERSITY OF CALGARY

Temporal Synchrony in the Processing of Social Stimuli: Bimodal Integration in Autism Spectrum Disorder

by

Jacqueline Ann Beatch

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA
SEPTEMBER, 2014

© Jacqueline Ann Beatch 2014
Abstract

The current study examined preferences for one’s own species (conspecifics) by presenting four-year-old typically-developing (TD) children and those diagnosed with Autism Spectrum Disorder (ASD) pictures and videos of human, monkey, and robot stimuli while examining the role temporal synchrony plays in audiovisual integration. Stimuli were presented on an eye-tracker then children’s verbal and non-verbal abilities were assessed. In total, TD children looked longer to faces but similar proportional looking between groups was found for static and dynamic stimuli. Preference for synchrony was found for TD children and reduced total looking across asynchronies by children with ASD may have inhibited their ability to notice asynchrony. Children’s verbal and non-verbal abilities were not related to looking behaviour during static and dynamic presentations. In general, atypical looking behaviour was found for children with ASD which may have cascading effects on their social and language development.
Acknowledgements

There are many people without whom this project would not have been possible. First and foremost I would like to thank my supervisor Dr. Suzanne Curtin for giving me an amazing opportunity to do what I love. Her unwavering support, guidance, and patience over the past few years has made this moment possible. My deepest thanks also go to my committee members Dr. Susan Graham and Dr. Suzanne Hala for all of their assistance and understanding throughout the thesis process.

Second I would like to thank everyone in the Speech Development Lab. The amount of hard work and dedication they put towards my project never ceased to amaze me. Thank you to Pat, Regina, Michelle, Lena, Will, and Kat for the countless hours you spent training and helping run participants and for your endless support. Special thanks go to Jennifer Ference and Lisa Henkel. Jennifer guided me not only throughout my thesis but helped me navigate life as a graduate student. Lisa was a continued source of support during moments of uncertainty and was always eager to help in anyway, with anything.

Lastly, I would like to thank my family and my husband. Their unconditional love and support throughout all of my studies has helped give me strength and confidence to pursue my passion. I would not be where I am today if not for their faith and support.
To Mom and Dad, for always believing in me
# Table of Contents

Abstract ........................................................................................................................... ii
Acknowledgements ........................................................................................................ iii
Dedication ........................................................................................................................ iv
Table of Contents ............................................................................................................. v
List of Tables ................................................................................................................... vi

**Introduction**.................................................................................................................. 1

1.1 Speech and Face Preferences ................................................................................... 2
1.2 Audiovisual Integration and Temporal Synchrony ...................................................... 5

**Present Study** ............................................................................................................. 10

**Task 1: Conspecific Preferences for Static Images** ...................................................... 13

3.1 Methods ...................................................................................................................... 14
3.1.1 Participants ........................................................................................................... 14
3.1.2 Materials ............................................................................................................. 15
3.1.3 Stimuli ................................................................................................................ 16
3.1.4 Procedure .......................................................................................................... 17

3.2 Results ....................................................................................................................... 18

3.3 Discussion ................................................................................................................ 20

**Task 2: Audiovisual and Temporal Integration** .......................................................... 22

4.1 Methods ...................................................................................................................... 22
4.1.1 Participants ........................................................................................................... 22
4.1.2 Materials ............................................................................................................. 23
4.1.3 Stimuli ................................................................................................................ 23
4.1.4 Procedure .......................................................................................................... 24

4.2 Results ....................................................................................................................... 24

4.3 Discussion ................................................................................................................ 28

**General Discussion** .................................................................................................. 30

5.1 Conclusion ............................................................................................................... 35
5.2 Limitations and Future Directions .......................................................................... 36

References ..................................................................................................................... 38

Appendix A: Visual Stimuli ........................................................................................... 46
Appendix B: Ethics Certification ..................................................................................... 47
List of Tables

Table 1: Demographic Information by Group for all Participants..........................16
Table 2: Mann-Whitney Test of Between Group Differences to Static Faces.............18
Table 3: Paired Sample t-test for TD Children’s Preference for Static Eye Region.......19
Table 4: Paired Sample t-test for Children with ASD’s Preference for Static Eye Region..19
Table 5: Mullen NVIQ and VIQ Scores for Static Stimuli between Groups..............20
Table 6: Mann-Whitney test of Between Group Differences to Dynamic Faces........24
Table 7: Paired Sample t-test for TD Children’s Preference for Dynamic eye Region....25
Table 8: Mann-Whitney Test for Preferences for Synchrony between Groups using
  Proportional Looking........................................................................................26
Table 9: Mann-Whitney Test for Preference for Synchrony between Groups using Total
  Looking Time........................................................................................................26
Table 10: Paired Sample t-test for Within Group Preference for Synchronous Stimuli....27
Table 11: Mullen NVIQ and VIQ Scores for Dynamic Stimuli between Groups.........28
Introduction

Autism Spectrum Disorder (ASD) is a broad diagnostic category which includes Autism disorder, Asperger’s disorder, and pervasive developmental disorder not otherwise specified (PDD-NOS; American Psychiatric Association, 1994). Recent research suggests that as many as 1 in 50 children in the United States will be diagnosed with ASD in a year (Blumberg et al., 2013). Rising rates have led to an increased importance of understanding the etiology, developmental course, and short and long-term outcomes for individuals with ASD.

Autism Spectrum Disorder is broadly characterized by social and communicative impairments (Volkmar, Chawarska, & Klin, 2005), and failure for children to acquire appropriate language by typical ages is one of the most frequently noticed symptoms by parents (Rapin, 1991). Language development is a complex process, making it difficult to determine the nature of the deficit or delay children with ASD experience. Deficits in communication abilities represent a core feature of ASD (APA, 1994) with many children having significant delays in their comprehension and production of language compared to typically developing (TD) peers and those with developmental delays (Weismer, Lord, & Esler, 2010). Heterogeneity in language skill development in children with ASD show some children performing at age-appropriate levels while others remain non-verbal into adulthood (Hudry et al., 2010; Watson et al., 2011). In general, TD children have greater receptive compared to expressive language abilities (Weismer et al., 2010). While some children with ASD display this typical profile, others show atypical patterns of greater expressive compared to receptive language abilities (Hudry et al., 2010; Weismer et al., 2010). Those exhibiting typical language profiles tend to have higher overall language functioning (Wynn & Smith, 2003) while those displaying atypical expressive-receptive profiles may rely on echolalia as a communication tool (Hudry et al., 2010). That is,
children with ASD with lower language functioning may use the rote repetition of expressing words to come to understand the word’s meaning based on the consistent responses from others (Hudry et al., 2010). These findings demonstrate the heterogeneity within this population and highlight the complexity in understanding language development and abilities in this group.

Recently, researchers have begun to explore the perceptual underpinnings of linguistic and social development in both typically and atypically developing populations. This research has found adaptive preferences exist in multiple domains, such as conspecifics, audiovisual integration, and temporal synchrony, which may aid in social and communicative development (Bebko, Weiss, Demark, & Gomez, 2006; Vouloumanos & Werker, 2007). Through the examination of a range of early perceptual skills such as, fundamental biases that direct attention to salient stimuli, bimodal processing abilities, and relationships between these skills, insight into atypical language and social development may emerge. These perceptual abilities will be examined in the current research in both TD children and those diagnosed with ASD.

1.1 Speech and Face Preferences

Fundamental biases that direct attention to socially relevant stimuli such as voices and faces are hallmarks of typical development (Johnson, Dziurawiec, Ellis, & Morton, 1991; Vouloumanos & Werker, 2004; 2007). From birth, TD newborns show preferences for human faces over other complex visual stimuli (Fantz, 1963) and human speech over non-speech analogues (Vouloumanos & Werker, 2007). Children with ASD, however, do not seem to demonstrate the same attentional biases for human faces or speech (Chawarska, Volkmar, & Klin, 2010; Kuhl, Coffey-Corina, Padden, & Dawson, 2005). Species-specific, or conspecific, preferences help infants focus on salient social information during complex events thereby assisting their social and communication development (Vouloumanos & Werker, 2007). By 3
months of age, TD infants prefer human faces and bodies to monkey faces and bodies (e.g. gorilla) when presented with a head only, body only, or both (i.e. a full length body; Heron-Delaney, Wirth & Pascalis, 2011). This suggests that by this age TD infants have a flexible mental representation of a human allowing them to recognize an incomplete picture (Heron-Delaney et al., 2011). Children with ASD, however, attend differently to social (human face) and non-social (mosaic) visual stimuli (Chawarska et al., 2010). Compared to TD children and those with developmental delays, children with ASD can more quickly disengage their attention from social stimuli to focus on a peripheral target, but do not differ in their looking behaviour on non-social trials (Chawarska et al., 2010). This suggests that children with ASD are less engaged by social compared to non-social stimuli (Chawarska et al., 2010). Similar patterns are found with auditory speech stimuli in that by 3 months, TD infants have narrowed their tuning of listening preferences and prefer human speech over rhesus monkey calls (Vouloumanos, Hauser, Werker, & Martin, 2010). Yet, children with ASD between the ages of 2 and 4 years show atypical auditory preferences for non-speech analogues, with matched acoustic characteristics of motherese, compared to motherese speech (Kuhl et al., 2005), suggesting that children with ASD seem to show a deficit, or delay, in preferring human speech. Lack of fundamental biases for conspecific information may have cascading effects on children’s social and language development (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). That is, for infants to learn how to properly interact within their social environment they first must learn to attend to salient information (Fantz, 1963; Watson et al., 2011).

In addition to overall looking and listening preferences, it is also informative to explore where on a visual stimulus participants tend to look and what factors might influence looking behaviour. For example, TD infant’s preferences for the eye or mouth regions are influenced by
their age and stage of language development. Infants under 6 months look longer to the eye region and between 6 and 12 months infants tend to look longer to the mouth region (Lewkowicz & Hansen-Tift, 2012) especially when speech is present (Tenenbaum, Shah, Sobel, Malle, & Morgan, 2012). Changes in looking behaviour around 6 months have been attributed to infant’s development of language or attentional skills (Lewkowicz & Hansen-Tift, 2012). Around 12 months infants seem to more flexibly shift between regions based on which region is providing salient social information (e.g., gaze cue information versus speech information; Tenenbaum et al., 2012), and by around 4 years of age children show increased scanning between the eye and mouth regions (Shic, Chawarska, Bradshow, & Scassellati, 2008). Increased exploration between regions may suggest increased attentional resources allowing children to attend to both social and linguistic information rather than simply focusing on one or the other type of information.

Attention to salient social and linguistic information may be difficult for individuals with ASD due to deficient, or delayed, attention to salient facial regions and changes in scanning patterns. Not only do children and adolescents with ASD focus more on external features, such as objects or the body (Chawarska & Shic, 2009; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Rice, Moriuchi, Jones, & Klin, 2012; Shic et al., 2008; Speer, Cook, McMahon, & Clark, 2007), their looking behaviour does not seem to develop to focus on salient areas of interest (AOI), such as the eye or mouth regions (Shic et al., 2008). Some research actually suggests that looking to salient facial regions declines over early childhood in children with ASD (Chawarska & Shic, 2009; Shic et al., 2008). This makes it less likely that children with ASD are delayed in changing their looking behaviour. Further, research studying adolescents with ASD also finds atypical fixation patterns possibly suggesting irregular looking behaviour is stable across childhood and adolescence (Klin et al., 2002; Speer et al., 2007).
Since differences in facial processing between TD individuals and those with ASD occur across childhood and adolescence, these differences may be important for predicting children who may develop a diagnosis of ASD. The strongest predictor of group membership (i.e. TD or diagnosed with ASD) in school age children is looking time to faces with children with ASD spending significantly less time looking at the face in general (Rice et al., 2012). In adolescence, individuals with ASD are significantly less likely to focus on the eye region (Klin et al., 2002). However, caution should be taken in predicting group differences as research suggests that looking behaviour in individuals with ASD may be influenced by the larger context in which faces are placed (Hanley, McPhillips, Mulhern, & Riby, 2013; Speer et al., 2007). For example, when a single face is isolated from a larger social context, TD individuals and those with ASD demonstrate similar looking patterns to whole faces, salient features of the face (eye and mouth) and non-salient features (hair, body, object, and background; Hanley et al., 2013; Speer et al., 2007). However, when faces are part of a larger social scene, which include a second face, differences in looking behaviour emerge. TD individuals look longer to faces in general, specifically the eye region, compared to those with ASD (Hanley et al., 2013; Speer et al., 2007). Further, when a face is part of a complex scene, individuals with ASD, compared to those with Williams Syndrome and TD peers, take longer to fixate on the face initially and overall attend to it less (Riby & Hancock, 2009). Thus, future research should use a variety of scenes involving both isolated and multiple faces, and social and non-social stimuli to better examine group differences.

1.2 Audiovisual Integration and Temporal Synchrony

Speech is often conceptualized as primarily an auditory task. However, research has shown that speech involves multiple sensory systems, such as auditory and visual systems (e.g.,
Patterson & Werker, 2003). Moreover, information from both auditory and visual modalities is integrated during speech production and proper integration is necessary for language development (Legerstee, 1990; Smith & Bennetto, 2007). At a broader level, multisensory information provides the individual with redundant information which increases the likelihood it will be perceived (Lewkowicz, Leo, & Simion, 2010).

Processing of audiovisual information begins early in life. TD infants as young as 2-months-old prefer congruent audio and visual information suggesting they can accurately integrate information from both sensory systems (Patterson & Werker, 2003). Typically developing newborns' ability to accurately integrate audio and visual information that is not species specific (Lewkowicz, Leo, & Simion, 2010) demonstrates an adaptive preference for matching bimodal information; this preference allows newborns to generalize their skills in a variety of contexts and access a wider knowledge base. However, difficulty integrating multisensory (e.g., audio and visual) information may be related to impairment in learning communicative or social information. For example, research suggests that audiovisual information is integrated differently in children with ASD and this may be related to deficits in their language abilities (Smith & Bennetto, 2007). Children with ASD are less accurate in reporting visual information and seem to rely more on auditory information than their TD peers, who rely more on visual information, suggesting they may have difficulty comprehending visual information (Iarocci, Rombough, Yager, Weeks, & Chua, 2010; Smith & Bennetto, 2007; Williams, Massaro, Peel, Bosseler, Suddendorf, 2004). Interestingly, when trained in speech-reading, children with ASD show increased accuracy in both visual only and audiovisual conditions suggesting speech reading abilities may affect audiovisual integration (Williams et al., 2004). Thus, difficulty with audiovisual integration in children with ASD may not be a deficit,
rather, it may be a difference or delay in processing which can be improved with training (Taylor, Isaac, & Milne, 2010; Williams et al., 2004).

The McGurk effect is a robust finding demonstrating the importance of audiovisual integration in speech perception (Taylor et al., 2010; Mongillo, Irwin, Whalen, Klaiman, Carter, & Schultz, 2008). To elicit this effect, the task typically uses mismatched audio (e.g. /ga/) and visual /ba/) information which, when integrated, results in the perception of a novel sound (e.g. /ða/; Mongillo et al., 2008). When presented with this task, children and adolescents with ASD seem to have more difficulty integrating audiovisual information compared to TD peers and those with Down Syndrome (Bebko, Schroeder, & Weiss, 2014). This task can also be conducted using non-linguistic information in which the visual of a ball bouncing (e.g. a rubber ball), either matches or mismatches the sound heard (e.g. a ping-pong ball bouncing; Mongillo et al., 2008). In this case, children and adolescents with ASD seem to accurately integrate audiovisual information (Irwin, Tornatore, Brancazio, & Whalen, 2011; Mongillo et al., 2008). Similarly, another study using beeps and flashes which, if accurately integrated produce an illusion, found that young adults with ASD were subject to this audiovisual illusion as were their TD peers (Van der Smagt, van Engeland, & Kemner, 2007). These findings using non-linguistic stimuli suggest that it is not a deficit in audiovisual integration in ASD per se, but a difference based on the stimuli presented. Further, research supports the notion that when looking at audiovisual integration across childhood and adolescence, atypical audiovisual processing may be a delay rather than a deficit (Taylor et al., 2010). Previous studies included wide age ranges (see Bebko et al., 2014; Irwin et al., 2011; Mongillo et al., 2008) and Taylor and colleagues (2010) sought to examine changes across childhood and adolescence using a traditional McGurk task. They found that while initially delayed, children with ASD showed a faster rate of audiovisual development
compared to their typically developing peers, resulting in similar audiovisual integration abilities in adolescence (Taylor et al., 2010). Thus, it appears that age as well as the specific stimuli presented influences the degree to which children with ASD integrate information.

Speech not only involves multiple modalities but, like many types of multimodal information, the information in the speech event occurs in-synch. That is, the visual information is in-synch with the auditory information. Information occurs in-synch (often enough) that TD infants by 10-to-16-weeks-old prefer synchronous, as opposed to asynchronous, events (Dodd, 1979). Children with ASD’s preference for synchrony, however, is contingent upon they type of stimuli presented. They demonstrate a preference for synchronous non-linguistic stimuli, but not a preference for synchrony when the stimuli are linguistic (Bebko et al., 2006). Bebko and colleagues propose two possible explanations for this difference. First, it is possible that children with ASD either do not detect the violation of synchrony or that they are slower in detecting a violation. Second, it may be that these children have not built up expectations about the synchrony of linguistic events. However, preliminary evidence from Bahrick and colleagues (2010) suggests that children with ASD show impairment in their ability to detect asynchrony in both social (human face and voice in neutral and positive affect) and non-social (object hitting a surface) stimuli. The latter results are preliminary and information such as the asynchrony times used are unknown; thus disparities may be due to methodological differences. In general, more information regarding children with ASD’s processing of asynchronous information is needed.

While TD children prefer synchrony (Bebko et al., 2006), they are able to detect asynchronous presentations. The time it takes infants to detect asynchrony seems to depend on how the stimuli are manipulated. If the stimuli presented differ in their use of linguistic or non-linguistic information, differences in children’s ability to detect asynchrony surface (Lewkowicz,
For example, TD infants are able to detect audiovisual asynchrony with as little as 350ms deviation when the event is non-linguistic (Lewkowicz, 1996) yet 4-year-old children require a minimum of 666ms asynchrony to detect a deviation in linguistic stimuli (Lewkowicz & Flom, 2013). By 6-years-old, children are able to detect asynchrony of linguistic stimuli with only a 366ms deviation (Lewkowicz & Flom, 2013) suggesting processing abilities increase significantly across development.

Linguistic information may require higher level processing abilities, and thus a longer presentation for deviant or discrepant information to be processed. For TD children age may be related to ability to detect asynchrony, but detection may also be related to general language or processing abilities. Unlike TD children, children with specific language impairment (SLI) fail to detect asynchrony at 666ms (Pons, Andreu, Sanz-Torrent, Buil-Legaz, & Lewkowicz, 2012). Slower processing of linguistic information may be reflected in delays in comprehension and production found in individuals with SLI. Thus, slower processing may require longer asynchrony presentations for a deviation to be noticed. This may suggest the ability to detect asynchrony lies on a continuum with individuals with processing impairments requiring longer presentations to detect asynchrony.

From an early age, TD infants have built-up expectations regarding multimodal events; for example, audio and visual information should be congruent (Patterson & Werker, 2003) and such information should be in-sync (Dodd, 1979). While TD children do not seem to differ in their ability to process audiovisual information or preferences for synchrony, children with ASD’s preferences and processing abilities seem to hinge on the social or non-social nature of the stimuli (Bebko et al., 2006). Thus, researchers may have to consider difficulties, such as their ability to generalize skills (i.e. looking behaviour on one stimuli to another) and language
abilities that may impact children with ASD’s processing of audiovisual information and preferences for synchronous events. Since many of these processes develop early and effortlessly in TD infants, examining other populations, such as children with ASD, in which delays or difficulties occur may help us to understand how these processes develop and what other abilities (e.g., speech-reading skills; Williams et al., 2004) may affect development.

**Present Study**

The present study examines preferences for conspecifics and audiovisual (AV) temporal processing in TD children and children with ASD and whether language ability is related to preferences and processing of AV information. Conspecific preferences are innate biases, which according to the social motivation theory (SMT), bias individuals toward social stimuli (e.g., human face; Chevalier et al., 2012). When social stimuli are interacted with it results in a pleasurable experience further motivating an individual to maintain social bonds (Chevalier et al., 2012). If indeed this is the case, then one might expect children to prefer human stimuli. The current study examines conspecific preferences using human, monkey and robot stimuli. However, it is not designed to test motivation for attending to any particular stimuli. Rather the current study forms the basis for exploring general preferences for species-specific stimuli. It is expected that children with ASD will not show conspecific preferences (see Chawarska et al., 2010; Kuhl et al., 2005), which contribute to the deficits or delays in social behaviour often seen in children with ASD (Chevallier et al., 2012).

Previous research examining children’s preferences for social (i.e. human faces) and non-social stimuli (e.g., children’s games; Bebko et al., 2006) has used stimuli that varies in many dimensions, such as their local and global features (see Lahaie, Mottron, Arguin, Berthiaume, Jemel, & Saumier, 2006). The current study will use stimuli with similar local and global
properties to the human face (e.g., an ‘eye region’) to examine whether TD children and those with ASD will prefer human (i.e. human face and speech) compared to monkey and robot stimuli. We also examine whether the same children’s preferences for synchrony are influenced by the type of stimuli presented. According to the weak central coherence theory (WCC), children with ASD display a cognitive profile with strengths and weaknesses (Happé et al., 2001; Happé & Frith, 2006). Stronger local as compared to global processing often found in children with ASD (Lahaie et al., 2006) may help explain their discrepancies in integrating and preferring non-social compared to social stimuli. Indeed, the non-social stimuli used in previous research (e.g., objects; Baharick et al., 2010) are less complex and contain fewer local features; thus, children with ASD are better able to focus on the global features (i.e. the whole object). Social stimuli are often faces, which contain many local features (i.e. eyes, mouth, hair, etc.) each of which could be focused on and require processing. The stimuli used in the current study vary along a continuum from human to monkey to robot but contain similar local features. It is not known how children will interpret each type of stimuli in terms of ‘social’ versus ‘non-social.’ If they do differentiate between them, then it is expected that TD children will prefer social, or human, stimuli while children with ASD may show preferences for or better processing of non-social stimuli. The stimuli also have similar local (e.g., an ‘eye region’) and global properties (e.g., a ‘face region’), allowing us to examine whether TD children and those diagnosed with ASD explore local and global regions similarly across the stimuli types.

Happé and Frith (2006) propose that individuals with ASD may have difficulty shifting between local (the default) and global processing making integrating social information difficult. Further, the local features of a face (e.g., eyes and mouth) do not disappear or change location on the face which allows individuals to scan between these features. Disengaging attention from one
area of the face (e.g., the mouth) to focus on another area present (e.g., the eyes) may be challenging for children with ASD. Disengagement requires executive functioning skills with which many children with ASD have difficulties (Elsabbagh et al., 2013; Landry & Bryson, 2004). Increased cognitive load needed to examine faces combined with poor cognitive flexibility (Verté, Geirts, Roeyers, Oosterlaan, & Sergeant, 2006) may restrict children with ASD’s ability to focus on salient facial features; reduced fixation to faces or salient regions on those faces would decrease the likelihood salient social or linguistic information would be integrated and asynchrony would be noticed. That is, if AV information is presented out of synch and children are focused on regions outside of the mouth, the asynchrony may not be detected. Moreover, if shifting between various regions on the face is difficult, then children with ASD may in general reduce the type and amount of information available to be processed. That is, if children with ASD become stuck on one area of the face (see Landry & Bryson, 2004), the majority of information available will come from that area. Further, difficulties shifting attention may influence how information is integrated. Flexible shifting between eye and mouth regions allows for information between the regions (e.g., linguistic and emotional information) to be integrated. Difficulties with disengaging attention may make the integration of multiple types of information less fluid. However, this should not suggest that flexibly shifting attention guarantees information will be accurately integrated; it is possible that children with ASD may show similar looking behaviour as TD children but differences in the processing or integration of the information could be found. The current study cannot directly test this hypothesis as looking behaviour gives no indication of how this information is being processed or used by the brain. The current study will use faces (human, monkey, and robot) to examine TD children and children with ASD’s audiovisual and temporal processing abilities.
Both social and language difficulties are a requirement for a diagnosis of ASD (APA, 1994). Children’s social and language skill development seem to be related in that conspecific preferences may aid in language development (Vouloumanos & Werker, 2007) and flexible shifting between salient AOI while looking at human faces may affect social-communicative development (Lewkowicz & Hansen-Tift, 2012). In this way, lack of social motivation toward human stimuli may limit children’s exposure to valuable learning experiences (Chevallier et al., 2012) which may affect their language comprehension and production. However, the current study will not examine children’s motivation for attending to stimuli thus the relationship between social motivation and language development will not be examined. We will, however, be examining whether or not a relationship between children’s language abilities and preferences for conspecifics and synchrony exists. As previously stated, TD infants shift their attention to salient facial regions (e.g., eyes or mouth) based on their current developmental stage (Lewkowicz & Hansen-Tift, 2012; Tenenbaum et al., 2012). Given that children with ASD are in general delayed in their language abilities (Weismer et al., 2010), it may be expected that they would use strategies similar to TD infants and increase looking to the mouth region (see Lewkowicz & Hansen-Tift, 2012). However, children with ASD have difficulty disengaging attention (Landry & Bryson, 2004) possibly reducing their ability to focus on salient facial regions (i.e., the mouth) which may affect their language development. The current study will use human, monkey, and robot static and dynamic stimuli to examine children’s looking behaviour and if this relates to their current language abilities.

**Task 1: Conspecific Preferences for Static Images**

In task 1, we examine preferences for conspecific stimuli by typically developing children and children diagnosed with ASD. It is hypothesized that TD children will look longer
to all stimuli in general compared to children with ASD (see Riby & Hancock, 2009). Second, we propose that compared to the monkey and robot, TD children will show a conspecific preference for the static human face (Vouloumanos et al., 2010) while children with ASD will not show this preference. Further, we postulate that TD children will look longer to the eye region on human and monkey faces while children with ASD will fixate longer on the mouth or background regions on these faces (Klin et al., 2002; Neumann, Spezio, Piven, & Adolphs, 2006). Lastly, we hypothesize that TD children will have stronger language and cognitive abilities compared to children with ASD, and these abilities will be related to looking behaviour.

3.1 Method

3.1.1 Participants

Written consent was given for all children by a parent and/or legal guardian prior to testing. Total looking time to faces was calculated in ms and a total of 27 TD children ($M = 49.51$ months, Range 47-58 months; 48% female) and 10 children with ASD ($M = 50.90$ months, Range 40-58 months; 22.2% female; 50% ASD, 10% PDD-NOS, 40% Autism) were included in analyses. Proportional looking time was calculated for the eye and mouth regions on each stimuli but only 25 TD children ($M = 49.28$ months, Range 47-58 months; 48% female) and 7 children with ASD ($M = 51.14$ months, Range, 40-58 months; 28.6% female; 42.9% ASD, 14.3% PDD-NOS, 42.9% Autism) were included in analyses. For proportional looking analyses, two TD children and three children with ASD were considered outliers with scores of +/- 2.5 standard deviations from the mean, and removed from analysis. Nine additional TD children or those diagnosed with ASD were tested but not included in either task’s analyses due to incomplete data collection ($n = 3$), experimenter error ($n = 2$), insufficient eye-tracking data ($n = 2$), or their primary language was not English ($n = 2$). Eye-tracking data were considered insufficient if
children tracked less than 30% of the total time stimuli were presented (Frank, Vul, & Saxe, 2012).

Typically developing children were recruited through the Child and Infant Learning and Development Research Group (Ch.I.L.D) at the University of Calgary. Children diagnosed with ASD were recruited through an ongoing longitudinal study and in partnership with the Society for Treatment of Autism. A diagnosis of ASD was confirmed either through parent report or by a multidisciplinary team at a treatment center (42.9% Autism; 14.3% PPD-NOS; 42.9% ASD). All children had no prenatal, birth, or neonatal complications, had normal or corrected vision and hearing, and had English as their primary language. Children diagnosed with ASD had no comorbid disorders or disabilities, such as attention-deficit/hyperactivity disorder, mental retardation, or communication disorders as indicated by parent report. As compensation for their participation, each child was given an age-appropriate toy and a certificate.

3.1.2 Materials

The Mullen Scales of Early Learning (MSEL; Mullen, 1995) was administered as a measure of children’s non-verbal and verbal abilities. The Mullen is a standardized measure and is appropriate for infants and children up to 68 months (Mullen, 1995). It has also been used with children diagnosed with ASD (Chawarska et al., 2010). The MSEL includes five subscales: gross motor, fine motor, visual reception, and receptive and expressive language, but the gross motor scale is not completed for children over 33 months (Mullen, 1995), as in the current study. A non-verbal intelligence quotient (NVIQ), comprised of the visual reception and fine motor scales, and verbal intelligence quotient (VIQ), consisting of the receptive and expressive language scales, were calculated for all children using a ratio IQ (i.e. age equivalent /
chronological age x 100; see Bishop, Guthrie, Coffing, & Lord, 2011 for calculation). A trained researcher administered the MSEL following the completion of both eye-tracking tasks.

A basic demographic form was given to all parents to complete. All but one parent of a child diagnosed with ASD completed the form. Questions addressed parent level of education, current employment, and languages spoken around the child (see Table 1). No significant group differences were found on any of the questions.

Table 1

Demographic Information by Group for all Participants

<table>
<thead>
<tr>
<th></th>
<th>ASD (n = 9)</th>
<th>TD (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mom age (mean in years)</td>
<td>35.00</td>
<td>36.33</td>
</tr>
<tr>
<td>Dad age (mean in years)</td>
<td>38.89</td>
<td>38.04</td>
</tr>
<tr>
<td>Mom Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or GED</td>
<td>10.00%</td>
<td>18.50%</td>
</tr>
<tr>
<td>Trade or Vocational</td>
<td>10.00%</td>
<td>3.70%</td>
</tr>
<tr>
<td>Associates or 2 years</td>
<td>10.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>College degree</td>
<td>60.00%</td>
<td>51.90%</td>
</tr>
<tr>
<td>Graduate Degree (Master's or Professional)</td>
<td>0.00%</td>
<td>25.90%</td>
</tr>
<tr>
<td>Mom Employment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working at a full time paid job</td>
<td>10.00%</td>
<td>33.30%</td>
</tr>
<tr>
<td>Working at a part-time paid job</td>
<td>30.00%</td>
<td>18.50%</td>
</tr>
<tr>
<td>Not currently working</td>
<td>20.00%</td>
<td>25.90%</td>
</tr>
<tr>
<td>Other</td>
<td>30.00%</td>
<td>22.20%</td>
</tr>
<tr>
<td>Household Income (mean)</td>
<td>$100,001 - $125,000</td>
<td>$125,001-$150,000</td>
</tr>
</tbody>
</table>

Note: 1 child from the ASD group (10%) is missing

3.1.3 Stimuli

Extending previous literature (e.g. Heron-Delaney et al., 2011), children were presented with three types of static stimuli: a human face, a monkey face, and a robot head (see Appendix B). The images were presented in black and white and measure 750x750 pixels in a .bmp format. The images were cropped in an oval shape surrounded by a black background. Images were
equated for luminosity, size, colour, and positioning. Two images were presented side-by-side on a screen, counterbalanced for side and stimuli paired with it. This resulted in six different stimulus pairings presented (e.g., human, monkey; monkey, robot).

3.1.4 Procedure

Stimuli were presented on a 21 inch Dell monitor with a resolution of 1680x1050. The SensoMotoric Instruments (SMI) eye tracker was placed directly below the screen and angled so there was minimal interference from blinking. Placed on top of the screen was a Logitech webcam so experimenters were able to monitor children and ensure they stayed on task. Eye-tracking data, or looking behaviour, was recorded at a rate of 120Hz using the SMI iViewX RED program. Eye tracking data was processed using the SMI BeGaze program. The experiment was run using a Dell laptop and the SMI experiment center program.

The preferential looking paradigm combined with eye-tracking technology (see Bebko et al., 2006) was used to reduce the cognitive and speech demands placed on children. This allowed researchers to assess lower functioning children on the Autism spectrum. Children were told they would see some pictures and videos but were not specifically instructed to watch them and their attention was not redirected to the screen if they looked away. Instruction was only given to children if the eye-tracker was unable to track their looking behaviour (e.g., “can you sit back please?”). Children were seated on a car seat that was strapped to a chair which raised them enough for the eye-tracker to track their eye movements. A two point calibration was presented first with a pulsing circle as the calibration point. Immediately following calibration, an attention getter was presented and when children’s eyes were tracked for 800ms the test trials began. Each set static side-by-side images was presented once for 3000ms for a total of six static images. The
order in which the sets of stimuli were presented was randomized for each child. Following the static images, the attention getter was presented, immediately followed by task 2 stimuli.

3.2 Results

A Mann-Whitney test was used due to unequal sample sizes and the non-normal distribution of the data. Total looking time to faces was calculated to examine between group differences. After correcting the alpha level, TD children looked significantly longer than children with ASD to human and robot static faces (see Table 2).

Table 2

*Mann-Whitney Test of between Group Differences to Static Faces*

<table>
<thead>
<tr>
<th>Face Stimuli</th>
<th>Group</th>
<th>Mdn</th>
<th>Mdn</th>
<th>U</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>TD (n = 27)</td>
<td>1051.00</td>
<td>383.67</td>
<td>56.00</td>
<td>-2.70*</td>
</tr>
<tr>
<td>Monkey</td>
<td>ASD(n =10)</td>
<td>941.00</td>
<td>623.44</td>
<td>82.00</td>
<td>-1.81</td>
</tr>
<tr>
<td>Robot</td>
<td></td>
<td>910.00</td>
<td>474.39</td>
<td>60.00</td>
<td>-2.57*</td>
</tr>
</tbody>
</table>

*Note: Two-tailed significance values were used. Median looking time is in ms. *p < .05

Conspecific preferences were examined using a paired sample t-test with a Bonferonni correction; total looking time to faces was used. No conspecific preferences were found for either group indicating neither TD children nor those with ASD looked significantly longer to the human compared to the monkey or robot static faces.

Proportional looking to the eye and mouth regions was calculated by taking the total time for that AOI (e.g., eye region) and dividing it by the total looking time to the whole face which equaled a percentage. Within groups an interesting result emerged in which both TD children and those with ASD showed preferences for the human and monkey eye regions, as compared to
the mouth regions (see Tables 3 & 4). Neither group preferred the robot eye compared to mouth region (see Tables 3 & 4).

Table 3

*Paired Sample t-test for TD Children's Preference for Static Eye Region*

<table>
<thead>
<tr>
<th>AOI</th>
<th>Eye</th>
<th>Mouth</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M(SD)</td>
<td>M(SD)</td>
<td></td>
</tr>
<tr>
<td>Stimuli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>73.92(16.33)</td>
<td>3.76(6.14)</td>
<td>18.25*</td>
</tr>
<tr>
<td>Monkey</td>
<td>65.62(22.86)</td>
<td>1.35(2.99)</td>
<td>13.08*</td>
</tr>
<tr>
<td>Robot</td>
<td>49.89(25.95)</td>
<td>30.17(20.17)</td>
<td>2.28</td>
</tr>
</tbody>
</table>

*Note: degrees of freedom equal 24. Proportion of looking in percent*

*p < .001

Table 4

*Paired Sample t-test for Children with ASD's Preference for Static Eye Region*

<table>
<thead>
<tr>
<th>AOI</th>
<th>Eye</th>
<th>Mouth</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M(SD)</td>
<td>M(SD)</td>
<td></td>
</tr>
<tr>
<td>Stimuli</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>62.33(34.76)</td>
<td>4.52(8.02)</td>
<td>4.05*</td>
</tr>
<tr>
<td>Monkey</td>
<td>60.63(35.26)</td>
<td>0 (0)</td>
<td>4.55*</td>
</tr>
<tr>
<td>Robot</td>
<td>46.81(37.48)</td>
<td>16.31(19.77)</td>
<td>1.97</td>
</tr>
</tbody>
</table>

*Note: degrees of freedom equal 6. Proportion of looking in percent*

*p < .05

As predicted, results of a Mann-Whitney test revealed TD children had higher NVIQ and VIQ scores compared to children with ASD (see Table 5). Unfortunately, after correcting the alpha level, NVIQ and VIQ scores were not significantly related to total looking time to faces or to proportionate looking time to eye and mouth regions.
Table 5

**Mullen NVIQ and VIQ Scores for Static Stimuli between Groups**

<table>
<thead>
<tr>
<th>Mullen score</th>
<th>Group</th>
<th>Median (Mdn)</th>
<th>Median (Mdn)</th>
<th>U</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TD (n = 24)</td>
<td></td>
<td>ASD(n = 6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NVIQ</td>
<td>106.72</td>
<td>70.00</td>
<td>23.00</td>
<td>-2.54*</td>
<td></td>
</tr>
<tr>
<td>VIQ</td>
<td>105.45</td>
<td>71.00</td>
<td>6.00</td>
<td>-3.42**</td>
<td></td>
</tr>
</tbody>
</table>

*Note: One TD child failed to complete the Mullen*  
*p < .05 **p < .001

### 3.3 Discussion

These results are important for further understanding differences in looking behaviour found in children with ASD and the influences stimuli have on looking behaviour. Children with ASD spent significantly less time looking to human and robot faces compared to TD children. This is consistent with previous research which has found TD children look longer to human faces compared to children with ASD (Shic et al., 2008). It is may also be that TD children displayed a novelty preference for the static robot face while children with ASD did not. Interestingly, while children with ASD looked less to faces in general, when using proportional looking times they showed similar proportional looking to both the eye and mouth regions on human and monkey stimuli. While this does not support our hypothesis, our hypothesis was based on previous research using total time spent looking to AOIs and found children with ASD looked less to key facial features (Shic et al., 2008). Use of proportional looking is fundamentally different and yields different results. While children with ASD spend less time looking at faces, when they are looking to faces they spend a similar proportion of time viewing salient regions. Thus, it is possible that children with ASD show similar developmental patterns to TD children in that by around 4 years of age they look longer to eye compared to mouth regions (Lewkowicz & Hansen-Tift, 2012). Given that the robot eye region does not look like
typical eyes, it is not surprising that children did not examine it longer than the mouth region. They may have been unsure of what information was being conveyed by the robot thus looked similarly to both regions. While it is possible that the reduced total looking time may affect the amount and type of information available to children, we are not able to conclude that children with ASD are not interested in the eye or mouth regions.

No conspecific preferences for static stimuli were found for either group. This is unexpected given previous research finding TD children, but not children with ASD, show conspecific preferences for human stimuli (Heron-Delaney et al., 2011; Chawarska et al., 2010). While the current study’s findings may seem to conflict with those of Heron-Delaney and colleagues (2011), a study by Di Giorgio, Leo, Pascalis, and Simion (2012) did not find conspecific preferences. The latter study used stimuli similar to the current study (monochrome, oval face shapes) and also did not find a preference for static human compared to monkey faces. Thus, the type of stimuli used may affect TD children’s conspecific preferences.

Lastly, results found that TD children had significantly higher NVIQ and VIQ scores compared to children with ASD. However, results from correlational analysis did not find either NVIQ or VIQ scores were related to total looking time to faces or looking behaviour to eye or mouth regions. It is interesting that children with higher verbal and non-verbal abilities, mostly TD children, did not look proportionally longer to the eye or mouth regions. This may not be surprising given the similar patterns of proportional looking by TD children and those with ASD. Thus, even though children with ASD had significantly lower language abilities they looked proportionally similar across stimuli. However, considering TD children did look significantly longer to human and robot faces in general, it may be surprising that no significant correlations for total looking time were found. This may suggest that children’s verbal and non-verbal
abilities are not related to total looking behaviour, but other areas of children’s language development should be examined to see if relationships exist.

**Task 2: Audiovisual and Temporal Integration**

In this task, we examine whether TD children and children with ASD show similar sensitivities to temporal synchrony of auditory and visual information. For the dynamic stimuli it is proposed that TD children will look longer to stimuli in general, compared to children with ASD, and show a conspecific preference for the human face regardless of asynchrony. We also hypothesize that TD children will look longer to the human eye region whereas children with ASD will look longer to the mouth or background. Due to the novelty of the robot and monkey stimuli in the context of temporal synchrony it is difficult to make specific hypotheses regarding looking behaviour. We do, however, expect TD children to prefer synchronous presentations of all stimuli (i.e. human, monkey, and robot) while children with ASD will not detect temporal asynchrony. Lastly, children’s language and cognitive abilities are predicted to be related to their preference for conspecific stimuli and synchrony. That is, preferences for human stimuli will be related to stronger language abilities as will preferences for synchronous stimuli.

**4.1 Methods**

**4.1.1 Participants**

Children were the same as those tested in task 1, however not all children completed task 2. Analyses examining total looking time to faces included 27 TD children ($M = 49.52$ months, Range 47-58 months; 48% female) and 10 children with ASD ($M = 50.90$ months, Range 40-58 months; 22.2% female; 50% ASD, 10% PPD-NOS, 40% Autism). Proportional looking time analyses included 23 TD children ($M = 49.04$ months, Range 47-58 months old; 39.1% female) and 6 children with ASD ($M = 51$ months, Range 40-58 months old; 0% female; 57.1% Autism,
42.9% ASD). For proportional looking analyses, four TD and three children with ASD were not included due to being outliers; their scores were +/- 2.5 standard deviations from the mean. Nine additional TD children and those diagnosed with ASD were tested but not included in analysis due to incomplete data collection (n = 3), experimenter error (n = 2), insufficient eye-tracking data (n = 2), or their primary language was not English (n = 2). The same criteria as in study 1 were used for determining insufficient eye-tracking data.

4.1.2 Materials

Parents completed one demographic form during their visit. Data from the demographic form collected on all the children are presented in Task 1 (see Table 1).

4.1.3 Stimuli

The same faces that were used in the first task were converted into dynamic videos and paired with audio stimuli in order to examine conspecific preferences for faces. This resulted in three types of novel audiovisual stimuli: a human face and speech (e.g., /kēf/), a monkey face and monkey vocalizations (e.g., coos), and a robot head with robotic sounds (e.g., beeps). Each stimulus was presented four times with varying asynchronies: 0ms, 660ms, 1000ms, and 1500ms with a total of 12 videos presented to each child. This allowed the researchers to see if tolerance for asynchrony was related to which stimulus was presented. Audiovisual asynchronies were selected based on previous research that used a minimum threshold of 666ms with children around 4 years of age (Lewkowicz & Flom, 2013; Pons et al., 2012). The longer asynchronies (i.e. 1000ms and 1500ms) were meant to ensure children with ASD had sufficient time to process and notice the asynchrony as previous research has found children with language difficulties may require longer processing times to detect asynchrony (Pons et al., 2012).
4.1.4 Procedure

As stimuli for task 2 immediately followed study one, children remained seated on the car seat facing the eye-tracker. No new instructions were provided and children were not prompted to look at the screen if they looked away. Videos ranged from 13000ms to 14000ms long and were semi-randomized. Videos were presented in order of asynchrony (0ms, 660ms, 1000ms, then 1500ms) but the stimuli order within an asynchrony (human, monkey, or robot) was randomized. After each set of asynchrony the same attention getter as in task 1 was presented and disappeared once children fixated on it for 800ms. At the end of the study, a five point validation with a pulsing white dot was used.

4.2 Results

This task was run to look at differences in audiovisual integration between groups and the effect temporal synchrony played in this integration. A Mann-Whitney test was run to look at between subject differences in looking behaviour to stimuli. Results found that TD children looked significantly longer than children with ASD on many of the stimuli, regardless of asynchrony (see Table 6).

Table 6

*Mann-Whitney Test of between Group Differences to Dynamic Faces*

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Group</th>
<th>TD(n = 27)</th>
<th>ASD(n = 9)</th>
<th>U</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Mdn</em></td>
<td><em>Mdn</em></td>
<td><em>U</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robot Face 0ms</td>
<td>9460</td>
<td>1025.00</td>
<td>40.00</td>
<td>-2.98</td>
<td></td>
</tr>
<tr>
<td>Human Face 660ms</td>
<td>10455.00</td>
<td>150</td>
<td>11.00</td>
<td>-4.04</td>
<td></td>
</tr>
<tr>
<td>Monkey Face 660ms</td>
<td>8469.00</td>
<td>902</td>
<td>8.00</td>
<td>-4.15</td>
<td></td>
</tr>
<tr>
<td>Robot Face 660ms</td>
<td>9585</td>
<td>1018</td>
<td>42.00</td>
<td>-2.90</td>
<td></td>
</tr>
<tr>
<td>Monkey Face 1000ms</td>
<td>7182</td>
<td>1035</td>
<td>44.50</td>
<td>-2.81</td>
<td></td>
</tr>
<tr>
<td>Monkey Face 1500ms</td>
<td>4636</td>
<td>725</td>
<td>23.00</td>
<td>-3.60</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* Two tailed exact significance values were used. Looking times are in ms
*p < .05  **p < .001

Paired sample t-tests were run looking at conspecific preferences based on total looking time to human, monkey, and robot stimuli at each of the asynchronies. Similar to task 1, neither children with ASD nor TD children looked significantly longer to human dynamic stimuli at any of the asynchronies.

Paired sample t-tests were also run looking at differences in proportional looking to the eye and mouth region on each stimulus. After correcting the alpha level it was found that TD children looked significantly longer to the monkey eye versus mouth region when the stimulus was synchronous (0ms) and 1500ms asynchronous (see Table 7). Children with ASD did not look significantly longer to the eye region on any of the stimuli at any asynchrony.

Table 7

*Paired Sample t-test for TD Children's Preference for Dynamic Eye Region*

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Eye M(SD)</th>
<th>Mouth M(SD)</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monkey 0ms</td>
<td>44.19(18.25)</td>
<td>19.42(9.45)</td>
<td>5.72**</td>
</tr>
<tr>
<td>Monkey 1500ms</td>
<td>43.85(23.03)</td>
<td>19.40(14.46)</td>
<td>3.61*</td>
</tr>
</tbody>
</table>

Note: degrees of freedom equal 22. Proportion of looking in percent
* *p < .05  **p < .001

To examine preferences for synchrony, proportional looking to the eye and mouth regions were collapsed for stimuli at each asynchrony (e.g., human eye 0ms, human mouth 0ms, monkey eye 0ms, etc. = 0ms). A Mann-Whitney test was run examining our hypothesis that TD children would look proportionally longer to stimuli when they were synchronous as compared to children with ASD; this hypothesis was not supported. It was, however, found that TD children looked significantly longer than children with ASD when stimuli were asynchronous by 660ms (see Table 8). Between group preferences were also examined using total looking time to
faces for each stimuli at all asynchronies (e.g., human, monkey, and robot faces at 0ms). A Mann-Whitney test found that TD children looked significantly longer than children with ASD to faces in general at each asynchrony (see Table 9).

Table 8

*Mann-Whitney Test for Preference for Synchrony between Groups using Proportional Looking*

<table>
<thead>
<tr>
<th>Asynchrony</th>
<th>TD (n = 23)</th>
<th>ASD(n = 6)</th>
<th>U</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>0ms</td>
<td>38.36</td>
<td>26.56</td>
<td>33.00</td>
<td>-1.94</td>
</tr>
<tr>
<td>660ms</td>
<td>40.57</td>
<td>24.23</td>
<td>11.00</td>
<td>-3.12*</td>
</tr>
<tr>
<td>1000ms</td>
<td>35.66</td>
<td>27.26</td>
<td>46.00</td>
<td>-1.24</td>
</tr>
<tr>
<td>1500ms</td>
<td>35.44</td>
<td>29.68</td>
<td>45.00</td>
<td>-1.29</td>
</tr>
</tbody>
</table>

*Note:* Two tailed exact significance values were used. Median proportional looking in percent

*p < .05

Table 9

*Mann-Whitney Test for Preference for Synchrony between Groups using Total Looking Time*

<table>
<thead>
<tr>
<th>Asynchrony</th>
<th>TD (n = 27)</th>
<th>ASD(n = 9)</th>
<th>U</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>0ms</td>
<td>9977.33</td>
<td>26.56</td>
<td>49.00</td>
<td>-2.65*</td>
</tr>
<tr>
<td>660ms</td>
<td>9572.67</td>
<td>24.23</td>
<td>16.00</td>
<td>-3.85**</td>
</tr>
<tr>
<td>1000ms</td>
<td>7562.33</td>
<td>27.26</td>
<td>50.00</td>
<td>-2.61*</td>
</tr>
<tr>
<td>1500ms</td>
<td>7562.33</td>
<td>29.68</td>
<td>40.00</td>
<td>-2.98*</td>
</tr>
</tbody>
</table>

*Note:* Two tailed exact significance values were used. Median total looking times in ms

*p < .05 **p < .001

Collapsing proportional looking time to stimuli at each asynchrony, within group differences were examined using paired sample t-tests to see whether or not children in each group looked longer to synchronous compared to asynchronous stimuli. Neither TD children nor
those with ASD looked significantly longer to the synchronous presentation compared to any of the asynchronies when using proportional looking times. However, when collapsing total looking time to faces across stimuli to examine within group preferences for synchrony significant differences are found. Typically developing children looked significantly longer to all synchronous faces compared to when stimuli were asynchronous at 1000ms and 1500ms (see Table 9). No differences in total looking time to synchronous or asynchronous stimuli were found for the ASD group (see Table 10).

Table 10

*Paired Sample t-test for Within Group Preference for Synchronous Stimuli*

<table>
<thead>
<tr>
<th>Group</th>
<th>Synchrony</th>
<th></th>
<th></th>
<th></th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0ms</td>
<td>1000ms</td>
<td>1500ms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TD (n = 27)</td>
<td>8847.93</td>
<td>7251.57</td>
<td>6213.04</td>
<td>2.79*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3385.96</td>
<td>2997.48</td>
<td>2136.30</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

*p < .05 **p < .001

Similar to task 1, TD children had significantly higher NVIQ and VIQ scores compared to children with ASD (see Table 11). Given that tests of within and between group differences consistently found distinctions in looking behaviour at 0ms and 1500ms asynchronous, proportional looking to eye and mouth regions and total looking times to faces for stimuli at those two asynchronies were correlated with NVIQ and VIQ scores. After correcting the alpha level, no significant relationships between children’s NVIQ and VIQ scores and looking behaviour were found.
Table 11

*Mullen NVIQ and VIQ Scores for Dynamic Stimuli between Groups*

<table>
<thead>
<tr>
<th>Mullen score</th>
<th>Group</th>
<th>Mdn</th>
<th>Mdn</th>
<th>U</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TD (n = 22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NVIQ</td>
<td>107.27</td>
<td>79.50</td>
<td>24.00</td>
<td>-2.35*</td>
<td></td>
</tr>
<tr>
<td>VIQ</td>
<td>105.45</td>
<td>70.25</td>
<td>6.00</td>
<td>-3.36**</td>
<td></td>
</tr>
</tbody>
</table>

*Note: One TD child failed to complete the Mullen
*p < .05 **p < .001*

4.3 Discussion

Between groups differences found TD children looked significantly longer to most faces across all asynchronies compared to children with ASD. It is interesting that differences were found for the monkey face only when it was asynchronous (660ms, 1000ms, and 1500ms) possibly indicating TD children cared more about, or were more interested in, processing asynchronous monkey stimuli than children with ASD. Typically developing children also looked significantly longer to the robot stimuli when it was synchronous (0ms) and asynchronous by 660ms. Due to the novelty of the robot stimulus TD children may have looked longer at 0ms and 660ms, however, by 1000ms stimuli were too asynchronous to process thus TD children significantly decreased their looking behaviour. Conversely, it was only at 660ms that significant group differences were found for the human face with TD children looking significantly longer. Possibly at 660ms TD children noticed the asynchrony on the human stimuli thus increased their looking to assist in processing the linguistic stimuli.

Similar to task 1, neither group showed conspecific preferences for human dynamic faces compared to monkey or robot faces at any of the asynchronies. This may indicate novelty preferences for both monkey and robot stimuli by both groups in which children looked longer to monkey and robot stimuli than expected leading to lack of conspecific preferences.
The current study hypothesized that for the human stimuli, TD children would look longer to the eye compared to mouth region while children with ASD would look longer to the mouth region. No specific hypothesis regarding looking to eye or mouth regions on monkey or robot stimuli were made. Unexpectedly, the only significant difference was found for TD children who looked proportionally longer to the monkey eye compared to mouth region when the stimulus was synchronous (0ms) and 1500ms asynchronous. When stimuli were synchronous TD children may have been more interested in the information provided by the monkey eye region, or less interested in the mouth region. However, by 1500ms audiovisual information may have been too incongruent for TD children to try to process, thus they looked to the eye region more than the mouth region.

Preferences for synchrony are found for TD children, regardless of stimuli presented, while children with ASD’s preferences may be conditional upon the type of stimuli presented (Bahrick et al., 2010; Bebko et al., 2006). Using total looking time to faces it was found TD children looked longer than children with ASD to all stimuli at each asynchrony. However, when using proportional looking time it was only at 660ms that TD children looked longer to salient facial regions than did children with ASD. While children with ASD are looking less to faces in general, they seem to look proportionally similar to salient regions compared to TD children, except at 660ms. Preferences for synchrony were also examined within groups. When using proportional looking collapsed across stimuli at each asynchrony neither the TD nor ASD group looked proportionally longer to synchronous compared to asynchronous stimuli. Total looking time to stimuli at each asynchrony found that TD children looked significantly longer to faces at 0ms compared to 1000ms and 1500ms asynchronous. Children with ASD did not differ in their looking to synchronous or asynchronous faces which is consistent with previous research.
(Bahrick et al., 2010). These results support our hypothesis that children with ASD did not notice or care that stimuli were asynchronous and that TD children preferred synchronous stimuli.

Similar results as task 1 for the NVIQ and VIQ scores were found in that TD children had higher verbal and non-verbal abilities but no relationship between these abilities and looking behaviour were found. Correlational analysis was only run looking at proportion of looking time to eye, mouth, and face regions at 0ms and 1500ms asynchronous as consistent differences during these times were found. Neither children’s NVIQ nor VIQ scores were related to proportional or total looking to dynamic stimuli. Children with ASD’s overall lower scores do not seem to be related to their looking behaviour.

**General Discussion**

The present study had two aims. The first was to examine TD children and children with ASD’s preferences for conspecific static and dynamic faces. The second was to examine the role temporal synchrony plays in the preference for and processing of audiovisual dynamic stimuli. In total, TD children look longer to face than children with ASD to most static and dynamic stimuli. Results support and extend previous findings that TD not only look longer than children with ASD to human (see Shic et al., 2008), but also to monkey and robot faces. Typically developing children may prefer the stimuli more, especially the novelty of the monkey and robot. Consistent with previous research, children with ASD may have more difficulty processing stimuli thus look to faces less; this may affect the availability of social cues which could impact learning (Riby & Hancock, 2009). Not only is children with ASD’s failure to orient to social stimuli greater than for non-social stimuli (Dawson, Webb, & McPartland, 2005), but their lack of bias for human stimuli (Chawarska et al., 2010; Kuhl et al., 2005) may suggests the current study’s use of faces may have reduced children with ASD ability to orient to stimuli due to their complex, social
nature; thus, children with ASD may not have been as intrinsically motivated to attend to the faces (see Chevallier et al., 2012). Further, reduced total looking may be related to difficulties with executive functioning, specifically disengagement of attention; children with ASD may get stuck attending to less relevant information (e.g., background; Rice et al., 2012) and take longer to disengage their attention (see Landry & Bryson, 2004) to focus on faces. Thus, children with ASD may take longer to attend to stimuli or attend to them less.

Proportional looking to eye and mouth regions indicates the amount of time children look to AOIs when viewing the faces. Results show similar proportional looking for both groups, TD and ASD, for static and dynamic stimuli. For static stimuli, both groups look proportionally longer to eye as compared to mouth regions on human and monkey stimuli. This result is important given that previous research often compares children with ASD’s looking to that of TD children (Klin et al., 2002; Rice et al., 2012) and may not examine within group behaviour. While children with ASD may spend proportionally less time viewing human faces or specific AOI (Klin et al., 2002; Rice et al, 2012), longer looking to eye regions on human and monkey stimuli may suggest children with ASD find typical eye regions more important, or interesting, than the mouth region on static stimuli. This is important for better understanding the type of information children with ASD have available to process. Another possible explanation is that increased looking to the eye region on static stimuli by children with ASD was actually a result of difficulties with executive functioning. Children with ASD have difficulty disengaging their attention (Landry & Bryson, 2004) and may have become ‘stuck’ on the eye region and had difficulty disengaging attention to the mouth region which resulted in similar proportional looking as TD children. Decreased overall looking to faces but similar proportional looking may also be explained by children with ASD’s greater local compared to global processing (see
Lahaie et al., 2006) in that they scan overall faces less but focus more intently on local features. Better understanding of children with ASD’s processing abilities is needed to determine the cognitive mechanisms contributing to ASD.

Conversely, for dynamic stimuli both TD children and those with ASD in general did not look longer to eye compared to mouth regions on stimuli. The only significant differences were found for TD children looking longer to the monkey eye versus mouth region at 0ms and 1500ms asynchronous. This may indicate the for TD children the type of stimuli and how it is presented may affect looking behaviour. It is possible that both TD children and those with ASD are able to flexibly shift their attention between eye and mouth regions on dynamic stimuli (Shic et al., 2008). This hypothesis is supported for TD children (see Shic et al., 2008) but may be unexpected for children with ASD due to their difficulties disengaging attention (Landry & Bryson, 2004). It may be children with ASD have had more difficulty shifting and disengaging attention but due to small sample size no significant differences are found. Thus, results should be interpreted with caution. Taken together it is interesting that while children with ASD look significantly less to faces, both static and dynamic, proportional looking reveals similarities between groups. Future research should consider examining both total looking time as well as proportional looking time to better understand where children with ASD are attending to when looking at faces.

Lack of conspecific preferences for static or dynamic stimuli were found for both TD children and those with ASD. Previous research, however, has found conspecific preferences in TD newborns (Heron-Delaney et al., 2011; Vouloumanos & Werker, 2007) possibly suggesting that by early childhood children may not prefer human stimuli over novel stimuli. Thus, TD 4-year-olds in the current study may have sufficient previous exposure with human faces that more
resources were allocated to novel monkey and robot stimuli. While fundamental biases which
direct infant’s attention to socially relevant stimuli may allow them to process salient
information, such as speech (Vouloumanos & Werker, 2007), and have cascading effects on their
development (Chevallier et al., 2012), fixating on novel stimuli may also be an adaptive learning
mechanism. Increased looking to novel stimuli may increase the type and amount of information
available to children. Lack of conspecific preferences may actually demonstrate children’s
flexibility in processing information in that processing is not based solely on fundamental biases
(see Chevallier et al., 2012), but also on stimuli’s novelty and possible importance. While the
current study was not created to directly test the social motivation theory, early biases for social
stimuli may be important for infants to quickly learn and adapt to their environment, but as
young children broader tuning or interest in stimuli may be more beneficial for learning. Further,
results from previous research and the current study may suggest TD children’s conspecific
preferences are contingent upon the type of stimuli presented. Previous research using coloured,
naturalistic pictures of human and monkeys (Heron-Delaney et al., 2011) versus monochrome
oval faces on a black matt– similar to the current study– (Di Giorgio et al., 2012) have found
different results. Thus, the current study’s use of cropped black-and-white faces may affect the
lack of conspecific preferences found.

Previous research has suggested that TD children are able to detect and prefer synchrony
regardless of the type of stimuli presented (e.g., social or non-social; Bahrick et al., 2010; Bebko
et al., 2006). Prior conceptualization of stimuli as social led to an understanding that children
with ASD may process social and non-social information differently which may affect their
preference for synchrony (Bebko et al., 2006). The current study’s use of faces reveal TD
children have longer total looking than children with ASD across stimuli at each asynchrony;
however, children with ASD look proportionally similar to AOIs as TD children at each asynchrony, except at 660ms. Reduced total looking to dynamic faces by children with ASD may affect their ability to notice asynchrony, hence why at 660ms there are differences in proportional looking. Possibly at 660ms TD children notice the asynchrony and increase looking to salient regions, while children with ASD do not change their looking behaviour. Within group differences also reveal that TD children notice asynchrony by looking longer to synchronous (0ms) compared to 1000ms and 1500ms asynchronous stimuli. Lack of significant differences from 0ms to 660ms asynchronous may indicate TD children do not notice the asynchrony and look similar between presentations; or that children notice the asynchrony and maintain long looking to assist in processing. The latter hypothesis is congruent with finding differences in proportional looking at 660ms where TD children may have increased their looking to salient regions to assist in processing asynchrony. Children with ASD, however, show no differences in looking to synchronous and asynchronous stimuli. This supports previous findings that children with ASD do not prefer synchrony (Bahrick et al., 2010; Bebko et al., 2006) and extending these findings by using both human and non-human faces.

As expected, for both tasks TD children had higher NVIQ and VIQ scores as compared to children with ASD. However, verbal and non-verbal abilities were not related to proportional or total looking behaviour to static or dynamic stimuli. Two factors should considered given these results. First, similar patterns of proportional looking between groups may explain why language abilities are not related to proportional looking. However, given significant differences in total looking time lack of significant correlations may be surprising. This leads to the second factor: these results bring into question the relationship between verbal and non-verbal abilities and looking behaviour. Namely, that proportional or total looking to faces may not necessarily mean
the information from these areas is being processed, or being processed in a typical manner. Eye-tracking information cannot reveal how information is being processed or used by the brain (Boraston & Blakemore, 2007). It may be that children with lower NVIQ and VIQ abilities, while not looking proportionally less to eye and mouth regions, may struggle to process the information provided. That is, attending to does not equal processing of stimuli. Thus, children with lower NVIQ and VIQ abilities, as typically found in children with ASD, may need to look just as long if not longer to salient AOI compared to children with higher NVIQ and VIQ scores to assist in processing. It is also possible that other areas of language development, for example pragmatics may be related to proportional looking behaviour to specific AOI on faces (Chawarska et al., 2010) but such areas were not examined in the current study.

5.1 Conclusion

The current study makes important distinctions between total and proportional looking which may affect how children with ASD are thought to process stimuli. Reduced total looking by children with ASD to static and dynamic stimuli may influence the amount of information available; yet, similar proportional looking patterns as TD children may suggest similar types of information are available to children with ASD. Further, reduced total looking across asynchronies for children with ASD may result in their inability to notice asynchrony, as reflected by TD children’s significantly longer looking at 660ms. Previous research has conceptualized TD and children with ASD’s processing and preference differences in relation to the type of stimuli used (i.e. social, human, or non-social, shapes; Bahrick et al., 2010) and the current results may be important for re-conceptualizing how stimuli are labelled. Given that results of proportional looking in general were not different between human and non-human stimuli, previous findings may be influenced by significant differences between stimuli used.
Typically, human faces contain many local and global features (see Lahaie et al., 2006) while non-social stimuli used in previous studies are more simple (see Bahrick et al., 2010). The current study’s use of faces, both conspecific and from other species, may help to examine whether or not looking behaviour in TD and children with ASD are affected by the type of face presented, or if looking behaviour is generalized to all face-like-stimuli. Children in both groups seem to show similar looking to human and non-human stimuli, with few exceptions for TD children; these subtle exceptions may relate to TD children’s greater cognitive flexibility compared to children with ASD (Verté et al., 2006). Overall, similar proportional looking and lack of preferences for conspecifics by both groups may highlight the influence stimuli plays in children’s preference and processing abilities.

### 5.2 Limitations and Future Research

Several limitations should be noted for this study. First, AOI-based approaches may not be sensitive enough to reveal significant group differences due to the larger area they encompass (Yi, Feng, Quinn, Ding, Li, Liu, & Lee, 2014). It is possible that in the current study analyzing each eye separately for the eye region AOI could reveal significant group differences; however, the practical utility of such knowledge (e.g., longer looking to the left eye versus the right) has yet to be determined. A second limitation is that the current study did not utilize saccade information which may be important for future research to understand not only where children look on faces but also the sequence or path of their looking behaviour (Yi et al., 2014). Reduced or atypical scanning between salient facial regions has been found in children with ASD (Yi et al., 2014) and may influence how social and communicative information are integrated and processed. A third limitation common to research with atypical populations is small sample size which significantly reduces power. Considering the ASD group consisted of less than 10
children, significant group differences found should be interpreted cautiously. Fourth, the only comparison group consisted of TD children. Future research should use additional control groups, such as children with developmental disabilities or specific language impairment, to further examine the roles cognitive functioning and language abilities play in processing and integrating abilities. Additional comparison groups would also help research disentangle which deficits (e.g., social, language, cognitive) seen in children with ASD are contributing to differences seen in looking behaviour. Lastly, the study design is limiting as results are based on looking behaviour which do not reveal what information and the way in which the information is being processed by the brain (Boraston & Blakemore, 2007). While studies combining brain imaging and eye-tracking technology are a useful next step, the current research is also important for creating a basic understanding about whether or not group differences exist and what these group differences may mean. Fundamental research like the current study are needed for future research to build upon and determine what may be causing or underlying these differences and how they may be related to other areas of development (Behrmann, Thomas, & Humphreys, 2006). Lastly, it is unclear whether or not the differences in looking behaviour in children with ASD are present at birth or as a result of the onset of other symptoms (Chawarska et al., 2010). Future research may examine younger siblings of children diagnosed with ASD to survey if these differences are present in infancy and continue in children who are later diagnosed.
References


Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth


Riby, D. M., & Hancock, P. J. B. (2009). Do faces capture the attention of individuals with


Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in

Yi, L., Feng, C., Quinn, P. C., Ding, H., Li, J., Liu, Y., & Lee, K. (2014). Do individuals with
and without autism spectrum disorder scan faces differently? A new multi-method look at
Appendix A: Visual Stimuli

Task 1 Visual Stimuli

Task 2 Visual Stimuli

Each stimulus was presented four times asynchronous: 0ms, 660ms, 1000ms, and 1500ms
Appendix B: Ethics Certification

See attached.