

ScholarWorks@GSU

Visual Scanning of Dynamic Affective Stimuli in Autism Spectrum Disorders

Authors	McManus, Susan M.
Citation	McManus, Susan M.. "Visual Scanning of Dynamic Affective Stimuli in Autism Spectrum Disorders." 2012. Dissertation, Georgia State University. https://doi.org/10.57709/3098238
DOI	https://doi.org/10.57709/3098238
Download date	2026-06-06 19:36:59
Link to Item	https://hdl.handle.net/20.500.14694/12830

VISUAL SCANNING OF DYNAMIC AFFECTIVE STIMULI IN AUTISM
SPECTRUM DISORDERS

by

SUSAN M. MCMANUS

Under the Direction of Diana L. Robins, Ph.D.

ABSTRACT

The accurate integration of audio-visual emotion cues is critical for social interactions and requires efficient processing of facial cues. Gaze behavior of typically developing (TD) individuals and individuals with autism spectrum disorders (ASD) was measured via eye-tracking during the perception of dynamic audio-visual emotion (DAVE) stimuli. This study provides information about the regions of the face sampled during an emotion perception task that is relatively more complex than those used in previous studies, providing both bimodal (auditory and visual) and dynamic (biological motion) cues. Results indicated that the ASD

group was less accurate at emotion detection and demonstrated less of a visual-affective bias than TD individuals. Both groups displayed similar fixation patterns across regions during the perception of congruent audio-visual stimuli. However, between-group analyses revealed that fixation patterns differed significantly by facial regions during the perception of both congruent and incongruent movies together. In addition, fixation duration to critical regions (i.e., face, core, eyes) was negatively correlated with measures of ASD symptomatology and social impairment. Findings suggest weaknesses in the early integration of audio-visual information, automatic perception of emotion, and efficient detection of affective conflict in individuals with ASD. Implications for future research and social skills intervention programs are discussed.

INDEX WORDS: Autism Spectrum Disorders, Emotion perception, Eye-tracking, Gaze behavior, Face processing

VISUAL SCANNING OF DYNAMIC AFFECTIVE STIMULI IN AUTISM
SPECTRUM DISORDERS

by

SUSAN M. MCMANUS

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2012

Copyright by
Susan M. McManus
2012

VISUAL SCANNING OF DYNAMIC AFFECTIVE STIMULI IN AUTISM
SPECTRUM DISORDERS

by

SUSAN M. MCMANUS

Committee Chair: Diana L. Robins

Committee: Erin B. Tone

David J. Marcus

Robert D. Lutzman

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2012

ACKNOWLEDGEMENTS

I would like to extend my gratitude to the faculty members who have provided mentorship and supported my training throughout graduate school. I would also like to acknowledge the Georgia State University Dissertation Grant program for the funding on this project, and Harder & Co. for assistance with stimulus development.

TABLE OF CONTENTS

ACKNOWLEDGMENTS		iv
LIST OF TABLES		vii
LIST OF FIGURES		ix
1 INTRODUCTION		1
1.1 Autism Spectrum Disorders (ASD)		2
1.2 Development of Emotion Perception		4
1.3 Impact of Emotion Perception Deficits on Social Functioning in ASD		7
1.4 Components of Emotion Perception		9
1.5 Gaze Behavior		18
1.6 Static vs. Dynamic Stimuli		21
1.7 Neural Correlates of Social Perception		23
1.8 Preliminary Findings		29
1.9 The Current Study		30
1.10 Hypotheses		35
2 METHOD		41
2.1 Participants		41
2.2 Experimental Stimuli		42
2.3 Diagnostic Measures		43
2.4 Inclusion and Exclusion Measures		46
2.5 Apparatus		51
2.6 Procedure		53
3 RESULTS		58

3.1	Hypothesis 1: Accuracy and Reaction Time during Congruent Movies	61
3.2	Hypothesis 2: Fixation Pattern during Congruent Movies	62
3.3	Hypothesis 3: Behavioral Responses and Reaction Time during Incongruent Movies	65
3.4	Hypothesis 4: Fixation Pattern during Incongruent Movies as Compared to Congruent Movies	67
3.5	Hypothesis 5: Exploratory Analysis of Initial Fixations	74
3.6	Hypothesis 6: Exploratory Analysis of the Relationship between Visual Fixations and Social Functioning	77
4	DISCUSSION	78
4.1	Behavioral Performance	78
4.2	Fixation Patterns: Congruent and Incongruent Movies	87
4.3	Social Cognition	95
4.4	Conclusions	98
4.5	Limitations and Future Directions	101
	REFERENCES	105
	APPENDIX	125

LIST OF TABLES

Table 1	<i>Summary of sample and task characteristics of relevant emotion perception studies</i>	33
Table 2	<i>Movie duration ranges in seconds</i>	43
Table 3	<i>Group performance on cognitive, perceptual, and emotional measures</i>	59
Table 4	<i>Number of significant elevations (T-score ≥ 70) on the BASC-2</i>	60
Table 5	<i>Group performance on subscales of interest for the BASC-2</i>	60
Table 6	<i>Group fixation duration (sec) data for face and non-face regions during congruent movies of Task 1</i>	62
Table 7	<i>Group fixation duration (sec) data for core and periphery regions during congruent movies of Task 1</i>	64
Table 8	<i>Group fixation duration (sec) data for eye, nose, and mouth regions during congruent movies of Task 1</i>	65
Table 9	<i>Group modality bias data for incongruent movies of Task 2</i>	66
Table 10	<i>Response time data for congruent and incongruent movies of Task 2</i>	67
Table 11	<i>Group fixation duration (sec) data for face and non-face regions during congruent and incongruent movies of Task 2</i>	68
Table 12	<i>Group fixation duration (sec) data for core and periphery regions during congruent and incongruent movies of Task 2</i>	70
Table 13	<i>Group fixation duration (sec) data for eye, nose, and mouth regions during congruent and incongruent movies of Task 2</i>	73
Table 14	<i>Group initial fixation data for eye, nose, and mouth regions during congruent and incongruent movies of Task 2</i>	75

Table 15	<i>Correlations between fixation duration data and measures of social functioning</i>	78
Table 16	<i>Summary of fixation duration findings for congruent movies of Task 1</i>	88
Table 17	<i>Summary of fixation duration findings for congruent and incongruent movies of Task 2</i>	89
Table 18	<i>Summary of initial fixation findings for congruent and incongruent movies of Task 2</i>	91

LIST OF FIGURES

<i>Figure 1.</i> Stimulus regions compared in fixation time analyses: (a) all defined regions,	37
(b) facial regions (shaded area) compared with non-facial regions (non-shaded),	
(c) core regions (shaded area) compared with periphery (non-shaded boxes on face),	
and (d) three core regions compared (eyes, nose, mouth).	
<i>Figure 2.</i> Illustration of expected three-way interaction between group, congruence,	39
and region for the core versus periphery analysis.	
<i>Figure 3.</i> Illustration of expected three-way interaction between group, congruence,	40
and region for the analysis between eyes, nose, and mouth.	
<i>Figure 4.</i> Schematic of ASL 6000 set-up and eye-tracking data collection	52
<i>Figure 5.</i> Total fixation duration for the face and non-face regions when viewing	63
congruent movies.	
<i>Figure 6.</i> Total fixation duration for the core and peripheral face regions when viewing	64
congruent movies.	
<i>Figure 7.</i> Total fixation duration for the eye, nose, and mouth regions when viewing.....	65
congruent movies.	
<i>Figure 8.</i> Total fixation duration for the face and non-face regions when viewing	69
all movies.	
<i>Figure 9.</i> Combined group data for total fixation duration to face and non-face regions	69
when viewing congruent and incongruent movies.	
<i>Figure 10.</i> Total fixation duration for the core and peripheral face regions when.....	71
viewing all movies.	
<i>Figure 11.</i> Combined group data for total fixation duration to the core and peripheral face.....	72

regions when viewing congruent and incongruent movies.

Figure 12. Total fixation duration for the eye, nose, and mouth regions when viewing.....74
congruent movies.

Figure 13. Total fixation duration for the eye, nose, and mouth regions when viewing.....74
congruent movies.

Figure 14. Number of first fixations devoted to the eye, nose, and mouth regions when.....76
viewing congruent movies.

Figure 15. Number of first fixations devoted to the eye, nose, and mouth regions when77
viewing incongruent movies.

1 INTRODUCTION

Affective perception occurs quickly and requires the seemingly automatic integration of multiple sources of information, including facial expression, semantic information, affective prosody, affective memory, and physiological responses. The accurate perception of facial affect is just one important component of affective perception, as facial expressions reveal a lot of information about one's state of mind, desires, and intentions. Researchers continue to explore the manner in which humans derive affective information from the facial expressions of others, but the process is not yet completely understood. In order to better understand the perception of affective information in individuals with autism spectrum disorders (ASD) and their typically developing (TD) peers, it is important to consider the underlying processes, including facial processing strategies and the integration of facial expression and affective prosody, as well as gaze behavior and the impact of biological motion on affect recognition. In accordance with this, research on emotion perception will be reviewed, taking into consideration what is known about the typical and atypical perception of visual and audio-visual information. Given that the emotion perception literature is fairly narrow and primarily focuses on perception of visual emotion cues, a broader perceptual literature provides a useful foundation from which predictions can be made about the strategies utilized during audio-visual emotion processing in ASD.

The data collected in this study provide information about the regions of the face (i.e., eyes, nose, mouth) that are sampled during a relatively more complex emotion perception task than those typically used in the literature. The relative increase in complexity of the stimuli to be used in the current study comes from the fact that stimuli provide both bimodal (auditory and visual) and dynamic (biological motion) cues. Results of this study will provide critical

information about different visual scanning approaches to the perception of dynamic audio-visual stimuli, and will increase understanding of emotion perception deficits in ASD. This knowledge will inform treatment intervention strategies that aim to improve social skills in ASD. Future directions will be discussed in the context of previous findings, as well as in terms of clinical implications for such research.

1.1 Autism Spectrum Disorders (ASD)

Autism spectrum disorders (ASD), or pervasive developmental disorders, are characterized by markedly abnormal or impaired development across three domains: social interaction, communication, and restricted interests or stereotyped/repetitive behaviors (American Psychiatric Association, 2000). Within the social domain, symptoms include limited social-emotional reciprocity, impaired use of nonverbal behaviors (e.g., facial expression, gestures, eye contact) to regulate social interactions, failure to develop age-appropriate peer relationships, and deficits in joint attention. Symptoms within the communication domain include language delay coupled with a failure to compensate with nonverbal communication, difficulty with reciprocal conversation, limited imitative or pretend play, and stereotyped or idiosyncratic use of language. The restricted, repetitive, and stereotyped patterns of behavior or interests are observed as intense preoccupations, rigidity of nonfunctional routines, stereotyped or repetitive motor movements, or preoccupations with parts of or sensory qualities of objects (American Psychiatric Association, 2000).

Abnormal functioning or a delay in the development of six behaviors across the three domains, with impairment in one of these areas (e.g., social interaction, social use of language, imitative play) by the age of three years, meets criteria for a diagnosis of Autistic Disorder. A diagnosis of Asperger's syndrome (AS) requires deficits in the social interaction and restricted

behaviors and interests domains, in the absence of language delay and cognitive impairment. Individuals who present with significant and pervasive impairments in one or more of these areas, but at a level that does not meet criteria for another specific ASD, may be diagnosed with Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS; American Psychiatric Association, 2000). In addition to these core symptoms, children with ASD may develop behavioral changes or comorbid disorders, such as problems with sleep, appetite, mood, anxiety, hyperactivity, irritability, anger regulation, and aggression, but these symptoms may be difficult to recognize due to the broader psychosocial deficits associated with the diagnosis of ASD (Ozonoff, Goodlin-Jones, & Solomon, 2005). Additional symptoms (e.g., guilt) that alert professionals to other psychiatric disorders (e.g., depression, anxiety) may not present in individuals with ASD (Ozonoff, Goodlin-Jones, & Solomon, 2005).

A recent epidemiological study conducted by the Centers for Disease Control and Prevention estimates the current prevalence rate for all ASDs at 1/88, representing a 23% increase in prevalence when compared to data collected in 2006 (Centers for Disease Control and Prevention, 2012). With regard to subtypes of ASD, one recent epidemiological study reports prevalence rates for Autistic Disorder at 22.0/10,000 and for AS around 10.5/10,000 (Saracino, Noseworthy, Steiman, Reisinger, & Fombonne, 2010). With the apparent rise in rates of ASD diagnosis, research that aims to explore early risk factors, identify methods for early detection, and develop early intervention strategies is becoming more widely conducted. Dawson (2008) describes a developmental model of such risk factors, risk processes, and outcomes in individuals diagnosed with ASD, highlighting the neurophysiological changes occurring with early experience and intervention and the potential for reduced symptomatology in ASD. With recent advances in intervention strategies, a number of research groups have reported

improvement in symptoms of ASD (using cognitive, adaptive, language, and/or general developmental measures) following early intervention (Howlin, 2005; Rogers & Vismara, 2008). Other reviews indicate significant improvement and treatment efficacy with early communication intervention (Brunner & Seung, 2009) and behavioral intervention (Howlin, Magiati, & Charman, 2009).

Social deficits are viewed as a primary symptom that is characteristic of all diagnoses within the autism spectrum (American Psychiatric Association, 2000). Individuals with ASD exhibit significant weaknesses in social-emotional reciprocity beginning early in development, with primary impairments in social orienting, joint attention, motor imitation, the processing of facial information, and the recognition of emotion from both visual and auditory cues (Dawson & Bernier, 2007; Loveland et al., 1997; Mundy, Sigman, Ungerer, & Sherman, 1986).

1.2 Development of emotion perception

Typical Development. Research shows that TD individuals are consistently able to identify basic emotions that are recognized across cultures (e.g., Ekman & Friesen, 1975). Researchers generally focus on six basic emotions when exploring the universality of emotion perception in cross-cultural studies: happiness, sadness, surprise, fear, disgust, and anger (e.g., Bryant & Barrett, 2008; Ekman, 1992; Izard, 1971, 1992; Wang et al., 2006). Basic emotions have been defined as those presumed to have innate neural substrates, a unique and universally recognized facial expression, and a unique feeling state (Izard, 1972, 1992; Ortony & Turner, 1990; Panksepp, 1982). In addition to early arguments regarding the existence of emotion- or system-specific biological correlates, evidence from recent neuroimaging and neurophysiological studies supports the existence of discrete, biologically-based emotion response mechanisms. For example, a recent review of neuroimaging studies examining neural responses to various

emotion experiences (e.g., mood induction, facial expression identification, perception of arousing pictures) found a unique pattern of neural activation for each of the six basic emotions; the most robust activation was found in the superior temporal gyrus for happiness, left medial frontal gyrus for sadness, left inferior frontal gyrus for anger, left amygdala for fear, and right insula and right inferior frontal gyrus for disgust. In addition, discrete patterns of autonomic nervous system (ANS) responding have been found for each of the six basic emotions, as indicated by electrodermal (i.e., skin-conductance, skin resistance, skin potential), thermovascular (i.e., skin blood flow, skin temperature), and respiratory measures (Collet, Vernet-Maury, Delhomme, & Dittmar, 1997; Kreibig, 2010).

The recognition of emotions requires understanding of nonverbal social cues, including facial expression, prosodic information or tone of voice, and body posture of others. The development of emotion perception skills during infancy and early childhood is observed in terms of the child's responsiveness to others' emotions and evidence of social referencing (Baldwin & Moses, 1996; Haviland & Lelwica, 1987). Research suggests that infants display a sensitivity to social cues (e.g., Grossman et al., 2008) and are able to discriminate between dynamic (i.e., in motion) displays of basic, prototypical emotions (i.e., happy, sad) by the age of six months old (de Haan & Nelson, 1997). Maternal expression of happy, fearful, or neutral emotions has been shown to influence an infant's behavior as early as 12 to 18 months of age. In situations designed to produce uncertainty in the child, maternal facial expression was found to influence the distance to which the infant would explore (Klinnert, 1984). By the age of four years, children are able to verbally label prototypical emotions (i.e., happy, sad, angry; Widen & Russell, 2003). Although the ability to recognize most prototypical emotional expressions is comparable to the adult level by ten years of age (Mondloch, Geldart, Maurer, & LeGrand,

2003), there continues to be improvement in the recognition of less intense emotions and increased sensitivity to more subtle displays of the basic emotions beyond adolescence (Thomas, DeBellis, Graham, & Labar, 2007). Researchers have found that emotion recognition proficiency, as defined by the speed of processing and the ability to detect subtle differences, continues to develop throughout adolescence and peaks in adulthood (Rump, Giovannelli, Minshew, & Strauss, 2009).

Atypical Development of Emotion Perception Abilities in Autism Spectrum Disorders.

Individuals with ASD encounter difficulties in basic visual processing of social situations (e.g., social use of eye gaze to interpret what others may be thinking, feeling, or communicating) and even young infants with ASD respond less to others, with evidence for the reduced salience of human faces (Klin, Jones, Schultz, & Volkmar, 2005; Osterling, Dawson, & Munson, 2002). However, the results of studies exploring emotion recognition abilities are not as definitive as those that highlight early deficits in social engagement. Some researchers assert that emotion recognition is intact in ASD, whereas others claim that it is impaired (see Capps, Yirmiya, & Sigman, 1992 and Celani, Battacchi, & Arcidiacono, 1999 for opposing findings). Equivocal findings can broadly be attributed to differences in methodology, task demands, and age and abilities of the participants being studied.

In general, individuals with ASD perform similarly to controls when they are presented with unambiguous, prototypical emotions and allowed sufficient processing time before a response is required (Capps, Yirmiya, & Sigman, 1992; Grossman, Klin, Carter, & Volkmar, 2000; Humphreys, Minshew, Leonard, & Behrmann, 2007; Ozonoff, Pennington, & Rogers, 1990). Weaknesses in emotion perception emerge when individuals with ASD are shown stimuli for brief periods of time or when emotional expressions are subtle (Critchley et al., 2000;

Greimel et al., 2010; Humphreys et al., 2007; Law Smith et al., 2010; Pelphrey, et al., 2002; Philip et al., 2010; Rump, Giavannelli, Minshew, & Strauss, 2009). In addition, children with ASD, regardless of specific diagnosis on the autism spectrum, perform worse than controls on emotion identification (Lindner & Rosen, 2006) and affect matching tasks (Boucher, Lewis, & Collis, 2000). However, some researchers have failed to find these deficits in emotion identification tasks with adolescents and adults with AS (e.g., Capps, Yirmiya, & Sigman, 1992; O'Connor, 2007; Rutherford & Towns, 2008).

A recent study attempted to reconcile these inconsistent findings by assessing emotion recognition abilities from a developmental perspective. Rump and colleagues (2009) found that children diagnosed with ASD, ages five to seven years, performed more poorly on an emotion recognition task than did their TD peers. When examining the developmental differences in the same study, the authors found that older children (ages 8-12 years) and adolescents (ages 13-17 years) with ASD performed similarly to controls; however, adults (ages 18-53 years) with ASD performed significantly worse than same-age controls. Results indicated that the adult controls showed significant improvement in terms of proficiency (i.e., processing speed, ability to detect subtle differences) over younger controls, but the ASD adults did not show the same advantage over older children and adolescents with ASD. As such, the researchers postulated, based on the data from this cross-sectional study, that the continued development of proficiency seen in typical development does not occur in ASD (Rump et al., 2009).

1.3 Impact of Emotion Perception Deficits on Social Functioning in ASD

Basic emotion recognition through nonverbal cues has been shown to relate to performance on brief measures of socio-cognitive abilities and social functioning and adjustment in school-aged children (Bonner, Hardy, Willard, Anthony, Hood, & Gururangan, 2008; Leppanen &

Hietanen, 2003). Inattention to important nonverbal social cues (e.g., faces, eyes) may have deleterious consequences for social functioning because of the roles that these cues play in signaling others' intentions. As such, ongoing deficits in age-appropriate affective recognition skills in children and adolescents with ASD can have adverse consequences that are additive and persist into adulthood. Studies with TD children provide evidence that peer rejection and victimization alters the way children process and respond to social cues by increasing sensitivity to hostile cues and increasing aggressive responding (Dodge, Lansford, Burks, Bates, Pettit, Fontaine, et al., 2003). For individuals with ASD, deficits in using nonverbal cues to make inferences about others' intentions and desires impact their social adjustment; some researchers have found that children with ASD respond to peer rejection with negative behaviors (e.g., isolation, oblivion, emotionally-charged accounts relayed to family, behavioral acting out) in the same way as typical children (Ames & Jarrold, 2007; Ochs et al., 2001).

Just as peer group factors (e.g., peer rejection) impact a child's social and academic adjustment, so do the resulting individual child factors (e.g., isolation, withdrawal; Ladd, 2006). Peer group size and peer group acceptance contribute to socio-emotional adjustment, academic competence, and self-concept in TD children (Vandell & Hembree, 1994). Despite reported involvement in peer networks, children with ASD report lower centrality, acceptance, companionship, and reciprocity (Chamberlain, Kasari, Rotheram-Fuller, 2007). Interestingly, some have shown that children with ASD report more loneliness than do their TD peers, albeit with a less complete understanding of the relationship between loneliness and friendship, whereas other report no difference in reports of loneliness between children with ASD and their TD peers (see Bauminger & Kasari, 2000 and Chamberlain, Kasari, & Rotheram-Fuller, 2007 for opposing findings). The impact of chronically negative social experiences reaches farther than

the child's peer group. Some researchers have shown that chronic peer rejection inhibits TD children's classroom participation (Ladd, Herald-Brown, Reiser, 1997). In addition, studies with TD children have shown that peer acceptance and friendship predicts school perceptions, avoidance, and performance (Ladd, 1990).

As a result of this pervasive deficit, a vast amount of research focuses on the social abnormalities of individuals diagnosed with ASD. However, Brunner and Seung (2009) provide a summary of recent studies and meta-analyses examining the efficacy of social skills interventions for children with ASD and conclude that adequate efficacy studies for social skills interventions (i.e., social skills training, social stories, and peer-mediated training) are as yet unpublished. As programs are developed and evaluated, effective social skill intervention strategies could benefit from more in depth exploration of specific deficits in emotion perception and recognition. These specific emotion perception deficits will be explored in greater detail in the following section.

1.4 Components of Emotion Perception

A poor understanding of atypical emotion perception by researchers and professionals, specifically with regard to children and adolescents with ASD, can result in inadequate treatment interventions. With complex and often disparate results from emotion perception studies in individuals with ASD, as discussed previously, researchers may benefit from looking more closely at the various components of emotion perception (e.g., audio-visual integration, face perception) in order to better understand deficits that do exist. Results from studies in the broader perceptual literature that examine these components more directly will inform hypotheses about emotion perception abilities in ASD on more complex, dynamic tasks and may provide a better

understanding of ways to improve the emotional and social functioning of children and adolescents with ASD.

Audio-visual integration. In everyday social situations, individuals are required to integrate information from both visual (i.e., facial expression) and auditory (i.e., tone of voice) modalities. Evidence shows that the integration of facial and vocal information occurs early in emotion perception, during the perceptual level of processing and not later during the response selection stage (Hietanen, Leppanen, Illi, & Surakka, 2004; Pourtois, Debatisse, Desplan, & de Gelder, 2002). When congruent visual and auditory cues are presented concurrently, this integration occurs, as measured by event related brain potentials (ERPs), as early as 110 ms after stimulus onset (Pourtois et al., 2000). In addition, congruent emotional information presented through multiple modalities has been shown to modulate the amplitude size and latency of two early ERP components (e.g., P200 and N300; Paulmann, Jessen, & Kotz, 2009). Specifically, researchers found a gradual decrease in amplitude of these ERPs with the addition of each modality (i.e., unimodal, bimodal, multimodal). Findings suggest that modulated amplitudes reflect a reduced and speeded processing effort for bimodal (facial expression and prosody) as compared to unimodal (facial expression) information.

Some researchers have found that one sensory modality has a greater influence on perceptual outcome than others. When presented simultaneously with visual and auditory stimuli and asked to respond by indicating whether they detected a visual stimulus, an auditory stimulus, or both, participants made a significant number of errors. Specifically, participants were prone to report visual only during trials in which both visual and auditory cues were present, termed the visual dominance effect (Collignon et al., 2008; Hecht & Reiner, 2008). In another study, this visual dominance effect emerged both when stimuli were semantically congruent and when they

were semantically incongruent (Koppen, Alsius, & Spence, 2008). In this study, participants made visual-only responses significantly more often than auditory-only responses on bimodal trials, regardless of the congruence of auditory and visual syllables. Similar support has been found for a visual dominance effect during emotion perception as well; a number of researchers have demonstrated that typically developing individuals showed a bias toward facial affective information when incongruent information was presented in another modality (e.g., tone of voice, semantic information; de Gelder & Vroomen, 2000; de Gelder, Vroomen, & Bertelson, 1998; Massaro, 1998; Massaro & Egan, 1996). This influence held when participants were instructed to attend to both modalities (Massaro & Egan, 1996) and when they were instructed to attend to prosody (deGelder & Vroomen, 2000; Vroomen, Driver, & deGelder, 2001). However, this influence did not hold when the face was presented upside-down. The processing of faces was disrupted when the face was inverted (discussed in the following section) and this finding highlighted the unique influence of facial expression on judgments of prosody (de Gelder, Vroomen, & Bertelson, 1998). The literature is equivocal on whether prosody has the same influence on perception of facial expression (see deGelder & Vroomen, 2000 and Massaro, 1998 for opposing findings).

Weaknesses in bimodal integration have been found in individuals diagnosed with ASD for non-emotional (Gepner, deGelder, & deSchonen, 1996; Mongillo et al., 2008) and emotional (Loveland et al., 1995) tasks. Children with ASD produced significantly fewer blends and fusions on an audio-visual speech task than did matched controls (Gepner, deGelder, & deSchonen, 1996). In a similar study, Mongillo and colleagues (2008) found that individuals with ASD show less visual-speech influence on their responses during a task with mismatching audio-visual syllables. Individuals with ASD also showed weaknesses in detecting affective

congruence between computer images and a simultaneously played auditory track (Loveland et al., 1995) or matching photographs with tape recordings containing affective prosody (Boucher, Lewis, & Collis, 2000). However, individuals with ASD are able to perform at levels comparable to that of TD controls on audio-visual emotion recognition tasks when cued by semantic content (Lindner & Rosen, 2006). Individuals with AS showed a bias toward visual-verbal information, not facial expression, suggesting that people with AS rely on verbal mediation as a compensatory strategy during complex emotion perception (Grossman et al., 2000).

In summary, studies of typically developing individuals show that visual information is dominant during visual-auditory stimuli (Collignon et al., 2008; Hecht & Reiner, 2008; Koppen, Alsius, & Spence, 2008) and visual affective information may have a greater influence on prosodic information than vice versa (de Gelder & Vroomen, 2000; Massaro, 1998; Massaro & Egan, 1996). However, given that this same dominance appears to be absent in individuals with ASD, future research should continue to explore the differences in visual perception strategies employed in ASD.

Face perception. Given that information from the visual modality appears to have a greater influence on audio-visual integration, exploration of facial processing strategies is bound to offer critical insight into typical and atypical bimodal emotion perception. Research indicates that a global or configural representation of faces was related to a greater recognition of facial emotion (Gross, 2005; Smith & Scott, 1998; White, 1999). A global or configural representation refers to one in which the perceiver utilizes information about the position of facial features, or configuration of the features, to a greater degree than information about the individual features of the face. To identify a particular face or facial expression, typically developing individuals generally rely on the spatial configuration of core or internal facial features (i.e., eyes, nose,

mouth), utilizing a holistic processing approach, as opposed to relying on individual features, which is associated with the processing of non-face objects (Pelphrey et al., 2002). Exploring these various approaches to face processing (i.e., configural, featural, holistic) in greater detail is important to furthering our knowledge of how and what information is extracted from the face during emotion processing tasks.

There is a large body of literature that is dedicated to the exploration of face processing strategies, particularly configural and featural strategies. Diamond and Carey (1986) are some of the researchers at the forefront of this literature and have attempted to explain the inversion effect, a manipulation frequently utilized in standard recognition memory paradigms. The inversion effect is essentially a difference in performance on recognition tasks between upright and inverted photographs of faces; the inversion effect occurs when faces that are presented upside-down during encoding and recognition are more poorly recognized than faces presented upright.

Diamond and Carey (1986) argued that the inversion effect emerges when two criteria are met: (1) members of a class to be encoded share a configuration and can be distinguished from one another based on second-order relational features and (2) perceivers are experts with a class of stimuli. Second-order relational features, to be discussed in detail in the next section, essentially refer to the spatial configuration of the individual features of a stimulus. With regard to the criterion of expertise, Diamond and Carey (1986) found that the inversion effect emerges when perceivers are as experienced with the class of stimuli presented for recognition as they are with human faces. They found that the recognition of dogs is as sensitive to the inversion effect as is the recognition of faces, provided that the perceivers are highly experienced with both dogs and human faces, such as is found in dog breeders. This argument was in contrast to an earlier

theory proposed by Yin (1969, 1970), who found that memory recognition for human faces is more vulnerable to stimulus inversion when compared to other classes of familiar objects (e.g., houses, airplanes, stick figures of people in motion) and argued that the inversion effect occurs because neural specialization has evolved such that humans possess superior recognition processing strategies for human faces. The first criterion proposed by Diamond and Carey (1986), regarding stimulus-class characteristics, provides insight into the underlying processes of facial and emotion processing and will be discussed in greater detail.

Distinguishing between first-order and second-order features is important to understanding the differential effects of inversion on recognition. Rhodes (1988) defined first-order features as the appearances of individual parts of the face, such as those labeled the eye, nose, and mouth. Second-order features were those said to have configural properties, including the spatial relations among first-order features, the position of first-order features, and information about the shape of the face (Rhodes, 1988). Rhodes' initial hypothesis of the mechanisms of face representation utilized similar terminology as that proposed by Diamond and Carey (1986), but the definitions appeared somewhat different. The distinction between the two hypotheses became less clear in a later paper written by Rhodes, Brake, and Atkinson (1993). As defined by Diamond and Carey (1986), first-order relational features refer to the relationship between parts of a stimulus that are relatively unconstrained (i.e., the trees in a landscape). When you have a class of stimuli that are homogenous in nature (i.e., human faces), the first-order relational features are relatively constrained (i.e., eyes are located horizontally above the nose, which is above the mouth; Diamond & Carey, 1986; Rhodes, Brake, & Atkinson, 1993). With stimulus classes that possess fixed first-order relational features, another relationship becomes crucial to the individuation of class members. The distinctive relations between the fixed features

of a stimulus (e.g., the distance between a specific individual's eyes) are referred to as the second-order relational features; this unique configuration is what is used to individuate faces (Diamond & Carey, 1986; Rhodes, Brake, & Atkinson, 1993).

There is a growing body of literature focused on contrasting the configural versus featural processing of objects, and in human faces in particular. A common assumption is that both information about isolated features (i.e., featural properties) and the relations among them (i.e., configural properties) are used as a means of encoding upright faces. The configural information, however, cannot be extracted from inverted faces and therefore results in less successful recognition (Valentine, 1988). Diamond and Carey (1986) argued that the large inversion effect that results from expertise and experience is due to individuals' reliance on the configural properties of stimuli (e.g., faces). Researchers have shown that the coding of second-order relational features is more affected by inversion than the coding of first-order relations or of isolated features (Rhodes, Brake, & Atkinson, 1993). Rhodes (1988) concluded that second-order features are critical to face perception and thus it is, at least in part, configural by nature.

In addition to the second-order relations, the properties of a stimulus configuration include the holistic, or gestalt, form. The overall structure or gestalt of human faces is of critical importance when considering the superior recognition or processing of faces as compared to other classes of objects and a number of theories have been proposed that attempt to explain how this information is represented. In addition to the theories of relational or configural information, others have proposed that gestalt form of the face is represented in parallel processing or through interactive coding (Bradshaw & Wallace, 1971; Diamond & Carey, 1986; Rhodes, 1988; Rhodes, Brake, & Atkinson, 1993; Sergent, 1984). Farah, Wilson, Drain, and Tanaka (1998) have attempted to explain how all of the parts of a face are processed, postulating that face

recognition does not involve, or does so only to a lesser extent, the explicit representations of the eyes, nose, and the mouth. They believe that faces are instead recognized as undifferentiated wholes. In support of this theory, Farah and colleagues (1998) found that individuals perceived upright faces in a holistic manner, significantly more so than inverted faces, houses, or words; participants had significant difficulty differentiating and comparing individual parts of the upright face independent of the whole face.

Others have found a similar whole-face advantage for upright faces. Homa, Haver, and Schwartz (1976) found that individuals recognized features more accurately when presented in the context of a natural face than if presented amongst scramble features. TD children recognized eyes and mouths in the context of the whole face, with eyes being recognized more efficiently than the mouth (Joseph & Tanaka, 2003). If these component features were isolated from the whole context of the face, recognition was more disrupted for upright faces than inverted or scrambled faces or of component features of houses (Tanaka & Farah, 1993). Gauthier and Tarr (1997) also examined recognition along this isolation-configuration continuum with their novel stimuli; parts were better recognized in the studied configuration than when presented in isolation, indicating that they were encoded holistically, and extensively trained viewers (i.e., experts) were more sensitive to configuration change than were novice viewers.

Researchers that have compared the face processing strategies of TD children to adults have found that configural processing strategies develop more slowly than feature-based processing strategies, with a shift from analytic to holistic processing occurring with typical development (Mondloch, Le Grand, & Maurer, 2002; Schwarzer, Huber, and Dummler, 2005). On a task requiring participants to judge whether two faces were the same or different, TD children made significantly more errors than adults in a spacing task in which the space between

eyes and nose differed (i.e., configural changes). In contrast, children were almost as accurate as adults in judgments of faces in which the external contour of the face changed and in judgments about faces with differing eye and nose shapes (i.e., featural changes).

With regard to the difference in face perception between TD individuals and individuals with ASD, individuals with ASD are slower and less accurate on face perception tasks (Schultz, 2005). In fact, deficits in the perception and recognition of complex social stimuli, such as faces, has been shown to be a specific impairment in individuals with ASD; whereas individuals with ASD are generally successful at recognizing nonsocial stimuli (i.e., simple objects, complex block patterns), they show impairments in the recognition of faces (Bradshaw, J., Shic, F., & Chawarska, K., 2011). In addition, individuals with ASD have consistently demonstrated a bias for and heavy reliance on feature-level analyses and are less prone to use configural or holistic information in the perception of faces, often evidenced by an abnormal or decreased inversion effect (Behrmann et al., 2006; Davies, Bishop, Manstead, & Tantam, 1994; Hobson et al., 1988; Joseph & Tanaka, 2003; Karatekin et al., 2007; Schultz, 2005). As previously discussed, individuals with little experience with a class of objects, as opposed to those who have attained expertise, have been shown to rely more on the featural aspects of the face. Behrmann and colleagues (2006) found that adults with ASD performed more slowly on a face discrimination task than controls, showed a strong preference for local components of stimuli, and that the slowed processing of both faces and objects correlated with the preference for local processing.

This slowed processing of global features is generally consistent with others' conclusion that some individuals with autism are able to develop configural representations of faces; this ability, however, appears to emerge in a qualitatively different way than TD individuals. Teunisse and de Gelder (2003) found that individuals with ASD performed similarly to TD

controls on a face inversion task under reduced memory demands and increased viewing times, but utilized contextual information in a face composite tasks (i.e., visual-search tasks). Similarly, Joseph and Tanaka (2003) found that children with ASD evidenced a typical whole-face recognition advantage, albeit in a different manner than the advantage found in TD children. The whole-face advantage only emerged for ASD children during trials in which feature-matching relied on mouth-recognition, whereas TD children demonstrated the whole-face advantage for eye-recognition and, to a lesser degree, mouth-recognition (Joseph & Tanaka, 2003).

1.5 Gaze behavior

Further exploration of physiological markers during emotion perception tasks can afford researchers a better understanding of the emotion perception abilities of individuals with ASD. Should group discrepancies in performance on emotion perception tasks occur because TD individuals and individuals with ASD approach these tasks in fundamentally different ways, direct measures of face processing and perception strategies can help to elucidate such differences. Noton and Stark (1971) stated that the most direct and real-time method of assessing processing strategies was to record visual scanpaths. The visual scanpath is defined as the pattern of eye movements that occurs when an individual processes a complex stimulus. Walker-Smith, Gale, and Findlay (1977) studied the eye movements of three participants to determine whether typical scanpaths emerged during a facial recognition task. Results showed that participants will normally adopt a regular and consistent scanning strategy when viewing faces for encoding and during recognition. Research examining gaze behavior and eye movements during viewing of faces regularly indicates a proportionally greater sampling of the internal region compared with the external region and point of regard, or fixation location, primarily to the core features (Luria

& Strauss, 1978; Macworth & Bruner, 1970; Manor et al., 1999; Mertens, Sigmeund, & Grusser, 1993; Stacey et al., 2005).

Records of eye movements have shown that foveal vision (i.e., central point of gaze) is reserved for elements containing essential information to the observer. In the case of face perception, most attention is paid to the eyes, nose, and mouth (Walker-Smith, Gale, & Findlay, 1977). Research with healthy individuals shows that participants first fixate on the eyes and mouth when they are provided with no task instructions (Bar-Haim, 2006; Groner, Walder, & Groner, 1984). Lansing and McConkie (1999) described the “Gaze Direction Assumption,” the theory that observers direct their gaze toward those parts of the stimulus from which visual information is being sought in order to carry out the task at hand. During emotion processing tasks, typical children and adults have a preference towards fixating on the eyes, mouth, and nose, regions argued to yield the greatest amount of information about the mental state of other individuals (Pelphrey et al., 2002; Walker-Smith, Gale, & Findlay, 1977). In the case of emotion perception, TD individuals looked significantly more towards the eye region than the mouth region when processing simple emotions (e.g., happy, surprise, angry, afraid). When presented with complex emotions (e.g., scheming, thoughtfulness, flirting), these individuals increased the overall time spent looking at the core regions of the face (both eyes and mouth combined). The researchers found a marginally significant effect of complexity, attributed to the weak statistical power, in that TD individuals greatly increased time spent looking to the eyes, with reduced time to the mouth region (Rutherford & Towns, 2008).

A number of researchers have shown that, in various psychiatric populations such as those with ASD, social phobia, and schizophrenia, the failure to attend to the important information revealed by the eye region is related to social difficulties (Horley, Williams,

Gonsalvez, & Gordon, 2003, 2004; Manor et al., 1999; Pelphrey et al., 2002). In a study in which participants were asked to identify the emotion portrayed by actors in static photographs, Pelphrey and colleagues (2002) found that the visual scan-path of typical individuals resembled an upside-down triangle between the eyes, the nose, and the mouth, whereas the visual scan-path of individuals with ASD resembled a random pattern that focused on only one or two features surrounding the face (i.e., an ear, a region of the hairline, or the chin). When individuals with ASD did focus on a core feature, they selectively attended to the mouth (Pelphrey et al., 2002). Other researchers have also found that, during emotion processing tasks, individuals with ASD depend more on information from the mouth region than from the eye region (Hernandez, 2009; Joseph & Tanaka, 2003; Spezio et al., 2007). Another group (Rutherford & Towns, 2008) found that individuals with ASD did fixate on the eye region during the presentation (2500 ms) of photographs displaying simple emotions at a rate similar to that of the group of TD individuals. However, differences in fixation pattern emerged in this study during the presentation of more complex emotions. As previously discussed, TD individuals increased fixations both eyes and mouth combined when presented with complex emotions, with a trend toward specific increase to the eye region. In contrast, the individuals with ASD decreased the time spent looking at the eyes and mouth combined, with a trend toward decreased fixations to the eye region (Rutherford & Towns, 2008).

Few researchers to date have utilized dynamic stimuli to track the visual-scanning patterns of individuals with ASD. Fixation patterns during the passive viewing of dynamic movies depicting emotionally-laden social interactions has been studied in toddlers, adolescents, and young adults (Jones, Carr, & Klin, 2008; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Norbury, Brock, Cragg, Einav, Griffiths, & Nation, 2009). In one study, adolescents with ASD

fixated less often on the eye region than did TD controls, with no group difference in the time spent fixating on the mouth region (Norbury et al., 2009). In another study, adolescents and young adults with ASD focused less on the eye region as compared to the TD controls and also fixated more on the mouth, body, and object regions (Klin et al., 2002). These findings were largely replicated in a group of toddlers with ASD, with a lesser percentage of fixations on the eye and more on the mouth region (Jones et al., 2008). Bal and colleagues (2010) attempted to examine visual strategies during an explicit emotion-labeling task using dynamic stimuli. When viewing stimuli that digitally morphed from neutral to emotionally expressive over a period of 15-33 seconds, individuals with ASD demonstrated similar fixation duration to the eyes as did TD individuals.

1.6 Static vs. Dynamic stimuli

Researchers have posited that dynamic emotion stimuli, as compared to static stimuli, improve the perception of facial expressions by providing critical temporal information, particularly in the case of subtle emotional expressions (Ambadar, Schooler, & Cohn, 2005; Hess & Kleck, 1995; Lander & Chuang, 2005; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Sato et al., 2004). Neuroimaging studies have shown that the perception of dynamic facial stimuli activates a more widespread and enhanced neural network when compared to the perception of static stimuli (Arsalidou, Morris, & Taylor, 2001; Trautmann, Fehr, & Herrmann, 2009).

In addition, dynamic audio-visual stimuli more realistically simulate the emotion decisions that are required during social interactions, both because they provide biological motion and require rapid processing of changing facial configurations believed to signal unique emotions (Aviezer et al., 2008; Kohler et al., 2004; Nusseck, Cunningham, Wallraven, & Bulthoff, 2008; Smith, Cottrell, Gosselin, & Shyns, 2005). For example, previous research

indicates that a raised upper eye-lid and raised outer-eyebrows are characteristic and uniquely qualifying features of fearful expressions (Kohler et al., 2004), the recognition of a happy facial expression depends primarily on the motion of the mouth and its surrounding regions (i.e., upper lip, cheeks; Kohler et al., 2004; Nusseck, Cunningham, Wallraven, & Bulthoff, 2008), and the recognition of anger is facilitated by furrowed or lowered eyebrows, wrinkled nose, and a depressed lower lip (Kohler et al., 2004). Research conducted with children with ASD shows that they do not preferentially attend to biological motion, but instead attended to phase-scrambled (i.e., motion trajectories of point-light stimuli played temporally out of sequence) and biological motion (i.e., point-light display of walking human figure) for similar amounts of time (Annaz et al., 2012). More importantly, the children with ASD in this study preferentially attended to object motion (i.e., point-light display of a spinning top) over biological motion stimuli.

The aforementioned evidence provides strong support for the importance of assessing emotion perception abilities with bimodal, audio-visual (i.e., face, voice) and dynamic (i.e., movies) stimuli. However, previous studies have generally utilized static photographs when attempting to determine which regions of the face provide the most information for individuals under conditions requiring judgments of emotion (e.g., Dalton et al., 2005; Falkmer, Bjallmark, Larsson, & Falkmer, 2011; Hernandez et al., 2009; Kirchner, Hatri, Heekeren, & Dziobek, 2011; Pelphrey et al., 2002; Philip et al., 2010; Rutherford & Towns, 2008; Sawyer, Williamson, & Young, 2012). In studies that have attempted to provide bimodal emotional information, printed or audio-taped recordings of sentences with emotional content or spoken in various tones of voice have most often been used with static photographs of faces (de Gelder & Vroomen, 2000; Dziobek, Bahnemann, Convit, & Heekeren, 2010; Grossman, Klin, Carter, & Volkmar, 2000;

Hietanen, Leppanen, Illi, & Surakka, 2004; Lindner & Rosen, 2006; Massaro, 1998; Massaro & Egan, 1996; Pell, 2005), which still lacks the dynamic quality in the visual stimulus. Some researchers (see previous section on gaze behavior in ASD) have examined passive viewing of dynamic movies depicting emotionally laden scenes (Jones et al., 2008; Norbury et al., 2009), explicit emotion labeling of dynamic (i.e., morphing) visual-only video clips (Bal et al., 2010), and passive viewing an examiner's face during an interactive conversation with that examiner (Falkmer, Bjallmark, Larsson, & Falkmer, 2011). However, few studies to date have used sets of brief, dynamic and audio-visual stimuli to explore the fixation patterns of individuals with ASD during explicit emotion identification tasks.

In accordance with the reviewed literature, the accurate perception of emotions during social interactions arguably requires the successful integration of several components (e.g., face perception, audio-visual information, biological motion). Understanding each of these discrete processes and their integration provides a framework for understanding the complex skill of emotion perception and deficits found therein. Similarly, emotion perception and, more broadly, social functioning involves a complex interplay of various neural regions. Therefore, the role of each of these discrete neural regions in emotion perception, and the proposed connectivity among these regions, is briefly discussed in the following sections.

1.7 Neural Correlates of Social Perception

A network of brain structures that has been described as central to social functioning includes the superior temporal sulcus (STS), fusiform gyrus (FG), amygdala, and prefrontal cortex (PFC; Zilbovicius et al., 2006). The presumed function of each of these regions will be discussed briefly in terms of the available literature with TD individuals, followed by an introduction to research findings with individuals diagnosed with ASD. Differences in the

structure and function of these specific “social brain” regions may help to explain atypical performances on emotion perception tasks by ASD individuals found in previous, and predicted in the current, studies. Given that few papers have used dynamic and audio-visual stimuli to study neural involvement in emotion perception, the broader neural literature is discussed here.

The STS plays a critical role in the perception of biological motion, including eye movements and other changeable aspects of the face, the analysis of visual and social information conveyed by gaze shifts, and in the identification of emotion (Haxby et al., 2000; Itier & Batty, 2009; Pelphrey, Adolphs, & Morris, 2004). The STS region, along with the superior temporal gyrus (STG), has also been implicated in the integration of visual and auditory perceptual cues and in emotion perception (Allison, Puce, & McCarthy, 2000; Ethofer et al., 2006; Robins, Hunyadi, & Schultz, 2009).

Researchers have identified a specific region of interest, known as the “fusiform face area” (FFA), within the lateral aspect of the middle-FG that responds significantly more strongly during the passive viewing of a human face than to other objects (Bukach, Gauthier, & Tarr, 2006; Gauthier et al., 1999; Kanwisher, McDermott, & Chun, 1997; Kanwisher & Moscovitch, 2000; Kanwisher & Yovel, 2006; Tarr & Gauthier, 2000; Tong et al. 2000). Specificity of this region is for the perceptual identification of face (e.g., face recognition or discrimination); although the FFA is engaged during emotion perception tasks, as indicated by greater activation for emotional facial expressions as compared to neutral expressions, this region is not required for the identification of emotion (Schultz, 2005; Vuilleumier, 2002, 2004). Whether the increased activation in this region results from the unique specialization of the FFA for face perception or from the extensive amount of experience we have with this class of stimuli (i.e., faces), is yet to be answered and is the subject of ongoing debate.

The amygdala has dense reciprocal connections with the ventral visual stream (e.g., “what pathway”), plays a critical role in the early stages of facial expression processing, and also modulates the FFA (Schultz, 2005). More specifically, this subcortical structure is important for face processing, emotion identification, perspective taking, social judgments, and detection of threat (Adolphs, Baron-Cohen, & Tranel, 2002; Adolphs, Tranel, & Damasio, 1998; Carter & Pelphrey, 2006; Grelotti et al., 2002; Whalen et al., 2004). Some of the most direct evidence for the role of the amygdala in social cognition comes from functional imaging studies with patients with amygdala lesions. Previous research has revealed profound impairments in emotion recognition, particularly with recognition and processing of the intensity of fearful and other negative emotions, and other social judgments, such as evaluating the trustworthiness and approachability of others (for a review, see Pelphrey, Adolphs, & Morris, 2004). Additional evidence suggests that amygdala activation may be mediated by the amount of attention paid to the eye region (Dalton et al., 2005).

The PFC is a region that is highly connected with limbic structures, including the amygdala and hippocampus (Price, 2006). The PFC, specifically the ventromedial PFC, including the orbitofrontal cortex and the ventral part of the anterior cingulate cortex, has been implicated in motivation, reward, and planning. In addition, increased activation in the medial PFC has been found during tasks designed to elicit empathy, theory of mind, and emotion processing (for a review, see Neuhaus, Beauchaine, & Bernier, 2010).

Some research has suggested that there may be a hierarchy of processing for socially-relevant information, wherein the subcortical structures (i.e., amygdala, ventral striatum/ventral visual stream) respond quickly and signal other regions to the emotional salience of stimuli and the cortical regions (i.e., STS and FG) are involved in the advanced perceptual processing of

social information (Adolphs, 2003; Pelphrey, Adolphs, & Morris, 2004; Schultz, 2005). It has been suggested that the functions of these primary regions (i.e., amygdala, FG, and STS) work in parallel and rapidly integrate information during typical social processing (Pelphrey, Adolphs, & Morris, 2004).

Schultz (2005) suggests that an early, congenital abnormality in the amygdala could decrease attention to faces that occurs very early in infancy. This would then carry over throughout childhood and reduce the “expertise” for faces in the FFA, leading to the social dysfunction in ASD. In a similar argument, Dziobek and colleagues (2010) argued that deficits in social perception are precursors to the structural changes in and resulting abnormal activity of the FG. Results of their study showed a cortical thickening of the FG that was negatively correlated with emotion perception skills and a reduced covariation between amygdala volume and cortical thickening. Given that typical development includes increasing covariance between these two regions with maturation, the authors concluded that the reduced structural connectivity found in ASD resulted from underlying pathology (Dziobek et al., 2010). Alternate hypothesized causes of abnormal connectivity include, but are not limited to, pre- and postnatal neuronal and axonal development (Geschwind & Levitt, 2007). Geschwind and Levitt (2007) describe a model of ASD in which aberrant axonal development in higher-order brain regions (e.g., frontal lobe) results in a “developmental disconnection syndrome”. The authors argue that, when considered with findings from neuroimaging studies, recent genetic findings in support of this theory provide a potential explanation for the developmental course and heterogeneity of ASD.

Research exploring dysfunctional connectivity in ASD is becoming increasingly undertaken and is broadly motivated and supported by findings of dysfunction in individual regions of the “social brain” in individuals with ASD. As previously mentioned, the amygdala

and the FG have been widely implicated in the social deficits of ASD. A number of researchers have found both structural and functional abnormalities in the amygdala, including atypical patterns of neuronal development, volume abnormalities, and reduced connectivity with other regions (Conturo et al., 2008; Critchley et al., 2000; Schultz, 2005; Schumann & Amaral, 2000). Smaller amygdala volume in adults with ASD has been linked to fewer fixations on the eye region (Munson et al., 2006). Similarly, imaging studies have revealed both structural and functional abnormalities within the FG and, more specifically, the FFA, with reduced activation during emotion processing tasks and decreased connectivity with frontal regions, amygdala, and thalamus, amongst others, when viewing faces (Kleinhans et al., 2008; Koshino et al., 2000; Neuhaus, Beauchain, & Bernier, 2010; Pelphrey, Morris, McCarthy, & LaBar, 2007; Schultz et al., 2000). It has been suggested that the decreased activation in the FFA is modulated by amount of time spent fixating on the eyes (Dalton et al., 2005).

Abnormalities in the STS have also been strongly implicated ASD, including abnormal grey matter volumes, cortical thinning, changes in cerebral blood flow (for a review, see Zilbovicius, Meresse, Chabane, Brunelle, Samson, & Boddaert, 2006). It has been suggested that ASD individuals' early failure of joint attention and impaired ability to interpret the mental state and intentions of others from eye gaze (e.g., Baron-Cohen et al., 1999) and emotional facial expressions (Knutson, 1996) may result from abnormal STS functioning, given its importance in social perception, eye-gaze processing, and perception of dynamic facial expressions (Allison, Puce, & McCarthy, 2000; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Pelphrey, Adolphs, & Morris, 2004). Additional research shows reduced or altered activation in the PFC during social processing tasks. For example, Hall, Szechtman, and Nahmias (2003) found that individuals with

ASD had an increased activation when matching facial and vocal affect despite a greater number of errors.

Findings from neurodevelopmental studies indicate that the amygdala (e.g., Schumann et al., 2004) and the FG (Aylward et al., 2005) continue to develop through late childhood and adolescence. In his review, Schultz (2005) described a study conducted by his research group that revealed an abnormal growth process of the FG that continues into adulthood for individuals with ASD. Results indicated that there is bilateral enlargement of the FG above and beyond that explained by whole-brain enlargement for a group of 15 and older adults only, but not in those younger than 15 years (Schultz, 2005). Taken together, the previously discussed evidence for the typical involvement in the amygdala-FG system in social and emotional perception, the continued development of these regions into adolescence, and the increase in efficiency (e.g., speed, subtlety) of emotion detection into adulthood highlights the need to examine emotion perception difference between young adults with ASD.

Researcher have more recently employed neuroimaging techniques to expand our understanding of broader theories of social functioning in ASD, such as theories of reduced social motivation and reward system dysfunction. Neuroimaging studies have found hypoactivation across a network of brain regions involved in reward system for both monetary and social rewards (i.e., mesocorticolimbic circuitry, midbrain, thalamus, amygdala, striatum, ACC; Kohls et al., 2012). In addition, a recent review highlighted disruptions in the orbitofrontal-striatal-amygdala circuitry in response to social stimuli or social judgments and in the dysregulation of neurochemicals (e.g., oxytocin, glutamate) in social reward (Chevallier et al., 2012).

1.8 Preliminary Findings

A preliminary study was conducted in order to explore the visual scanning strategies of a group of TD young adults during the perception of dynamic audio-visual emotion (DAVE) stimuli to be used in the current study (McManus, Robins, Tone, & Washburn, Unpublished Thesis). These stimuli were digitally altered to allow for the presentation of both congruent and incongruent audio-visual affect and will be discussed in more detailed in the method section. Fixation duration was consistently greater for the core versus peripheral regions and for the eye region as compared to the other regions of interest. This preference for the eye region varied somewhat by emotion, with a significant fixation time devoted to the nose in some cases (e.g., neutral; McManus et al., Unpublished Thesis).

Behavioral responses from this preliminary study also indicated that participants were consistently able to identify angry, fearful, happy, and neutral expressions. Furthermore, the increased task difficulty of the incongruent conditions was evidenced by the increase in response times, wherein subjects showed a bias toward the visual affective information when asked to determine the emotion portrayed (McManus et al., Unpublished Thesis). Other preliminary studies have compared the behavioral performance (i.e., accuracy, modality bias, response time) of individuals with ASD to the performance of TD individuals on emotion perception tasks utilizing DAVE stimuli. Both TD and ASD groups made significantly more correct than incorrect responses when responding to congruent movies. When presented with incongruent movies, TD individuals showed a strong bias towards the facial expression, whereas individuals with ASD did not demonstrate such bias, instead responding comparably toward facial expression and prosody (McManus et al., 2008). With regard to response times, results have shown that both TD and ASD groups responded more quickly during congruent movies than

incongruent movies. In addition, findings indicated that response times to provide an emotion label for DAVE stimuli were significantly greater for individuals with ASD than TD individuals (Banks et al., 2008).

1.9 The Current Study

Previous research suggests that individuals with ASD may be able to accurately identify, at a level comparable to their TD peers, simple emotions presented as static visual stimuli and under extended viewing conditions (Capps, Yirmiya, & Sigman, 1992; Grossman, Klin, Carter, & Volkmar, 2000; Humphreys, Minshew, Leonard, & Behrmann, 2007; Ozonoff, Pennington, & Rogers, 1990). However, research also suggests that individuals with ASD evidence weaknesses in the integration of audio-visual integration (Boucher, Lewis, & Collis, 2000; Gepner, deGelder, & deSchonen, 1996; Loveland et al., 1995; Mongillo et al., 2008), show no dominance effect for visual information (Mongillo et al., 2008), show reduced salience of important affective information provided by the eye region (Hernandez, 2009; Jones et al., 2008; Joseph and Tanaka, 2003; Klin et al., 2002; Norbury et al., 2009; Pelphrey et al., 2002; Rutherford & Towns, 2008), and do not demonstrate the same level of proficiency as TD individuals on emotion perception tasks, particularly as they reach adulthood (Humphreys et al., 2007; Pelphrey, et al., 2002; Rump, Giavannelli, Minshew, & Strauss, 2009). Because the combined weaknesses in emotion perception have lasting effects on peer and interpersonal relationships (Bonner, Hardy, Willard, Anthony, Hood, & Gururangan, 2008; Leppanen & Hietanen, 2003), can potentially lead to affective disorders (Vandell & Hembree, 1994), and impede academic performance (Ladd, Herald-Brown, Reiser, 1997), it is imperative that researchers aim to explicate the factors contributing to such deficits and provide appropriate interventions at an early age.

Based on the reviewed literature, it appears that a gap exists in the direct comparison of physiological measures of face processing strategies during brief, dynamic emotion perception stimuli between individuals with ASD and TD individuals. Although dynamic audio-visual stimuli more realistically simulate the emotion decisions required during social interactions, both because they provide biological motion and critical temporal information that signals unique emotions (Ambadar, Schooler, & Cohn, 2005; Aviezer et al., 2008; Hess & Kleck, 1995; Kohler et al., 2004; Nusseck, Cunningham, Wallraven, & Bulthoff, 2008; Smith, Cottrell, Gosselin, & Shyns, 2005), previous studies have generally utilized static photographs in emotion identification tasks (e.g., Dalton et al., 2005; Hernandez et al., 2009; Pelphrey et al., 2002; Rutherford & Towns, 2008).

The current study addressed this gap in the literature by comparing the gaze behavior of individuals with ASD and TD individuals when viewing dynamic emotional movies and making emotion judgments about the actor in each video. Such research extended findings from studies that utilized static photographs to dynamic audio-visual emotion stimuli, which more realistically simulate emotions that are presented during real-world social interactions (Table 1). Gaze behavior measured using eye-tracking technology provided information as to the regions of the face sampled during emotion perception tasks. The current study included two tasks: the first task contained only congruent movies (e.g., happy face, happy voice) and the second task contained both congruent and incongruent movies (e.g., happy face, angry voice). The first task always preceded the second in presentation to allow for the exploration of face processing strategies during a congruent, dynamic emotion perception task without being influenced by the presence of incongruent movies. The second task provided information about perceptual changes that occur when some stimuli present conflicting auditory information. It was expected that

adolescents with ASD would evidence different processing under relatively more complex conditions (i.e., dynamic, brief presentation). As assessment of emotion perception under dynamic conditions more closely simulates real-life social interaction, this knowledge was expected to inform strategies which can be used to assess and treat social skills deficits in individuals with ASD. For example, social skills intervention programs designed for children and adolescents with ASD will benefit from increased knowledge about the type of explicit instructions that may improve the ability to extract information about the intentions of others.

Table 1
Summary of sample and task characteristics of relevant emotion perception studies.

Study	Participants		Experimental Procedure			
	Diagnoses (Mean Age)	Stimulus Type	Task	Presentation Time	Response Period	Eye- Tracking
Capps, Yirmiya, & Sigman (1992)	HF ASD (12.5) TD (12.0)	Static, Visual only	Labeling simple and complex emotions	Unlimited	Unlimited	No
Grossman, Klin, Carter, & Volkmar (2000)	AS (11.8) TD (11.5)	Static, Visual only	Forced-choice emotion recognition, face only and face with matching, mismatching, or irrelevant label	Unlimited	10 seconds	No
Hernandez et al. (2009)	Autism (24.09) TD (22.7)	Static, Visual only	Passive viewing of emotionally expressive faces (human and virtual faces)	4 seconds	N/A	Yes
Humphreys, Minshew, Leonard, & Behrmann (2007)	HF ASD (24) TD (28)	Static, Visual only	Forced-choice emotion identification, continuum of morphed facial expressions	Unlimited	Unlimited	No
Jones, Carr, & Klin (2008)	ASD (2.28) TD (2.03) DD (2.06)	Dynamic, Audio- visual	Passive viewing of actress playing the role of caregiver	10-38 seconds per clip	N/A	Yes
Klin et al. (2002)	AD (15.4) TD (17.9)	Dynamic, Audio- visual	Passive viewing of content-rich social interaction during a movie scene	30-60 seconds per clip	N/A	Yes
Ozonoff, Pennington, & Rogers (1990)	LF ASD (6.4) VMA TD (3.0) NMA TD (4.1)	Static, Visual only	Sorting by identity/emotion, Matching by emotion	Unlimited	Unlimited	No

Table 1 (cont'd)

Summary of sample and task characteristics of relevant emotion perception studies.

Study	Participants		Experimental Procedure			
	Diagnoses (Mean Age)	Stimulus Type	Task	Presentation Time	Response Period	Eye- Tracking
Pelphrey et al. (2002)	HF ASD (25.2) TD (28.2)	Static, Visual only	Passive viewing of faces, Forced-choice emotion recognition task	2.0 seconds	5 seconds	Yes
Rump, Giavannelli, Minshew, & Strauss (2009)	HF ASD (6.4) TD (6.0)	Dynamic, Visual-only	Matching by emotion, Progression from neutral to emotional expression (4 levels of intensity)	0.5 seconds	Unknown	No
Rutherford & Towns (2008)	HF ASD (25.6) TD (25.7)	Static, Visual only	Forced-choice emotion identification, simple and complex emotions	2.5 seconds	Unlimited	Yes

* Mean age in years

ASD = Autism Spectrum Disorder

AD = Autistic Disorder

AS = Asperger's syndrome

DD = Developmentally delayed, non-autism

HF = High-functioning

LF = Low-functioning

LI = Language-impaired

LN = Language normal

VMA = Verbal mental age-matched

NMA = Nonverbal mental age-matched

1.10 Hypotheses

Hypothesis 1 – Accuracy and Reaction Time during Congruent Movies. When identifying the emotions portrayed in congruent movies, it was predicted that both TD and ASD groups would respond accurately more often than they would respond inaccurately. There were no expected differences between TD and ASD groups in accuracy for the congruent movies (e.g., Grossman, Klin, Carter, & Volkmar, 2000; Humphreys, Minshew, Leonard, & Behrmann, 2007). However, it was predicted that individuals with ASD would respond more slowly than the TD individuals (e.g., Banks et al., 2008; Rump, Giavannelli, Minshew, & Strauss, 2009).

Hypothesis 2 – Fixation Pattern during Congruent Movies. Fixation patterns during congruent movies (Task 1) were examined across increasingly specific stimulus regions. First, face versus non-face fixations were compared across groups. Within face fixations, core versus periphery were compared across group, and finally, within core fixations, eyes versus nose versus mouth were compared across group. Please see Figure 1 for an example of how boxes were placed on the stimulus in order to calculate fixations. Each analysis involved comparing a subset of boxes to another subset, defined in Figures 1b-1d.

Hypothesis 2a. Face versus Non-Face. Previous research indicates that the visual modality (i.e., face) plays a critical role in the accurate identification of emotion (e.g., Collignon et al., 2008). As such, within-group analyses conducted at the broadest level were expected to show that both groups fixated for longer on facial regions than non-face regions. However, a Group x Region interaction was predicted for fixation duration to the facial regions. When viewing congruent movies, it was predicted that the overall fixation duration for facial regions (core + periphery) would be greater for TD individuals than ASD individuals, whereas individuals with ASD were not expected to show as strong of a preference for the facial regions.

Hypothesis 2b. Core versus Periphery. Among the facial regions, it was predicted that TD individuals would fixate for longer periods of time on core regions of the face (i.e., eyes, nose, mouth) than the peripheral regions (Luria & Strauss, 1978; Pelphrey et al., 2002; Figure 1c). In contrast to the prediction for the TD group, the individuals with ASD were expected to spend more time fixating on the peripheral regions (e.g., chin, forehead, cheeks) of the face than the core regions of the face (i.e., Group x Region interaction).

Hypothesis 2c. Eyes versus Nose versus Mouth. Within the core facial regions (eyes, nose, mouth), the eyes were expected to reveal the greatest amount of information about the affective state of another individual. Previous research has shown that TD individuals look significantly more towards the eye region than the mouth region when processing simple emotions (e.g., Rutherford & Towns, 2008). In contrast, previous research has shown that individuals with ASD do not attend to the critical facial regions (i.e., eyes) at a level comparable to that of TD individuals (e.g., Norbury et al., 2009; Pelphrey et al., 2002). When viewing congruent movies, it was therefore predicted that the total fixation duration for TD participants would be greater in the eye region than in any other core regions of the face (Figure 1d). In contrast, it was predicted that individuals with ASD would not show the same preference for the eye region; of the time spent fixating on the core regions of the face, they were expected to spend significantly less time fixating on the eye region (i.e., Group x Region interaction).

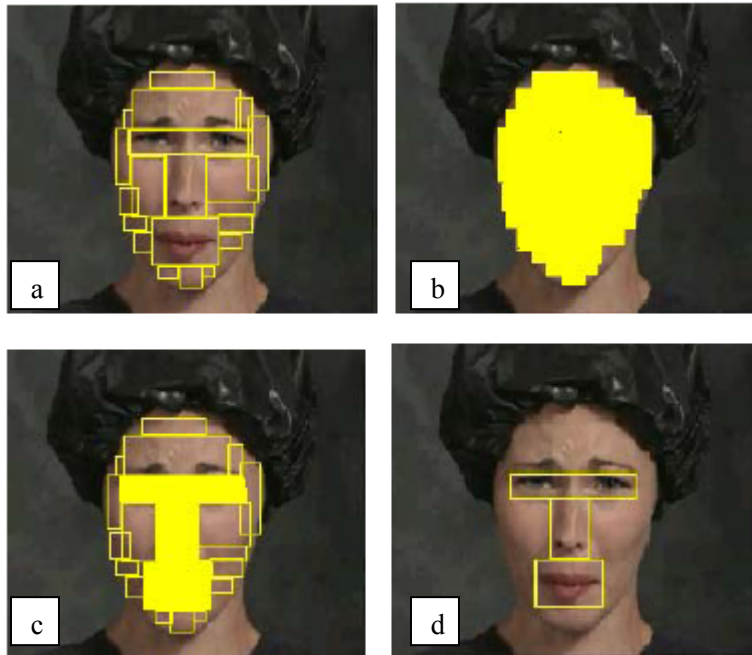


Figure 1. Stimulus regions compared in fixation time analyses: (a) all defined regions, (b) facial regions (shaded area) compared with non-facial regions (non-shaded), (c) core regions (shaded) compared with periphery (non-shaded boxes on face), and (d) three core regions compared (eyes, nose, mouth).

Hypothesis 3 –Behavioral Responses and Reaction Time during Incongruent Movies.

When identifying the emotions portrayed in incongruent movies, it was predicted that TD participants would be significantly more likely to choose labels consistent with facial expressions than with vocal prosody (de Gelder & Vroomen, 2000; Hecht & Reiner, 2009; Massaro, 1998; Massaro & Egan, 1996). In contrast, it was hypothesized that individuals with ASD would not show the same bias toward facial information and would choose labels consistent with facial expression and prosody at similar rates (i.e., Group x Modality interaction; McManus et al., 2008; Mongillo et al., 2008). Response times to label emotions were expected to be greater during the incongruent movie condition than during the congruent movie condition for both TD and ASD groups (Banks et al., 2008; Hietanen, Leppanen, Illi, & Surakka, 2004; Massaro & Egan, 1996; McManus et al., Unpublished Thesis; Pell, 2005). As previously mentioned, studies

have demonstrated greater response times for individuals with ASD during emotion perception tasks. As a result, the increase in response time for the incongruent movies was expected to be most pronounced for the individuals with ASD, demonstrating a significant interaction between group and congruence.

Hypothesis 4 – Fixation Pattern during Incongruent Movies as compared to Congruent Movies. As previously described for congruent movies (Hypothesis 2), fixation patterns examined across increasingly specific stimulus regions. These fixation patterns were then compared between congruent and incongruent movies (Task 2). Because of the increase in relative complexity of incongruent movies as compared to congruent movies, it was expected that TD individuals would increase fixations to the most salient facial feature for emotion judgments during incongruent movies, which was expected to be the eye region (Buchan et al., 2007; Lansing & McConkie, 1999; Rutherford & Towns, 2008). However, previous studies have shown that weaknesses in emotion perception and differences in fixation patterns emerge for individuals with ASD during more complex tasks (e.g., Critchley et al., 2000; Greimel et al., 2010; Philip et al., 2010; Rump, Giavanelli, Minshew, & Strauss, 2009). As a result, individuals with ASD were not expected to show the same increase to these most salient regions.

Hypothesis 4a. Face versus Non-Face. It was hypothesized that the group differences in pattern of fixations for facial (core + periphery) versus non-facial regions would be comparable during both the congruent and incongruent movies, in that the overall fixation duration for facial regions would be greater for TD individuals than ASD individuals (Figure 1b). This hypothesis would be supported by a two-way, Group x Region interaction.

Hypothesis 4b. Core versus Periphery. A three-way interaction between group, congruence, and region was predicted to emerge (Figure 2) for the fixation time to the core

regions of the face as compared to the peripheral regions of the face (Figure 1c). Specifically, it was predicted that TD participants would increase fixation time to the core regions of the face during the incongruent movies as compared to congruent movies whereas the ASD group would decrease fixation time to the core regions of the face during the incongruent movies as compared to congruent movies (i.e., Group x Region x Congruence interaction). This decrease in fixation duration to the core facial regions for the ASD group was expected to result from the decreased overall salience of the facial modality during this complex emotion perception task (e.g., McManus et al., 2008; Mongillo et al., 2008).

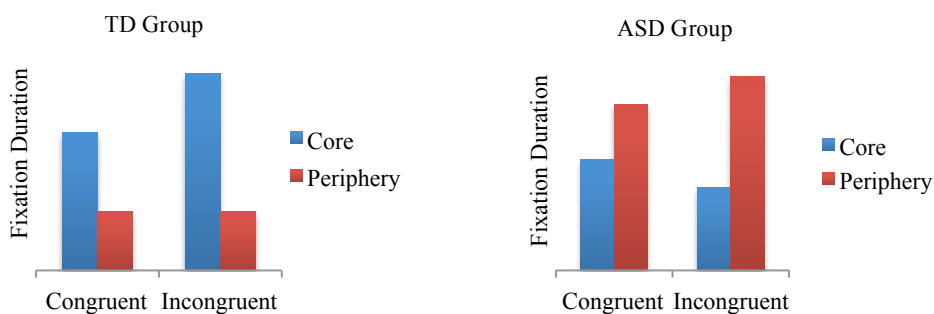


Figure 2. Illustration of expected three-way interaction between group, congruence, and region for the core versus periphery analysis.

Hypothesis 4c. Eyes versus Nose versus Mouth. With regard to group differences in fixation duration amongst the three core facial regions (eyes, nose, mouth; Figure 1d) for the congruent as compared to the incongruent movies, a three-way interaction was predicted (Group x Region x Congruence). It was predicted that TD participants would again fixate for longer on the eye region than other core regions of the face for incongruent versus congruent movies, with an expected increase in fixation time to the eyes during the incongruent as compared to the congruent movies. In contrast, it was hypothesized that individuals with ASD would not show a preferential increase in fixation duration to the eye region, for incongruent versus congruent

movies (e.g., Rutherford & Towns, 2008). The participants with ASD were instead expected to show a decrease in their fixation time to the eye region during the incongruent as compared to the congruent movies. For both TD and ASD groups, it was predicted that fixation time to the mouth region will remain comparable during congruent versus incongruent movies (McManus, Robins, Tone, & Washburn, Unpublished Thesis; Rutherford & Towns, 2008).

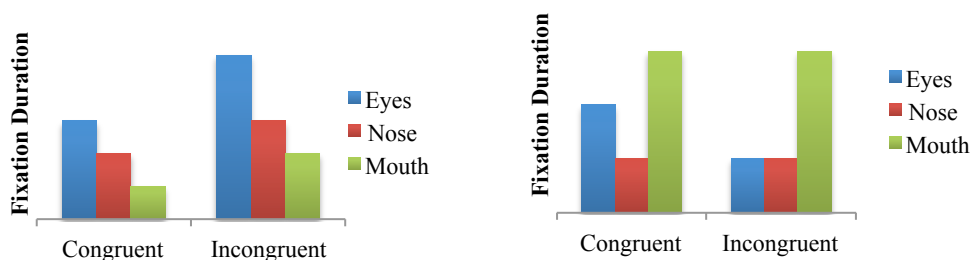


Figure 3. Illustration of expected three-way interaction between group, congruence, and region for the analysis between eyes, nose, and mouth.

Hypothesis 5 – Exploratory Analysis of Initial Fixations. As some research suggests that initial fixations differ for TD and ASD groups, exploratory analyses examined fixation sequences for both groups in congruent and incongruent movie conditions. It was predicted that TD individuals would first fixate on the eyes more often than ASD individuals, whereas ASD would more often initially fixate on the mouth region (i.e., Group x Region interaction).

Hypothesis 6 – Exploratory Analysis of the Relationship between Visual Fixations and Social Functioning. It was proposed that the visual scanning of faces during an emotion identification task would serve as an indirect measure of social functioning. As such, exploratory analyses examined the relationship between measures of visual scanning (i.e., fixation duration to face, core, and eye regions) and measures of social functioning (e.g., Autism-Spectrum Quotient, Autism Diagnostic Observation Schedule). It was predicted that scores on social functioning measures, with higher scores indicating greater impairments, would be negatively

correlated with fixation duration to critical facial regions. Specifically, it was expected that measures of social functioning would be negatively correlated with fixation time to the core region and, in particular, to the eye region.

2 METHOD

2.1 Participants

Thirty-six participants were recruited for the current study and, in accordance with exclusion criteria, a final sample of 34 participants were included for analyses. A post-hoc power analysis was conducted to determine the power afforded by the sample size recruited for the present study using the program G*Power3 (Faul et al., 2007). According to the standards set by Cohen (1988), a sample size of 34 provided sufficient power to detect a medium effect using repeated measures ANOVA (within-between subjects interaction; $\eta_p^2 = .25$ $\alpha = .05$, power $(1-\beta) = .95$).

The TD group included 6 females and 11 males, and ranged in age from 15.33-23.32 years ($M=19.502$, $SD=2.946$). The sample of individuals with ASD consisted of 5 females and 12 males, ranging in age from 15.08-23.25 years ($M=19.558$, $SD=2.967$). Based on administered diagnostic measures, described in detail in a subsequent section, individuals in the ASD group were classified as currently meeting criteria for either autism ($n=9$) or an autism spectrum disorder ($n=8$). There was no significant age difference between groups, $t(32)=-0.056$, $p=.956$. Individuals were not matched on IQ, as research suggests that using IQ as a matching variable can obscure important neurocognitive differences in research with neurodevelopmental populations (Dennis, Francis, Cirino, Schachar, Barnes, & Fletcher, 2009). Previous literature

suggests that proficiency of emotion perception and the brain regions critical to this skill continue to develop in TD individuals through adolescence and into adulthood. This research also suggests that deficits on these more difficult, subtle emotion perception tasks emerge in adults with ASD, with structural abnormalities in critical brain emerging around the age of 15 years (Aylward et al., 2005; Rump et al., 2009; Schultz, 2005; Schumann et al., 2004). As a result, individuals between the ages of 15 and 25 years of age were recruited for the current study in order to capture the group differences in emotion perception. Participants were recruited from the undergraduate research pool at Georgia State University, from community centers, word of mouth, schools and organizations serving individuals with ASD, and participants from previous studies conducted by our research group.

2.2 Experimental Stimuli

Dynamic audio-visual emotion (DAVE) stimuli were developed and validated at the Yale Child Study Center (Robins & Schultz, 2004). Movies were filmed and digitized with the help of a local production studio, using professional actors. Each stimulus consists of one actor delivering an emotionally ambiguous sentence, allowing for each emotion to be considered congruent with the semantic content. Actors delivered 10 different sentences in each of four emotional tones: angry, fearful, happy, and neutral. The complete set consists of 320 movies: 80 matching and 240 mismatching. Example sentences are: “The door is open” and “Look in the box.” To illustrate how the sentences are emotionally ambiguous, the sentence “Look in the box” could be construed as happy if one person is directing another to a birthday present, fearful if there is a snake in the box, angry if the glassware that someone ordered arrived broken, or neutral if one person were asking another if there was room to pack one more book. The digitized movies were then edited to create a set of clips with matching emotional presentations

in face and voice (e.g., happy voice, happy face) and a set of clips with incongruent, mismatching emotional presentations in face and voice (e.g., happy voice, angry face). Digitization allowed for the synchronization of lip movement and sound. Each clip is approximately 2 seconds long. Movie duration ranges can be found in Table 2. Only movies between 1.00 and 2.13 seconds long were included in the current study.

Table 2
Movie duration ranges in seconds.

Sentence	Minimum Duration	Maximum Duration
“Look in the box”	1.05	1.20
“Clouds are in the sky”	1.14	2.13
“It’s dark already”	1.10	1.50
“The dog is barking”	1.05	2.03
“The door is open”	1.08	1.23
“I didn’t expect you”	1.06	2.00
“It might happen”	1.05	1.29
“Put it down”	0.24	1.02
“It’s across the street”	1.07	2.02
“Turn off the television”	1.12	2.06

2.3 Diagnostic Measures

The following two measures were administered in order to confirm the diagnosis of an ASD prior to participation in the current study. Participants not meeting criteria, as described in the sections below, for an ASD on either one or both of these diagnostic measures were excluded from the ASD group of the current study.

Autism Diagnostic Interview – Revised (ADI-R; Rutter, LeCouteur, & Lord, 2003). The ADI-R focuses on three domains of functioning: language and communication, social interactions, and restricted, repetitive, and stereotyped behaviors and interests. The instrument uses diagnostic algorithms that allow for discrimination of either autistic disorder (AD) or one of the broader autism spectrum disorders (ASD), specifically Asperger’s Syndrome (AS) or

Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS). In accordance with DSM-IV-TR criteria, individuals in the current study were classified as AD if they met full criteria on all domains of the ADI-R (n=13). Individuals meeting full criteria on the language and communication, social interactions, and restricted, repetitive, and stereotyped behaviors and interests domains only, without evidence of language impairments prior to 36 months, were classified as AS (n=1). Individuals meeting full criteria on the language and communication and social interactions domains only, with sub-threshold evidence of restricted, repetitive, and stereotyped behaviors and interests and without evidence of language impairments prior to 36 months, were classified as PDD-NOS (n=1). The ADI-R is a parent interview that can be completed in 60 – 120 minutes. Algorithms used for the current study are appropriate for children, adolescents, or adults ages 10 years, 0 months and older. Items are coded by the clinician as 0, 1, or 2 and subsequently converted into algorithm scores. The instrument possesses good inter-rater reliability, with a weighted kappa of .72 across items, and test-retest reliability, with weighted kappas ranging from .74 to .91. The ADI-R has shown excellent reliability in diagnosing individuals with ASD in numerous studies (Rutter, LeCouteur, & Lord, 1994).

Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999). The ADOS is a semi-structured assessment of communication, social interaction, and imaginative use of materials or play (Lord et al., 1999). The ADOS consists of four modules, appropriate for different ages and/or levels of communication. Only Modules 3 and 4 were administered in the current study, as they are appropriate for use with verbal individuals ages 15 and older. The assessment can be completed in approximately 45-60 minutes and entails an observation during which the clinician provides opportunities to exhibit behaviors associated with a diagnosis of

autism spectrum disorders. The behaviors of interest are recorded and coded and the clinician will rate each domain according to the ADOS protocol. Inter-rater and test-retest reliability are good to high, with weighted kappas ranging from .82 to .93 and .59 to .82, respectively.

Individual item validity is high, with correlations from .76 to .93. Correlations between algorithm items and overall domains exceed .50 (Lord et al., 1999). Individuals included in the current study were classified as having autism (n=9) or an autism spectrum disorder (n=8).

Individuals who did not meet criteria for autism or an autism spectrum disorder were excluded from the current study (n=2).

Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003). The SCQ is a 40-item parent-report screening measure that assesses symptomatology associated with ASDs (Rutter et al., 2003). The SCQ consists of yes/no questions and can be completed in approximately 10 minutes. The SCQ has two forms available, appropriate for use in relation to behavior throughout the child's lifetime (Lifetime form) or for use in relation to the most recent three months (Current form). A score of 15 or greater indicates a probability of an ASD. Questions included on the SCQ were selected on the basis of the ADI-R and assess three domains; reciprocal social interaction, communication, and restricted, repetitive, and stereotyped patterns of behavior. Concurrent validity between the SCQ and the ADI-R was high, with Pearson inter-correlations of .92, .73, and .89 for the Reciprocal Social Interaction, Communication, and Restricted, Repetitive, and Stereotyped Patterns of Behaviors domains, respectively. Every effort was made to administer the ADI-R to all participants to all participants included in the ASD group of the current study. However, parents of two participants were unable to complete the ADI-R, but were willing to complete the brief SCQ screener. The

Lifetime form was administered for participants in the current study ($n=2$), and both exceeded the threshold for ASD.

2.4 Inclusion and Exclusion Measures

The following cognitive, perceptual, and emotional measures were administered to all participants in each of the two groups of the current study. As previously discussed, research suggests that individuals with psychopathology (i.e., anxiety, depression, psychosis) or a pervasive developmental disorder employ atypical face processing strategies. In addition, individuals with cognitive or perceptual disorders may employ atypical face processing strategies or be unable to complete an emotion decision task. As such, the following measures were primarily used to determine eligibility for the TD group in the current study; individuals who exceeded predetermined cutoffs for each of the following measures were excluded from the TD group to reduce potential confounding variables. However, due to the expected emotion perception deficits and high comorbidity rates (e.g., anxiety) in the ASD group, individuals who exceeded predetermined cutoffs on the following measures were not excluded from the ASD group; however, rates of comorbidity were minimized to the extent possible.

The Autism-Spectrum Quotient (AQ; Autism Research Centre (ARC); Baron-Cohen et al. 2001). The AQ is a 50-item self-administered questionnaire that assesses for symptoms associated with Autism Spectrum Disorders (ASD) in individuals with normal intelligence. A Total AQ score of 32 or greater indicates a high probability of ASD. The Total AQ is comprised of 5 five subscales; Communication, Social, Imagination, Local Details, and Attention Switching. Test-retest reliability was shown to be good ($r=.7$) and self-report scores were shown to correlate with parent-reports of symptomatology. Although the AQ was normed on a sample

of individuals ages 16 and older, this questionnaire was administered as a screening measure for ASD symptomatology of all participants in the current study. Participants were excluded from the TD group of the current study if they exceeded the cutoff score of ≥ 32 ($n=0$).

Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004). The Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004) is an assessment via self-report of the participant's adaptive and problem behaviors in both the community and the home setting. The current study used the Self-Report of Personality – Adolescent (SRP-A) or the Self-Report of Personality – College (SRP-COL), which are appropriate for the age level of participants. The BASC-2, SRP-A is a 176-item self-report questionnaire that assesses a variety of behavioral difficulties, providing the following Domain scores: School Problems, Internalizing Problems, Inattention/Hyperactivity, Emotional Symptoms Index, and Personal Adjustment. The BASC-2, SRP-COL is a 185-item self-report questionnaire that assesses for a variety of behavioral difficulties. Domain scores are attained for: School Problems, Internalizing Problems, Inattention/Hyperactivity, Emotional Symptoms Index, and Personal Adjustment. The BASC-2 has strong psychometric properties to support its use. Research has shown good reliability. Internal consistency levels across normative samples range from .80 to .95 for composite scores and .80 to .87 for individual scales. These reliabilities are essentially the same in the clinical norm sample. The BASC-2 shows good test-retest reliability, with coefficients ranging from .80 to .90. Strong scale inter-correlations are indicative of the validity of the BASC-2. The BASC-2 correlates strongly with other similar measure of behavior. For example, the BASC-2 and Achenbach System of Empirically Based Assessment (ASEBA) have correlation coefficients ranging from .65 to .84. The composite and scale scores on the BASC-2 were used to assess the severity of any DSM-IV-TR symptoms and to determine

whether participants had significant behavioral or psychological problems. Because of previous research findings that symptoms of depression or anxiety, or psychosis may bias or impede an individual's perception of emotions, those participants exhibiting clinically significant symptoms on the Depression and/or Anxiety scales, as indicated by T-scores of ≥ 70 , were excluded from the TD group of the current study ($n=0$). Data was not available for all subscales due to incomplete ($n=2$) or unreturned ($n=1$) measures; specifics will be provided in the results section. Those participants who self-reported significant symptoms of psychopathology were provided with information about and referred to the Georgia State University Counseling Center and Psychology Clinic.

Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Participants were administered the WASI to estimate overall cognitive ability. The WASI consists of four subtests (Vocabulary, Block Design, Similarities, and Matrix Reasoning) used to assess various aspects of intelligence, including Full Scale IQ (FSIQ), Verbal IQ (VIQ), and Performance IQ (PIQ). These four subtests were selected for the abbreviated scale because they possess the highest loadings (i.e. $> .70$) on g, or general intelligence, and this maximizes the correlation between the WISC-IV (FSIQ) and the WAIS-III FSIQ. Furthermore, the reliability of these four subtests has been shown to be exceptionally high (Wechsler, 1999). The WASI is a psychometrically valid assessment tool. The reliability coefficients range from .84 to .98 for the subtests and the IQ scales. The WASI has also been highly correlated with other overall cognitive ability measurements (Wechsler, 1999). The WASI can be administered as a four-subtest or two-subtest form; the two-subtest form was administered for this study. The two-subtest version of the WASI consists of the Vocabulary and Matrix Reasoning subtests and yields only the FSIQ score. Inter-correlations of FSIQ scores between the WASI two- and four-subtest forms are .94 for

individuals 17-19 years old and .95 for individuals 20-24 years old. To ensure complete understanding of the tasks, individuals who scored lower than 70 on this measure were excluded from both the TD and ASD groups of the current study (n=0).

Benton Facial Recognition Test (BFRT; Benton, 1994). The BFRT assesses an individual's ability to accurately recognize human faces without requiring recall. Each participant was exposed to three different conditions in which they are asked to correctly match human faces: matching of identical front views, matching of a front view with a three-quarter view, and matching a front views under two different lighting conditions. On the long-form of the BFRT, there are 54 separate matches, using a total of 22 stimulus cards. The short-form, from which a total-score is extrapolated, consists of 27 separate matches and uses a total of 13 stimulus cards. The short-form of the BFRT was administered for the current study. On the short-form, there are six items in which the participant is asked to select the single stimulus card that correctly matches a sample photograph. There are seven items in which the participant is asked to select the three stimulus cards that correctly match a sample photograph.

Neuropsychologists commonly use the test and the authors report extensive normative data for individuals over age 16 from various studies indicating good reliability and validity (Benton et al., 1983). The test was used to screen out individuals who did not possess typical facial processing abilities. Individuals scoring below the cutoff (≤ 38) on this measure were excluded from the TD group (n=0). Given the documented difficulty on face recognition tasks, however, individuals were not excluded from the ASD group due to a low score on the BFRT.

Seashore Rhythm Test (Seashore, Lewis, & Saetveit, 1960). This task requires participants to determine if two rhythms presented are matching or mismatching. The task is taken from a subtest of the Halstead-Reitan Neuropsychological Test Battery and requires

participants to discriminate between 30 pairs of rhythmic beats. Classified as a measure of non-verbal auditory discrimination, the reaction time is highly sensitive to the participant's ability to attend to and concentrate on the task. Test-retest differences are minimal and internal reliabilities (split-half and odd-even) are reported as .77 and .62, respectively. This test was administered in the current study to ensure that final participants possess within normal abilities of auditory processing. Participants whose data indicated a mean number of errors greater than or equal to two standard deviations below the mean of a normative sample (i.e., score ≤ 20) were excluded from the both the TD and ASD groups of the current study (n=0).

Diagnostic Analysis of Nonverbal Accuracy, Second Edition (DANVA2; Nowicki, 2004). In order to assess the participants' ability to detect nonverbal cues, participants underwent testing on two of the four subtests of the DANVA2. Each subtest includes 24 trials and asks the viewer to identify, in a forced-choice format, the facial expression depicted in the photograph (Adult Facial Expressions) or the emotion portrayed in the actor's (both male and female) voice (Adult Paralanguage). Two additional subtests, Child Facial Expressions and Child Paralanguage, were not administered in the current study. On each of the subtests, the participants are given the same four response choices (happy, sad, angry, and fearful). The DANVA2 is a computer-based test and the responses are recorded according to which of the four emotion labels are selected with the mouse. Results were found to be consistent over time, with test-retest reliability ranging from .81 to .84 for facial expressions and .73 to .93 for paralanguage. Internal consistency was good, with coefficient alpha ranging from .78 to .83. Criteria related validity has been shown to be good, such that scores on the DANVA2 correlated closely with scores on various psychological measures. See Nowicki & Duke (1994) and Nowicki (2004) for further information on psychometrics and scale creation. These subtests were

administered to assess an individual's ability to decode nonverbal information. Individuals scoring below the cutoff (i.e., greater than 2 standard deviations below the mean) were excluded from the TD group of the current study ($n=0$). Given the documented difficulties with identification of facial expression and affective prosody, individuals were not excluded from the ASD group due to low scores on the DANVA-2.

Adult Paralanguage (DANVA2-AP)—The Adult Paralanguage subtest of the DANVA2 consists of male and female actors stating the sentence “I’m going out of the room now, but I’ll be back later.” After the presentation of this recorded sentence, the participant is asked to indicate the emotion portrayed in the voice by selecting from one of four choices: happy, sad, angry, and fearful. Sentences are played one at a time and can be repeated as often as the participant requires.

Adult Facial Expressions (DANVA2-AF)—The Adult Facial Expressions subtest of the DANVA2 consists of 24 photographs of male and female adults displaying one of four facial expressions: happy, sad, angry, and fearful (Nowicki & Carton, 1993). During administration of this subtest, photographs are displayed one at a time, for 2.5-seconds each. Participants are asked to identify the emotion portrayed in the face.

2.5 Apparatus

Eye-tracking data was collected using a bright pupil image technique and hardware and software created by Applied Science Laboratories[®] (ASL; Bedford, MA). The ASL 6000 tracking system uses a table-mounted chin-rest to eliminate head movement during recording. The eye-tracker camera produces a near infrared light beam that shines into the eye. A portion of that light is reflected off of the retina through the pupil and back to the camera, which rests in front of the presentation computer. The camera uses this reflected light to localize both the

participant's pupil and an image reflection on the cornea. A schematic can be viewed in Figure 4 (Applied Science Laboratories, 2005). By computing the vector between the corneal reflection (a fixed image) and the papillary position (an indicator of direction of gaze), the camera can identify and record the participant's point of gaze (POG). Points of gaze were collected and recorded by EYENAL, the data analysis software, at a rate of 60 samples per second, providing 60-127.8 data points for each stimulus (stimulus length ranges from 1.00 to 2.13 seconds). Eye blinks by the participant or loss of the corneal reflection, identified by a loss of data from the pupil-corneal camera, were excluded from final analysis. A blink was defined as fewer than 12 consecutive lost data fields, where the pupil diameter equaled zero. When a blink occurs during a fixation, then the loss of data did not cause the fixation to end. Consecutive lost fields, where the pupil diameter equaled zero, greater than or equal to 12 were considered losses. When losses occurred during a fixation, the fixation measurement ended. When a loss of corneal reflection was detected, despite valid pupil data, the recorded values were interpreted as invalid data.

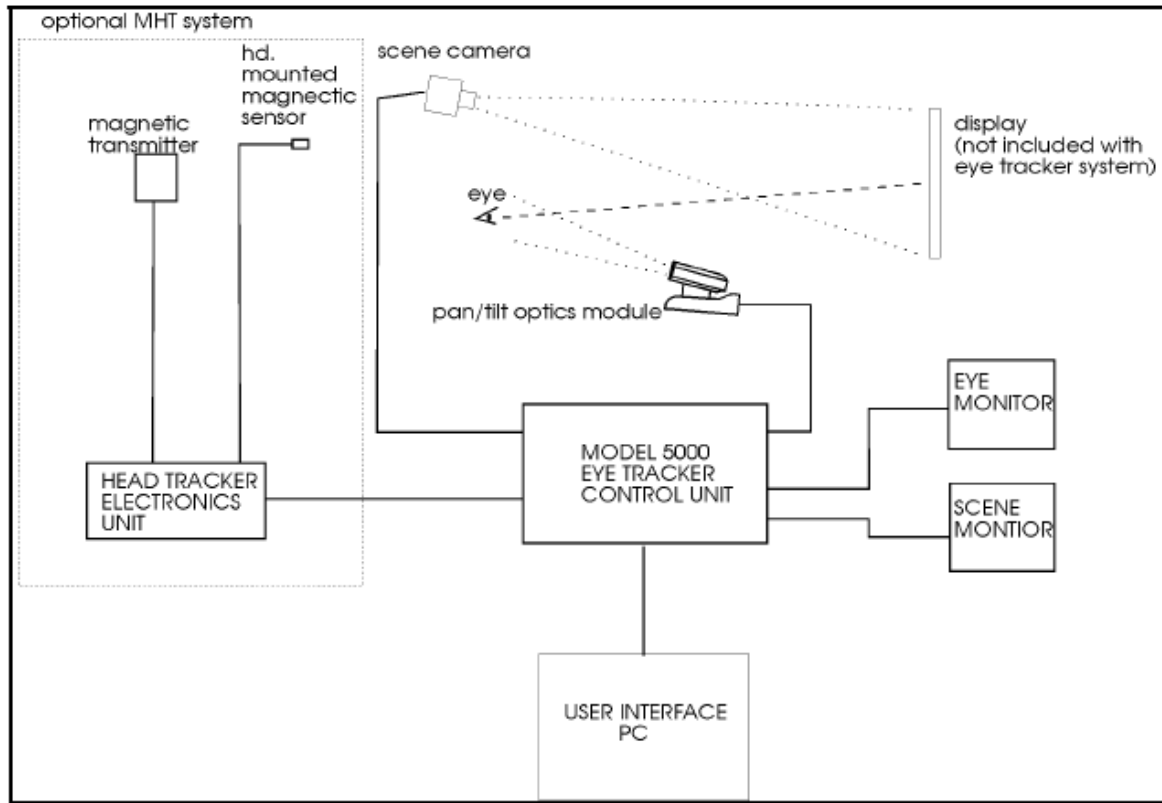


Figure 4. Schematic of ASL 6000 set-up and eye-tracking data collection.

Prior to the start of each movie, a white cross-hair appeared in the center of the screen to ensure that all perceivers began viewing each stimulus from the same point (Manor et al., 1999). Data was collected continuously throughout each of the two tasks; external data signals were sent from DirectRT to Eynal to mark the beginning and end of each movie stimulus. Eynal software provided both spatial and temporal eye gaze data for analysis. Spatial data consisted of the location of fixation points on each DAVE stimulus.

2.6 Procedure

Participants were recruited from the undergraduate participant pool at Georgia State University using an advertisement on the Sona Experiment Managements System (SONA), from community centers, word of mouth, schools and organizations serving individuals with ASD,

and participants from previous studies conducted by our research group. All tasks took place in a research laboratory in the Psychology Department of Georgia State University, Atlanta, Georgia. The current study was approved by the Institutional Review Board of Georgia State University. All individuals participated in testing during one session, with informed consent procedures at the start of the session. Upon arrival at the Psychology Department, the principal investigator explained the entire study and procedures to the participant and his or her guardian, as necessary. The consent and assent forms were reviewed verbally by the PI and signed by the participant and his or her guardian, as necessary. The PI addressed any concerns before obtaining written consent.

The researcher used several precautions to ensure the confidentiality of information obtained during the course of this study. In no case did the participants' actual name or other identifying information appear on any computerized or written test form or on any other data files created in this experiment. Each participant was identified using a unique alphanumeric tag, assigned to participants in the order of enrollment. The signed consent and assent forms served as the only paperwork that connected an individual participant's name to his or her assigned tag; the consent forms were filed in a cabinet separate from the data collected and accessible only to researchers. Computers housing any information pertaining to the study were password-protected and locked in the research lab. Any files that were printed were stored in a locked filing cabinet, accessible only to researchers.

Experimental Procedure. The current study consisted of diagnostic testing, neuropsychological testing, and the experimental tasks. With the exception of diagnostic testing, administration order of measures was counterbalanced across participants. Diagnostic measures were administered by trained and reliable graduate students supervised by a trained and reliable

licensed psychologist prior to all other inclusion/exclusion measures, given that in past research we have found that not all individuals with community diagnoses for ASD meet the stricter research criteria. Additional assessments included the previously described cognitive, perceptual, and emotional tests. Testing for individuals in the ASD group lasted approximately 2-2.5 hours, with child and parent measures completed simultaneously, and testing for individuals in the TD group lasted approximately 1-1.5 hours. Participants were excluded from the TD group if they possess a perceptual disorder that was expected to interfere with the detection of stimuli or if they indicated risk of meeting DSM-IV-TR criteria at the time of screening for a form of major psychopathology (e.g., depression, anxiety, ASD). Individuals were offered periodic breaks and were told that they may request a break at any time during testing.

The experimental part of the session consisted of the non-invasive eye-tracking of participants while they viewed the emotional movie stimuli. Participants were allowed to acclimate to sitting in the chair for as long as required for their comfort. The PI guided each participant through a practice run, watching non-experimental stimuli and responding with the four designated keys on the keyboard. The eye-tracking apparatus was then calibrated before beginning the experiment. Each participant was administered three experimental movie tasks; however, only the first two of these tasks were included in the current study. Instructions for each task were provided before each task began. The total viewing time was approximately 30-45 minutes, for a total session time of 3-3.5 hours for the ASD group and 2-2.5 hours for the TD group.

In this study, the participants were shown DAVE stimuli one at a time and, during a five second response period, asked to make a response. If the participant responded prior to the five-second limit, the task advanced automatically. Stimuli were presented using DirectRT

(Empirisoft Corporation, 2006) on a Dell Intel ® Pentium ® 4 CPU desktop computer using Windows XP operating system. Each stimulus was presented at 500x420 pixels. DirectRT recorded the participant's behavioral responses (emotion identification) and response time. Participants were seated approximately 60-cm from the computer screen and stimuli subtended a visual angle of approximately 36° vertical x 30° horizontal. Before beginning the eye-tracking tasks, participants were provided with general instructions. Prior to each of the two individual tasks of interest, participants were instructed to "Pay attention to the movie" and provided with specific instructions for that task. The script to be used for delivery of instructions can be found in the Appendix.

Task 1. In the first task, the participants were shown 40 congruent movies and asked, "How does this person feel?" Movies were selected to include ten movies portraying each of four emotions. Half of each emotion set consisted of sentences delivered by the male actor and the other half by the female actor, with randomized selection of the sentences delivered by each. Because the participants were unable to respond verbally without moving and disrupting the eye gaze data collection, these tasks employed a forced choice response format. After each movie, a screen was displayed that provides four emotion choices: "Happy", "Angry", "Fearful", and "Neutral." The individual were seated in front of a keyboard on which only four labeled keys were available (i.e., S, C, N, K). Each emotion word cue corresponded spatially to the key that the participant was to use to respond with that emotion (i.e., the words and keys corresponded to each other from left to right). The participant was instructed to respond within the 5-second response period provided after each movie; the task automatically proceeded to the next movie after a response was made or at the end of this 5-second response period. The remaining keys

were covered with a rubber cover. The participant's hands were positioned so that he/she could rest four fingers (e.g., index and middle fingers) comfortably on the operable keys.

Task 2. In the second task, participants were shown both congruent and incongruent movies and asked, "How does this person feel?" The participant was shown a total of 64 movies (32 congruent/32 incongruent) and the order of movie presentation was randomized. Half of each movie type consisted of sentences delivered by the male actor, the other half by the female actor. For the congruent movies, each actor delivered four movies in each of the four emotions, with sentences chosen at random. For the incongruent movies, sentences were again chosen at random. Incongruent movies were selected to ensure that both actors facially express each of the four emotions four times. Across the entire task, each actor facially expressed each emotion eight times (i.e., four congruent, four incongruent). For example, the male actor delivered happy-congruent sentences in four movies (i.e., happy face/happy voice) and four of the incongruent movies included the male actor with a happy facial expression (i.e., happy face/angry voice, happy face/fearful voice). After each movie, the participant was asked to respond using a forced-choice format, with one of the four emotion choices: "Happy", "Angry", "Fearful", and "Neutral". The response procedure was the same as in *Task 1*, with only four labeled keys available and the instructions to respond within a 5-second response period. Because of the forced-choice format of this study, individuals were not afforded the opportunity to respond with a blended emotion. The response choice indicated which of the two modalities had a stronger influence on emotion perception.

These two tasks were administered in the same order for all participants; counterbalanced presentation of the tasks would have contaminated *Task 1*. Because the initial task aimed to examine an individual's fixation pattern when congruent movies were presented without

exposure to incongruent movies, *Task 2* needed to follow *Task 1*. The third task, which was not being used in analyses for the current study, always occurred after Task 2, and thus could not contaminate the tasks of interest. Following the administration of the experimental tasks, participants were debriefed about the purpose of the study. The participants were afforded the opportunity to ask questions regarding any measures or stimuli used in the current study and qualitative feedback was collected regarding procedures and approach to the task.

Data Analysis Procedure. A fixation was defined as a set of consecutive data points during which the point of gaze (POG) had a standard deviation of no more than 0.5 degrees, in either the horizontal or vertical plane, from the original POG. Each fixation began when six sequential gaze positions met these criteria, a period of approximately 102 ms, and ended when the number of sequential gaze positions deviated from the original POG by a degree angle of one or more in either the horizontal or vertical plane. For the current study, spatial data were compared in four areas of interest, defined individually for each movie and based on the location of three core facial features and non-core face. A static image, or one frame of each movie clip, was used to define these regions. The three core regions of interest were the eyes, nose, and mouth. The remainder of the face was considered peripheral, or non-core, face (Pelphrey, 2002). An example of the defined core facial regions can be seen in Figure 1.

Fixation duration was calculated separately for each of the three predefined core regions of the face: eyes, nose, and mouth. The fixation duration for the eyes, nose, and mouth (core) was also combined to allow for comparison to fixation duration for all other regions of the face (peripheral or non-core). The sum of core and peripheral facial regions was totaled for analyses examining total looking time at the face for face vs. non-face contrasts. Total fixation duration

for each region was computed by summing the total fixation duration across all movies within each condition.

3 RESULTS

Data was collected by DirectRT and EYENAL software. Behavioral data collected by DirectRT included emotion selected by key press and response time. EYENAL software provided both spatial and temporal eye gaze data for analysis. Spatial data consisted of the location of fixation points on each stimulus. A fixation was defined as a set of consecutive data points during which the point of gaze (POG) had a standard deviation of no more than 0.5 degrees, in either the horizontal or vertical plane, from the original POG. The start of a fixation began when this consecutive set of data points equaled six or more, a period of approximately 102 ms. The fixation ended when the number of sequential gaze positions deviated from the original POG by a degree angle of one or more in either the horizontal or vertical plane. Temporal data included fixation duration for each of the four predefined facial regions: eyes, nose, mouth, and periphery. Total fixation duration was computed, for each region separately, by summing the total fixation duration across all movies within that condition. Unless otherwise noted, all assumptions of statistical analyses were satisfied.

All individuals were administered a variety of cognitive, perceptual, and emotional measures, and performance of the two groups were compared directly. The TD and ASD groups did not differ significantly on any of the cognitive or perceptual measures (Table 3). However, an expected group difference emerged on a measure of symptoms associated with ASDs (i.e., AQ), with higher scores for the ASD group as compared to the TD group, $t(32) = -5.358, p < .001, \eta_p^2 = .473$.

Table 3

Group performance on cognitive, perceptual, and emotional measures

Measure	Group						t	df	p	η_p^2
	TD			ASD						
	N	Range	M (SD)	N	Range	M (SD)				
FSIQ	17	79-128	101.706 (13.959)	17	73-122	99.118 (15.190)	.517	32	.609	.008
BFRT	17	41-50	45.588 (3.043)	17	37-52	44.059 (4.616)	1.141	32	.263	.039
Seashore	17	23-30	26.882 (1.933)	16	22-30	26.000 (2.066)	1.268	31	.214	.049
DANVA-2 AF	17	13-22	18.118 (2.395)	17	12-22	18.588 (3.083)	-.497	32	.623	.008
DANVA-2 AP	17	7-23	17.353 (3.605)	17	8-21	16.765 (3.308)	.496	32	.623	.008
AQ	17	11-26	17.529 (4.215)	17	16-34	26.706 (5.665)	-5.358	32	<.001	.473
ADI*	--	--	--	15	28-61	47.533 (11.070)	--	--	--	--
ADOS**	--	--	--	17	4-13	7.940 (2.968)	--	--	--	--

*Total Algorithm Score

** Social + Communication Algorithm Score

Typically developing group (TD); Autism Spectrum Disorder group (ASD); *Wechsler Abbreviated Scale of Intelligence* – Full Scale IQ (FSIQ); *Benton Facial Recognition Test* (BFRT); *Seashore Rhythm Test* (Seashore); *Diagnostic Analysis of Nonverbal Accuracy, Second Edition – Adult Faces and Adult Paralanguage* (DANVA-2 AF and DANVA-2 AP); *The Autism-Spectrum Quotient* (AQ); *Autism Diagnostic Interview* (ADI); *Autism Diagnostic Observation Schedule* (ADOS)

In addition, a number of elevations were observed in the ASD group on a measure of emotional and behavioral functioning (i.e., BASC-2; Table 4). As previously mentioned, no individuals were excluded from the TD group based on results on specific subscales of interest to this study (i.e., Anxiety, Depression). Elevations in the TD group were isolated to one subscale (i.e., Social Stress), and this subscale was therefore included in group-comparison analyses. Comparison of group performance on subscales of interest revealed that individuals with ASD demonstrated higher levels of depressive symptomatology and social stress (i.e., BASC-2, Depression and Social Stress subscales) than did their age-matched peers (Table 5).

Table 4

Number of significant elevations (T-score ≥ 70) on the BASC-2

Subscale	Group	
	TD	ASD
Anxiety	--	1
Depression	--	2
Social Stress	1	3
Atypicality	--	4
Locus of control	--	5
Somatization	--	1
Attention Problems	--	2
Hyperactivity	--	4

Table 5
Group performance on subscales of interest for the BASC-2

Measure	Group						t	df	p
	TD			ASD					
	N	Range	M (SD)	N	Range	M (SD)			
Anxiety	16	33-61	49.813 (8.440)	16	37-70	52.063 (11.228)	-.641	30	.527
Depression	17	34-55	44.882 (5.243)	16	38-72	55.875 (11.307)	-3.546	31	.002
Social Stress	17	34-74	46.353 (9.950)	15	36-84	56.933 (13.879)	-2.528	30	.017

3.1 Hypothesis 1: Accuracy and Reaction Time during Congruent Movies.

It was predicted that, when viewing congruent movies (i.e., Task 1), both TD and ASD groups would be able to correctly identify the emotion portrayed by the actors. In order to test this hypothesis, the number of correct responses for each group was summed and compared in an independent samples t-test. Exploration of the data indicated that the assumption of normality had been violated. As a result, analyses for this hypothesis was performed on log-transformed (i.e., Log10) accuracy data. Levene's test was considered for between group analyses and indicated that the assumption of equal variances had been upheld. Although accuracy was very high for both groups (95% for TD group, 91% for ASD group), analyses revealed that TD

individuals responded correctly significantly more often ($M=38.132$, $SD=1.938$) than individuals with ASD ($M=36.529$, $SD=2.832$) when viewing congruent movies, $t(32)=-2.074$, $p=.046$, $\eta_p^2=.118$.

An exploratory analysis was conducted to determine whether accuracy rates were consistent (i.e., TD group accuracy was greater than ASD group accuracy) when viewing congruent movies in isolation (i.e., 40 congruent stimuli in Task 1) and in the context of incongruent movies (i.e., 32 congruent stimuli in Task 2). When viewing congruent movies in Task 2, analyses indicated that TD individuals again responded correctly significantly more often ($M=30.882$, $SD=1.219$) than did ASD individuals ($M=26.882$, $SD=7.490$), $t(32)=-2.991$, $p=.005$, $\eta_p^2=.218$. The TD individuals correctly responded to 94% of the movies, and individuals with ASD correctly responded to 87% of the movies.

In order to test the prediction that individuals with ASD would respond more slowly than the TD individuals following the perception of congruent movies (Task 1), an independent-samples t-test compared the response time between the two groups. Response times to label congruent movies were comparable for the TD group ($M=807.918$ ms, $SD=238.458$ ms) and the ASD group ($M=853.445$ ms, $SD=227.086$ ms), $t(32)= -.569$, $p=.573$, $\eta_p^2=.010$. These findings failed to support the hypothesis. For both TD and ASD groups, response time to label congruent movies in Task 2 (complete data presented in Hypothesis 3) was significantly shorter than response time to label congruent movies in Task 1, $F(1,32)=8.933$, $p=.005$, $\eta_p^2=.218$.

3.2 Hypothesis 2: Fixation Pattern during Congruent Movies.

A series of two-way ANOVAs were conducted to explore group difference in fixation patterns during the perception of congruent movies (Task 1).

Hypothesis 2a. Face versus Non-Face. The first of these analyses, conducted to compare face (sum of core and periphery) and non-face regions, was a 2 x 2 (Group x Region) mixed design ANOVA (Table 6). In this analysis, the between-group variable (Group) consisted of two levels: TD and ASD. The within-group variable (Region) also consisted of two levels: Face and Non-Face. There was a significant main effect of Region, $F(1,32)=182.292, p<.001, \eta_p^2=.851$, with greater fixation duration on the face as compared to non-face region. The main effect of Group was non-significant, $F(1,32)=2.119, p=.155, \eta_p^2=.062$, indicating that both TD and ASD individuals fixated on the stimuli for similar amounts of time. The Group x Region interaction was non-significant, $F(1,32)=1.723, p=.199, \eta_p^2=.051$. Findings failed to support the hypothesis; both TD and ASD individuals fixated for longer on the facial regions as compared to non-face regions, and both groups fixated on facial regions at a comparable rate (Figure 5).

Table 6

Group fixation duration (sec) data for face and non-face regions during congruent movies of Task 1

Group	Region	
	Face	Non-Face
TD	68.822 (16.952)	7.671 (8.331)
ASD	59.686 (19.917)	9.372 (6.892)

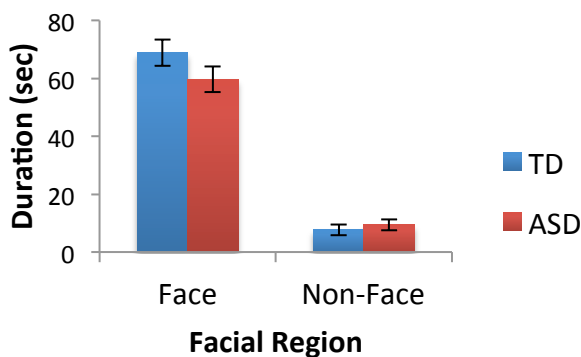


Figure 5. Total fixation duration for the face and non-face regions when viewing congruent movies.

Hypothesis 2b. Core versus Periphery. In order to compare fixation time among the facial regions, a second analysis was a 2 x 2 (Group x Region) mixed design ANOVA (Table 7). The between-group variable (Group) again consisted of two levels: TD and ASD. In this analysis, the within-group variable (Region) consisted of two face regions: Core and Periphery. There was a significant main effect of Region, $F(1,32)=179.728, p<.001, \eta_p^2=.849$, with greater fixation duration on core regions of the face (i.e., eyes, nose, mouth) as compared to the peripheral regions (e.g., chin, forehead, cheeks). There was no main effect of Group, $F(1,32)=2.074, p=.160, \eta_p^2=.061$, or Group x Region interaction, $F(1,32)=1.860, p=.182, \eta_p^2=.055$. Findings indicated that overall fixation duration was comparable for TD and ASD individuals, and both groups showed a similar preference for core as compared to peripheral facial regions (Figure 6).

Table 7

Group fixation duration (sec) data for core and periphery regions during congruent movies of Task 1

Group	Region	
	Core	Periphery
TD	62.299 (19.730)	6.523 (4.659)
ASD	52.581 (19.848)	7.105 (4.686)

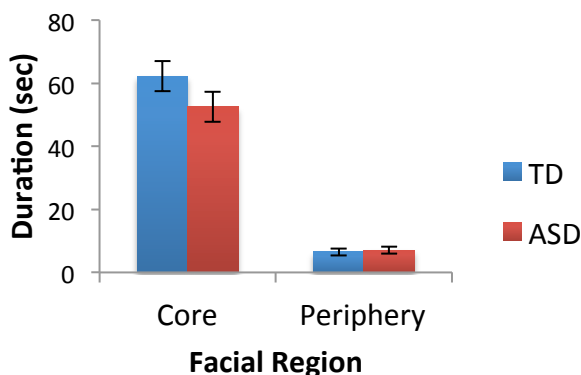


Figure 6. Total fixation duration for the core and peripheral face regions when viewing congruent movies.

Hypothesis 2c. Eyes versus Nose versus Mouth. A 2 x 3 (Group x Region) mixed-design ANOVA was conducted to explore the fixation pattern among core regions of the face during the perception of congruent movies (Table 8). The between-group variable in this analysis (Group) consisted of two levels: TD and ASD. The within-group variable (Region) consisted of three levels: eyes, nose, and mouth. Mauchly's test indicated that the assumption of sphericity had been violated for within-subject analyses (i.e., there are significant differences between the variances of differences). Therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity, $\chi^2(2)=10.815, p=.004, \epsilon=.772$. There was a significant main effect of Region, $F(1.545,49.439)=11.621, p<.001, \eta_p^2=.266$. Contrasts indicated that fixation duration was significantly greater for eyes than nose, $F(1,32)=8.858, p=.006, \eta_p^2=.217$, and greater for nose than mouth, $F(1,32)=4.590, p=.040, \eta_p^2=.125$ (Figure 7).

Table 8

Group fixation duration (sec) data for eye, nose, and mouth regions during congruent movies of Task 1

Group	Region		
	Eyes	Nose	Mouth
TD	34.541 (25.015)	17.645 (16.072)	10.113 (16.382)
ASD	29.003 (19.569)	14.972 (13.305)	8.606 (9.446)

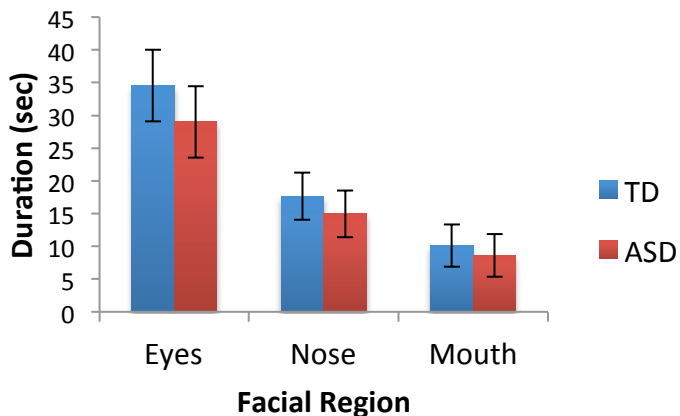


Figure 7. Total fixation duration for the eye, nose, and mouth regions when viewing congruent movies.

There was no significant main effect of Group, $F(1,32)=2.050$, $p=.162$, $\eta_p^2=.060$, or a Group x Region interaction, $F(2,64)=0.095$, $p=.910$, $\eta_p^2=.003$. Findings failed to support the hypothesis; analyses indicated that overall fixation duration was comparable for both TD and ASD individuals, and that both groups demonstrated the same preference for fixating on the eyes.

3.3 Hypothesis 3: Behavioral Responses and Reaction Time during Incongruent Movies.

When identifying the emotions portrayed in incongruent movies, it was predicted that TD participants would be significantly more likely to choose labels consistent with facial expressions than with vocal prosody. In contrast, it was hypothesized that individuals with ASD would show no bias toward facial information and would choose labels consistent with facial expression and prosody at similar rates. In order to examine the pattern of responses, a 2x2 (Group x Modality) mixed-design ANOVA was performed (Table 9). The between-group variable (Group) consisted of two levels: TD and ASD. The within-group variable (Modality) also consisted of two levels: Face and Voice. Please note that responses matching neither the face nor the voice (e.g., happy response for angry face/fearful voice) were excluded from analyses.

Table 9
Group modality bias data for incongruent movies of Task 2

Group	Region	
	Face	Voice
TD	23.412 (6.893)	6.059 (5.517)
ASD	18.059 (6.986)	8.941 (4.828)

There was a significant main effect of Modality, $F(1,32)= 43.437, p<.001, \eta_p^2=.576$, with a greater number of responses consistent with the facial expression than with the vocal prosody. The main effect of Group approached significance, $F(1,32)= 4.008, p=.054, \eta_p^2=.111$, suggesting that TD individuals were marginally more likely to respond with emotions consistent with either the facial expression or the vocal prosody than the ASD group, who made marginally more responses consistent with neither the face nor the voice (e.g., responded “angry” for a happy face, fearful voice movie). Analyses also revealed a significant Group x Modality interaction, $F(1,32)= 4.20, p=.049, \eta_p^2=.116$. These findings provide support for the hypothesis; TD individuals were significantly more likely to respond with emotions consistent with the facial expression, whereas the individuals with ASD did not show as strong a preference for the facial affective information. Of note, the ASD group did not respond with labels consistent with the face and those consistent with the voice at equal rates, but the facial bias was significantly weaker for individuals with ASD as compared to TD individuals.

In order to test the hypothesis that response time to label emotions would be greater during the incongruent movie condition than during the congruent movie condition for both TD and ASD groups and that the increase in response time for the incongruent movies would be most pronounced for the individuals with ASD, a 2 x 2 (Group x Congruence) mixed-design ANOVA was conducted. The between-group variable (Group) consisted of two levels: TD and ASD. The within-group variable (Congruence) also consisted of two levels: Congruent and

Incongruent. There was a significant main effect of Congruence, $F(1,32)= 42.894, p<.001, \eta_p^2=.573$, with significantly longer response times during incongruent movies as compared to congruent movies. The main effect of Group was non-significant, $F(1,32)= 0.374, p=.545, \eta_p^2=.012$. There was also no significant Group x Congruence interaction, $F(1,32)= 1.274, p=.267, \eta_p^2=.038$. Results provided partial support for the hypothesis, as response times were greater during the incongruent movies for both groups. In contrast to the prediction however, the increase in response times during incongruent movies was comparable for both the TD and ASD groups (Table 10).

Table 10

Response time data for congruent and incongruent movies of Task 2

Group	Congruence	
	Congruent	Incongruent
TD	672.748 (258.214)	813.588 (431.312)
ASD	661.612 (261.891)	904.653 (306.356)

3.4 Hypothesis 4: Fixation Pattern during Incongruent Movies as compared to Congruent Movies.

A series of ANOVAs were used to examine group differences in fixation patterns among predefined regions of the face during congruent and incongruent movies.

Hypothesis 4a. Face versus Non-Face. The first of these 2 x 2 x 2 (Group x Region x Congruence) mixed-design ANOVAs compared the group differences between face (sum of core and periphery) and non-face regions during both movie conditions (Table 11). The between-group variable (Group) consisted of two levels: TD and ASD. The first within-group variable (Region) consisted of two levels: Face and Non-Face. The second within-group variable (Congruence) consisted of two levels: Congruent and Incongruent.

Table 11

Group fixation duration (sec) data for face and non-face regions during congruent and incongruent movies of Task 2

Group	Region			
	Face	Non-Face	Face	Non-Face
	Congruent		Incongruent	
TD	59.697 (10.029)	3.083 (2.539)	62.205 (6.681)	2.110 (1.480)
ASD	46.464 (18.961)	6.757 (4.164)	50.548 (18.649)	5.244 (4.875)

Results revealed significant main effects of Group, $F(1,32) = 4.339, p = .045, \eta_p^2 = .119$, Region, $F(1,32) = 308.817, p < .001, \eta_p^2 = .906$, and Congruence, $F(1,32) = 9.926, p = .004, \eta_p^2 = .237$. Overall fixation duration was greater for the TD individuals as compared to ASD individuals, for incongruent movies as compared to congruent movies, and for face regions as compared to non-face regions. There was a significant Group x Region interaction, $F(1,32) = 7.625, p = .009, \eta_p^2 = .192$. During both congruent and incongruent movies, the difference between fixation duration to the face and non-face regions was greater for the TD individuals than ASD individuals (Figure 8).

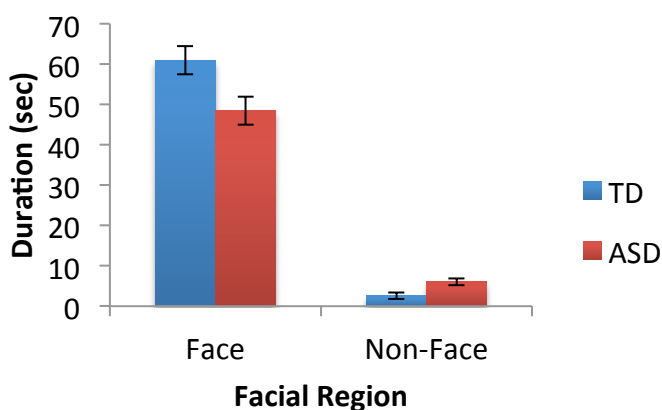


Figure 8. Total fixation duration for the face and non-face regions when viewing all movies.

In addition, there was a significant Region x Congruence interaction, $F(1,32) = 23.705, p < .001, \eta_p^2 = .426$. The difference between fixation duration to the face regions as compared to

non-face regions was greater during incongruent movies than congruent movies (Figure 9).

There was no significant Group x Congruence interaction, $F(1,32) = 0.632, p = .433, \eta_p^2 = .019$, or

Group x Region x Congruence interaction, $F(1,32) = 1.286, p = .265, \eta_p^2 = .039$.

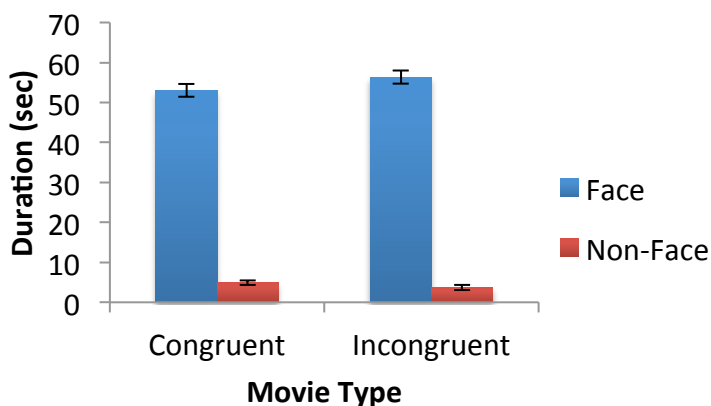


Figure 9. Combined group data for total fixation duration to face and non-face regions when viewing congruent and incongruent movies.

In summary, the significant Group x Region and Region x Congruence interactions, in the absence of a Group x Region x Congruence interaction, supported the prediction that the difference between fixation duration to the face and non-face regions would be greater for the TD individuals than ASD individuals across both movie types. Results indicated that preference for fixation on the face region (i.e., greater difference between fixation to face as compared to non-face) was greater for TD than ASD individuals. This pattern remained consistent during both congruent and incongruent movies.

Hypothesis 4b. Core versus Periphery. The second of the 2 x 2 x 2 (Group x Region x Congruence) mixed-design ANOVAs compared the group differences among the facial regions during both movie conditions (Table 12). The between-group variable (Group) consisted of two levels: TD and ASD. The first within-group variable (Region) consisted of two levels: Core and

Periphery. The second within-group variable (Congruence) consisted of two levels: Congruent and Incongruent.

Table 12

Group fixation duration (sec) data for core and periphery regions during congruent and incongruent movies of Task 2

Group	Region			
	Core	Periphery	Core	Periphery
	Congruent		Incongruent	
TD	53.868 (10.977)	5.829 (3.015)	56.354 (7.645)	5.851 (2.450)
ASD	39.461 (16.822)	7.004 (3.853)	42.442 (16.615)	8.106 (5.611)

Analyses revealed significant main effects of Group, $F(1,32) = 6.293, p = .017, \eta_p^2 = .164$, Region, $F(1,32) = 329.135, p < .001, \eta_p^2 = .911$, and Congruence, $F(1,32) = 23.267, p < .001, \eta_p^2 = .422$. Overall fixation duration was significantly greater for TD individuals as compared to ASD individuals, for the core region as compared to the peripheral region, and during the incongruent as compared to the congruent movies. There was also a significant Group x Region interaction, $F(1,32) = 12.137, p = .001, \eta_p^2 = .275$. The preference for fixation on the core face region was greater for the TD individuals than the ASD individuals (Figure 10).

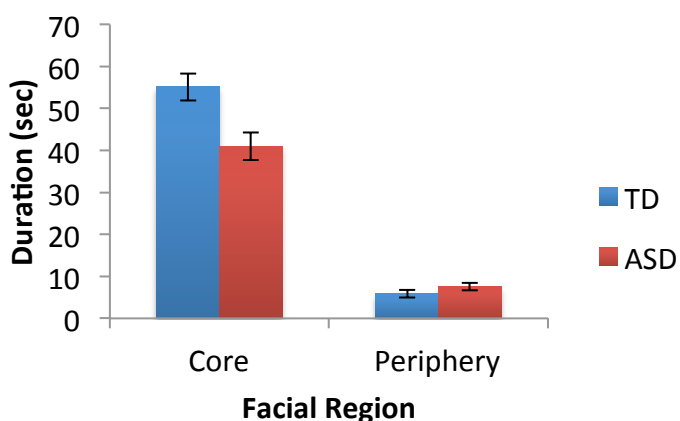


Figure 10. Total fixation duration for the core and peripheral face regions when viewing all movies.

The Region x Congruence interaction approached significance, $F(1,32) = 3.976, p = .055, \eta_p^2 = .111$. This trend suggests that the difference between fixation duration between core and peripheral regions was marginally greater during incongruent movies for both TD and ASD groups (Figure 11). The Group x Congruence interaction was non-significant, $F(1,32) = 1.334, p = .257, \eta_p^2 = .040$. Similarly, the Group x Region x Congruence interaction was non-significant, $F(1,32) = 0.072, p = .790, \eta_p^2 = .002$. Contrary to the predicted change in fixation pattern (i.e., individuals with ASD would demonstrate a decrease in fixation duration to the core region during incongruent movies), both groups showed a non-significant increase in fixation duration to the core region during incongruent movies.

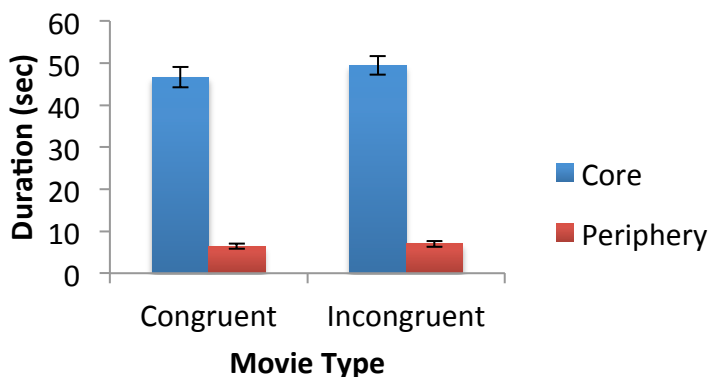


Figure 11. Combined group data for total fixation duration to the core and peripheral face regions when viewing congruent and incongruent movies.

The non-significant Group x Region x Congruence interaction failed to support the prediction that TD participants would increase fixation time to the core regions of the face during the incongruent movies as compared to congruent movies whereas the ASD group would decrease fixation time to the core regions of the face during the incongruent movies as compared to congruent movies. The results of this analysis instead indicated that the TD group's preference for the core regions remained comparable in both congruent and incongruent movies. In addition,

both TD and ASD individuals showed a marginally significant increase in core fixation duration during incongruent movies as compared to congruent movies.

Hypothesis 4c. Eyes versus Nose versus Mouth. The third analysis conducted was a 2 x 3 x 2 (Group x Region x Congruence) mixed-design ANOVA comparing fixation duration among the core facial regions during both movie types (Table 13). The between-group variable (Group) consisted of two levels: TD and ASD. The first within-group variable (Region) consisted of three levels: eyes, nose, and mouth. The second within-group variable (Congruence) consisted of two levels: Congruent and Incongruent. Mauchly's test indicated that the assumption of sphericity had been violated for within-group analyses for the variable Region (i.e., there are significant differences between the variances of differences). Therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity, $\chi^2(2)=12.679, p=.002, \epsilon=.749$.

Table 13

Group fixation duration (sec) data for eye, nose, and mouth regions during congruent and incongruent movies of Task 2

Group	Region					
	Congruent			Incongruent		
	Eyes	Nose	Mouth	Eyes	Nose	Mouth
TD	26.357 (21.304)	17.933 (15.232)	9.579 (13.375)	29.046 (21.016)	16.708 (14.056)	10.601 (13.567)
ASD	20.379 (16.405)	12.226 (10.963)	6.856 (7.493)	22.533 (16.423)	12.424 (11.738)	7.485 (7.965)

Analyses revealed significant main effects of Group, $F(1,32)=9.475, p=.004, \eta_p^2=.228$, Region, $F(1.497,47.916)=7.509, p=.003, \eta_p^2=.190$, and Congruence, $F(1,32)=13.739, p=.001, \eta_p^2=.300$. Fixation duration was greater for TD individuals as compared to ASD individuals and greater during incongruent movies as compared to congruent movies. Contrasts revealed that fixation duration was greater for eyes compared to nose, $F(1,32)=4.201, p=.049, \eta_p^2=.116$, and

for nose as compared to mouth, $F(1,32) = 5.285, p = .028, \eta_p^2 = .142$. There was also a significant Region x Congruence interaction, $F(2,64) = 8.767, p < .001, \eta_p^2 = .215$. Contrasts indicated that the preference for the eyes as compared to the nose was greater during incongruent movies than congruent movies; individuals increased fixation duration to the eyes during incongruent movies, whereas fixation duration to the nose was decreased, $F(1,32) = 16.572, p < .001, \eta_p^2 = .341$. There was no Group x Region interaction, $F(2,64) = .082, p = .921, \eta_p^2 = .003$, or Group x Congruence interaction, $F(1,32) = .113, p = .739, \eta_p^2 = .004$. Similarly, there was no Group x Region x Congruence interaction, $F(2,64) = 1.209, p = .305, \eta_p^2 = .036$. Findings failed to support the hypothesized group differences amongst the three core facial regions. Contrary to the predicted decrease in fixation duration to the eye region for individuals with ASD during incongruent movies, both TD and ASD individuals showed increased fixation duration to the eye region during incongruent movies (Figures 12 and 13).

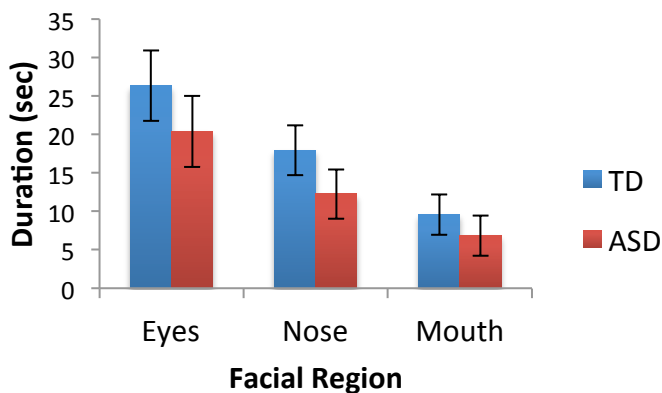


Figure 12. Total fixation duration for the eye, nose, and mouth regions when viewing congruent movies.

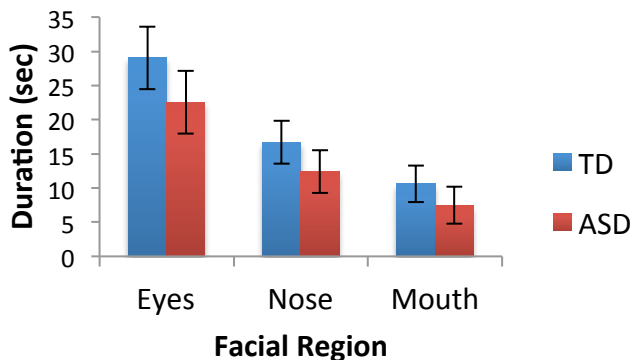


Figure 13. Total fixation duration for the eye, nose, and mouth regions when viewing incongruent movies.

3.5 Hypothesis 5: Exploratory Analysis of Initial Fixations

Differences in location of initial fixations were explored for TD and ASD groups in a 2 x 3 x 2 (Group x Region x Congruence) mixed-design ANOVA. This analysis compared the group differences in the total number of first fixations (i.e., number of movies) devoted to each of the core regions (i.e., eyes, nose, and mouth; Table 14). For this analysis, each individual participant could have potentially produced 32 first fixations during congruent movies and 32 first fixations during incongruent movies, provided that he or she fixated on each stimulus at least once. As the focus of this analysis was on differences in initially fixations to core facial regions, initial fixations that were devoted to other stimulus regions (i.e., peripheral or non-face regions) were not included.

Table 14

Group initial fixation data for eye, nose, and mouth regions during congruent and incongruent movies of Task 2

Group	Region					
	Eyes	Nose	Mouth	Eyes	Nose	Mouth
	Congruent			Incongruent		
TD	9.529 (5.088)	12.647 (6.363)	2.882 (2.913)	11.412 (4.797)	12.706 (5.301)	2.588 (3.261)
ASD	8.059 (6.210)	8.941 (6.179)	2.353 (2.473)	8.824 (5.151)	8.941 (5.573)	2.000 (1.768)

The between-group variable (Group) consisted of two levels: TD and ASD. The first within-group variable (Region) consisted of three levels: eyes, nose, and mouth. The second within-group variable (Congruence) consisted of two levels: Congruent and Incongruent. Mauchly's test indicated that the assumption of sphericity had been violated for within-group analyses for the variable Region (i.e., there are significant differences between the variances of differences). Therefore, the degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity, $\chi^2(2)=8.848, p=.012, \epsilon=.801$. There was a significant main effect of Region, $F(1.602,51.270) = 25.692, p < .001, \eta_p^2 = .445$. Contrasts indicated that a greater number of first fixations were directed toward the eyes than to the mouth, $F(1,33) = 34.973, p < .001, \eta_p^2 = .517$. There was also a main effect of Group, $F(1,32) = 9.185, p = .005, \eta_p^2 = .223$, with a greater number of first fixations located on these core regions (i.e., on the eyes, nose, or mouth) for the TD individuals than the ASD individuals. There was no main effect of Congruence, $F(1,32) = 2.785, p = .105, \eta_p^2 = .080$. The Region x Congruence interaction was significant, $F(2,64) = 3.726, p = .029, \eta_p^2 = .104$, indicating that individuals differentially increased first fixations to the eyes, nose, and mouth regions during the incongruent movies. Contrasts indicated that the difference between number of first fixations on the eyes as compared to the nose decreased, $F(1,32) = 4.203, p = .049, \eta_p^2 = .116$, and the difference between the eyes as compared to the mouth, $F(1,32) = 7.338, p = .011, \eta_p^2 = .187$, increased during incongruent movies compared to congruent movies. Specifically, whereas number of first fixations remained comparable for the nose and mouth regions, individuals demonstrated an increase in the number of movies during which they first fixated on the eye region. There were no significant interaction effects for Group x Region, $F(2,64) = .808, p = .450, \eta_p^2 = .025$, Group x Congruence, $F(1,32) = 1.003, p = .324, \eta_p^2 = .030$, or Group x Region x Congruence, $F(2,64) = .463, p = .632, \eta_p^2 = .014$. Results failed to

provide support for the hypothesis (i.e., TD group would first fixate on the eye region, ASD group would first fixate on the mouth region), as both TD and ASD individuals first fixated on the eye or nose regions during congruent movies (Figures 14 and 15). In addition, both groups increased fixations to the eye regions during incongruent movies compared to congruent movies.

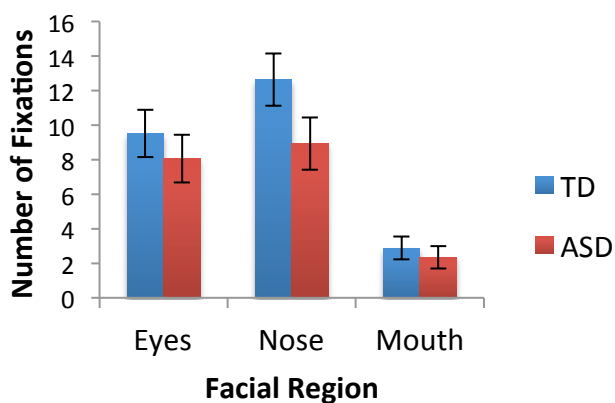


Figure 14. Number of first fixations devoted to the eye, nose, and mouth regions when viewing congruent movies.

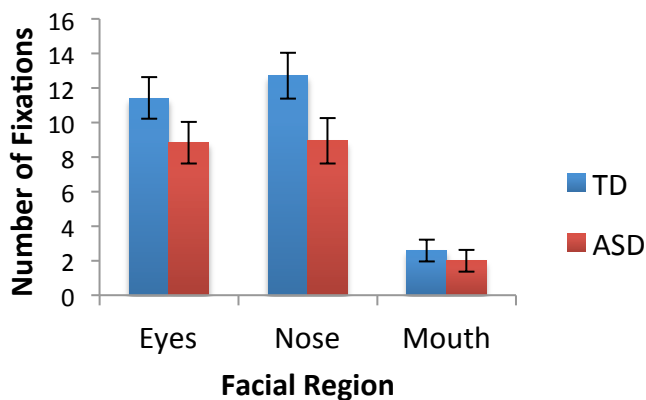


Figure 15. Number of first fixations devoted to the eye, nose, and mouth regions when viewing incongruent movies.

3.6 Hypothesis 6: Exploratory Analysis of the Relationship between Visual Fixations and Social Functioning

The relationship between measures of visual scanning (i.e., fixation duration to face, core, and eye regions) and measures of social functioning (e.g., Autism-Spectrum Quotient,

Autism Diagnostic Observation Schedule) were explored through a series of correlations (Table 15). Across all participants, analyses revealed that the AQ Total Score was negatively correlated with fixation duration to the critical regions. Similarly, within the ASD group, scores on the ADOS Reciprocal Social Interaction domain were negatively correlated with fixation duration to the critical regions. Furthermore, data showed a trend toward greater strength of relationship between several visual fixation measures (e.g., Core, Face) and measures of social functioning during Task 2 as compared to Task 1. As the AQ was normed on a sample of participants ages 16 years and older, correlation analyses were also repeated after excluding participants aged 15 years ($n=3$). These follow-up analyses revealed only marginal decreases in significance of correlations between AQ scores and fixation duration to the eyes during congruent ($r= -.294$) and incongruent ($r= -.332$) of Task 2.

Table 15

Correlations between fixation duration data and measures of social functioning

Measure	Task 1			Task 2					
	Congruent			Congruent			Incongruent		
	Face	Core	Eyes	Face	Core	Eyes	Face	Core	Eyes
AQ Total	-.422*	-.422*	-.382*	-.469**	-.510**	-.301*	-.465**	-.549**	-.337^
ADOS ⁺ Social	-.577*	-.582*	-.574*	-.622**	-.683**	-.570*	-.584*	-.730**	-.553*

⁺ ASD group only

* significant at $p<.05$ level

** significant at $p<.01$ level

^ approaches significance at $p=.052$

Autism-Spectrum Quotient (AQ); Autism Diagnostic Observation Schedule (ADOS)

4 DISCUSSION

4.1 Behavioral Performance

Affective Judgments. Social interactions require the integration of both visual and auditory information, a process that occurs quickly and, seemingly automatically, during the

perceptual stage of processing (Hietanen et al., 2004; Pourtois et al., 2000; Pourtois, Debatisse, Desplan, & de Gelder, 2002). In the current study, TD individuals were able to explicitly label the emotion when provided with congruent stimuli. This is consistent with previous research showing that TD individuals are consistently able to identify the basic emotions, such as the six basic emotions (i.e., happiness, sadness, anger, fear, disgust, and surprise) that are recognized across cultures (e.g., Ekman & Friesen, 1975; e.g., Bryant & Barrett, 2008; Ekman, 1992; Izard, 1971, 1992; Wang et al., 2006). Contrary to the hypothesis, however, individuals with ASD performed more poorly than did TD individuals, with regard to accuracy of emotion identification, when viewing congruent movies. Despite the fact that individuals with ASD were generally successfully (87% or better) in the identification of emotions in the DAVE stimuli, which portrayed basic emotions and provided congruent information in both the visual and auditory modalities, the ASD individuals still struggled to accurately identify these emotions at a rate comparable to that of their same-aged peers (95% or better). Although this effect is rather small, it seems important to consider this effect, along with other findings (e.g., fixation patterns), and it will, therefore, be discussed further in term of the implications of this research.

These findings are not consistent with that of other studies, in which ASD and TD groups showed no difference in emotion recognition abilities (Kirchner, Hatri, Heekeren, & Dziobek, 2011). In general, individuals with ASD perform similarly to controls when they are presented with unambiguous, prototypical emotions and allowed sufficient processing time before a response is required (Capps, Yirmiya, & Sigman, 1992; Grossman, Klin, Carter, & Volkmar, 2000; Humphreys, Minshew, Leonard, & Behrmann, 2007; Ozonoff, Pennington, & Rogers, 1990). However, some other studies have shown that weaknesses in emotion perception emerge when individuals with ASD are shown stimuli for brief periods of time or when emotional

expressions are subtle (Critchley et al., 2000; Greimel et al., 2010; Humphreys et al., 2007; Law Smith et al., 2010; Pelphey, et al., 2002; Philip et al., 2010; Rump, Giovannelli, Minshew, & Strauss, 2009). In one recent study, researchers found that individuals with Asperger's disorder were less accurate in emotion identification when presented with full face, eyes only, or mouth only photos, for both basic and complex emotions, even when allowed unlimited time to observe the photos prior to making a response (Sawyer, Williamson, & Young, 2012).

Researchers have posited that dynamic emotion stimuli, as compared to static stimuli, improve the perception of facial expressions and more closely simulate social interactions by providing biological motion and requiring rapid processing of changing facial configurations believed to signal unique emotions (Ambadar, Schooler, & Cohn, 2005; Aviezer et al., 2008; Hess & Kleck, 1995; Kohler et al., 2004; Lander & Chuang, 2005; LeBar, Crupain, Voyvodic, & McCarthy, 2003; Sato et al., 2004; Nusseck, Cunningham, Wallraven, & Bulthoff, 2008; Smith, Cottrell, Gosselin, & Shyns, 2005). In addition, it has been suggested that providing congruent emotional information through multiple modalities (e.g., facial expression and prosody as compared to facial expression alone) results in reduced and speeded processing effort for affective judgments (Paulmann, Jessen, & Kotz, 2009). Findings in the current study suggest that individuals with ASD do not show the same benefit, in terms of improved emotion identification of dynamic audio-visual stimuli, that TD individuals do. This supposition is supported by previous studies showing that ASD individuals do not show a preference for biological motion over object motion (Annaz et al., 2012) and that they responded less accurately even when presented with video clips that morphed or transitioned from neutral to emotional expressions (Evers, Noens, Steyaert, & Wagemans, 2011; Law Smith et al., 2010).

Other researchers have postulated that emotion recognition improves with age. Specifically, data have suggested that TD adults are more proficient (i.e., processing speed, ability to detect subtle differences) than younger controls, whereas ASD individuals did not show this same linear trend across development (Rump et al., 2009). The results of the current study provide similar findings; TD adolescents and young adults were able to maintain extremely high rates of accuracy even when presented with brief stimuli. In contrast, the adolescents and young adults with ASD in the current study were unable to perform at a level comparable to their age-matched peers. One possible explanation for this difference in accuracy is that the performance of the individuals with ASD was negatively impacted by slowed speed of processing during the perceptual stage. Although some researchers have failed to find deficits in static emotion identification tasks with adolescents and adults with Asperger's syndrome, the current findings suggest that deficits in the accurate identification of brief, dynamic stimuli persist in older individuals with ASD (e.g., Capps, Yirmiya, & Sigman, 1992; Rutherford & Towns, 2008).

Response Time. Following from the proposition that individuals with ASD demonstrate slowed processing speed on emotion identification tasks, results of response time measures are important to consider. Contrary to the hypothesized difference, response time to label congruent movies was comparable for TD and ASD groups in this study. These results are consistent with other studies that have reported similar response times between TD and ASD adolescents (Kirchner, Hatri, Heekeren, & Dziobek, 2011; Law Smith 2010). These findings, however, fail to support preliminary results obtained in research using the DAVE stimuli, as well as other studies demonstrating increased response times in individuals with ASD as compared to TD individuals (Bal et al., 2010; Sawyer, Williamson, & Young, 2012). It is important to note, however, that these previous studies reporting increased RTs for individuals with ASD employed tasks in

which response times were calculated when participants first pressed a button to indicate that they had identified the emotion. Participants were then presented with a response screen, from which they indicated which emotion they had identified in the video. These findings suggest that, when provided with the opportunity for extended viewing times (e.g., ≥ 15 seconds), individuals with ASD take longer to identify emotions; however, they may provide a response choice in similar amounts of time following the removal of a stimulus (Bal et al., 2010; Sawyer, Williamson, & Young, 2012). In addition, a preliminary study that utilized the DAVE stimuli employed a forced-choice yes/no response task; participants were asked to determine whether the emotion presented in the stimulus matched an emotion word cue presented prior to stimulus onset (Banks et al., 2008). It would be expected that this task required additional cognitive abilities (e.g., working memory, switching) that may have resulted in response time disparities for the individuals with ASD.

We sometimes encounter situations, albeit at a relatively low frequency, in which we receive social cues that appear inconsistent with one another. We then have to decipher these perplexing cues and make a decision about how to proceed in the social interaction. A portion of the current DAVE stimuli allowed for the simulation of this relatively more complex emotion perception, as some were digitally altered to contain mismatching facial expression and vocal prosody. The finding that all participants were slower to respond to the incongruent movies than congruent movies in an emotion decision task provides evidence of the increased difficulty of such responses. These findings are commensurate with other reports of increased response times during incongruent emotion (deGelder, 2000; Hietanen, Leppanen, Illi, & Surakka, 2004; Massaro & Egan, 1996; Pell, 2005), low intensity emotion (Law Smith et al., 2010), and incongruent non-emotion tasks (Hu et al., 2012; Plank et al., 2012).

These DAVE stimuli also allowed for examination of the types of decisions individuals made when faced with conflicting affective cues. In real-life social situations, individuals have the opportunity to blend visual and auditory speech information or affective cues, in a manner illustrated by the McGurk Illusion (McGurk & MacDonald, 1976). This was a forced-choice emotion identification task, so participants were unable to provide blended responses. In the current study, responses of the TD individuals were more strongly influenced by the visual modality than the auditory modality than were those of the ASD individuals. Although both groups responded more often with emotion judgments matching the facial expression, this dominance effect was more pronounced for the TD group. These findings replicate those studies with TD individuals that have found that visual information is dominant during visual-auditory stimuli (Collignon et al., 2008; Hecht & Reiner, 2008; Koppen, Alsius, & Spence, 2008) and visual affective information has a greater influence on prosodic information than vice versa (de Gelder & Vroomen, 2000; Massaro, 1998; Massaro & Egan, 1996). Results are also commensurate with other reports of reduced influence of visual speech information on auditory speech (Irwin, Tornatore, Brancazio, & Whalen, 2011; Mongillo et al., 2008). Similarly, these findings are consistent with preliminary findings using the DAVE stimuli, indicating that visual information is not as salient to ASD individuals when they are presented with incongruent visual-auditory affective information.

Weakness in Early Integration and Reduced Accuracy. Perhaps what is most interesting is the finding that individuals with ASD were responding quickly, at a rate comparable to their age-matched peers, but were still unable to perform at a comparable level in terms of accuracy. It could be argued that this discrepancy implies that there was a trade-off between speed and accuracy for the individuals with ASD, a finding opposite of that which you expect to see with

emotion recognition judgments in a typical control group, where accurate responses are performed faster than inaccurate ones (Kirouac & Dore, 1983).

One study of response latency for universally recognized emotions showed that emotion recognition occurs quickly (i.e., approximately 600 ms when correct) in TD individuals. Tracy and Robins (2008) found that participants were generally accurate in the identification of universally recognized emotions even when given a time-limited response period (i.e., 1 second). Participants in this study also demonstrated improvements in accuracy following a period of deliberation (i.e., up to 8 seconds) for some more difficult emotions (i.e., those with relatively lower accuracy rates under limited response periods). Of particular importance was the finding that performing this same emotion identification task under cognitive load (i.e., holding a 7-digit number in working memory for the duration of the task) led to increased response time for participants (i.e., approximately 700 ms) without a negative influence on accuracy of responses. This finding led researchers to propose that emotion recognition occurs efficiently and automatically (Tracy & Robins, 2008). Evidence from neuroimaging (Plank et al., 2012), electrophysiological (Giard & Peronnet, 1999; Hu et al., 2012; Paulmann & Pell, 2010; Pourtois, 2000) and behavioral (deGelder & Vroomen, 2000; deGelder, Vroomen, & Bertelson, 1998) studies provide further support for this supposition; the early integration and automatic processing of emotional information occurs during the perceptual stage of processing rather than the later response stage in TD individuals.

In light of the evidence for the automaticity of emotion identification in TD individuals, two important considerations emerge from the current results. Given the trade-off between speed and accuracy that occurred in the ASD group, in that performance was worse under similar response times, it might be suggested that their accuracy would have improved following

deliberation (Tracy & Robins, 2008). Alternatively, perhaps there is a weakness in audio-visual integration in the ASD group (Boucher, Lewis, & Collis, 2000; Gepner, deGelder, & deSchonen, 1996; Loveland et al., 1995; Mongillo et al., 2008) that attenuates the automatic processing of emotional information shown to occur during the perceptual stage (e.g., Hu et al., 2012; Paulmann & Pell, 2010; Pourtois, 2000). This might suggest that weaknesses in automatic processing of emotional stimuli demonstrated by ASD individuals preempts their ability to respond correctly when the perception is time-limited. Despite being unable to glean the same critical information from the brief stimuli as age-matched controls, the fact that the ASD group did not deliberate over responses for longer periods of time (average response time was approximately 800 ms out of 5000 ms allowed to respond) suggest that they were likely unaware of their inaccuracies.

Weakness in Early Integration and Conflict Resolution. Perhaps when typically developing participants are faced with bimodal information that does not quickly register as belonging to a particular affective category, they recruit additional cognitive resources to arrive upon an affective judgment. For example, some have found that, when participants are presented with spatially congruent auditory sound sources and semantically matching visual stimulus, there is increased activation in left inferior frontal cortex in addition to activation in the right middle and superior temporal gyri (Plank et al., 2012). Authors noted that the left inferior frontal cortex, near the left ventrolateral and orbitofrontal cortex, has been shown to play a role in combining auditory and visual information from the “what” and “where” (Plank et al., 2012). In another example, the perception of congruent prosody and lexico-semantic information, relative to prosody-only, has shown bilateral superior temporal gyri and left middle temporal gyrus, as well as the left ventrolateral prefrontal cortex (Mitchell, 2006). However, when participants were

required to identify emotional prosody in the presence of incongruent lexico-semantic information, there was increased activation in the left inferior frontal gyrus, left ventrolateral prefrontal cortex, bilateral middle temporal gyri, right superior temporal gyrus, as well as other regions. Activation in the right middle and superior temporal gyri and right inferior frontal gyrus has been associated with decoding of emotional prosodic cues (Mitchell, 2003), whereas the left inferior frontal gyrus has been associated with the syntactic processing in sentence comprehension (Bilenko, Grindrod, Myers, & Blumstein, 2008). Neuroimaging studies have also identified a neural network involved in the processing of cognitive conflict. Increased activation in the right dorsal anterior cingulate cortex (related to interfering information), right superior temporal gyrus, and right superior temporal sulcus when listening to incongruent as compared to congruent prosody and semantic content (Wittfoth et al., 2010).

Taken together, these findings suggest that the perception of incongruent bimodal information leads to additional processing in anatomical regions associated with unimodal perception, particularly for the non-dominant modality (speech comprehension), in typically-developing individuals. As applied to the current study, Robins, Hunyadi, and Schultz (2009) found increased activation in the posterior superior temporal sulcus during the perception of congruent DAVE stimuli when compared to unimodal presentation of the DAVE stimuli. The superior temporal sulcus (STS) region and the superior temporal gyrus (STG) have been implicated in the integration of visual and auditory perceptual cues and in emotion perception (Allison, Puce, & McCarthy, 2000; Ethofer et al., 2006; Robins, Hunyadi, & Schultz, 2009). Perhaps the perception of incongruent stimuli would lead to further processing in regions associated with decoding facial expressions (e.g., fusiform gyrus; Gauthier et al., 1998; Gauthier & Tarr 1997; Kanwisher, McDermott, & Chun, 1997; Kanwisher & Moscovitch, 2000) and,

more importantly, the decoding of prosody (lateral temporal lobes, frontal-temporal region; Mitchell, 2003).

If this is the case in TD individuals, perhaps abnormal functioning in these regions plays a role in the relative weakness in the integration of facial affective and prosodic cues in ASD. For example, abnormalities in the STS have been strongly implicated in ASD, including abnormal grey matter volumes, cortical thinning, changes in cerebral blood flow (for a review, see Zilbovicius, Meresse, Chabane, Brunelle, Samson, & Boddaert, 2006). Additional research shows reduced or altered activation in the PFC during social processing tasks. For example, Hall, Szechtman, and Nahmias (2003) found that individuals with ASD had an increased activation when matching facial and vocal affect despite a greater number of errors. Furthermore, it is possible that abnormal functioning in regions associated with the unimodal decoding of facial expressions and prosody result in aberrant processing of incongruent bimodal information in ASD individuals. Imaging studies with ASD individuals have revealed both structural and functional abnormalities within the fusiform gyrus (FG), with reduced activation during emotion processing tasks and decreased connectivity with frontal regions, amygdala, and thalamus when viewing faces (Kleinmans et al., 2008; Koshino et al., 2000; Neuhaus, Beauchain, & Bernier, 2010; Pelphrey, Morris, McCarthy, & LaBar, 2007; Schultz et al., 2000). Others have found reduced activation in frontal-temporal regions, including the left middle temporal gyrus, left medial prefrontal cortex, and left precuneus, during a prosodic perception task (Hesling et al., 2010). These differences in neural activation and the proposed relation to processing of emotional cues may account for the disparate modality responses found in the current study. Specifically, these differences may lead to less efficient processing and reduced saliency of critical information presented in multiple modalities for individuals with ASD.

4.2 Fixation Patterns: Congruent and Incongruent Movies

Congruent Movies. The results provide a great amount of information about the pattern of fixations used by individuals during the perception of congruent emotional stimuli. Contrary to hypothesized group differences in visual scanning patterns, the TD individuals and individuals with ASD in this study demonstrated similar patterns of fixations when viewing congruent movies alone (i.e., Task 1; Table 16). Consistent with previous research with TD individuals, all participants fixated significantly longer on the face regions and, more specifically, the core regions of the face (eyes, nose, mouth) than other regions (Luria & Strauss, 1978; Macworth & Bruner, 1970; Manor et al., 1999; Mertens, Sigmund, & Grusser, 1993; Stacey et al., 2005). Among these core regions, participants demonstrated a bias toward the eye region, in that they looked more often toward the eye region than both the nose and the mouth regions.

Table 16

Summary of fixation duration findings for congruent movies of Task 1

Regions	Congruent Movies	
	Result	Interpretation
Face v. Non-Face		Face > Non-Face
Core v. Periphery	Region*	
	Group (ns) Region x Group (ns)	Core > Periphery
Eyes v. Nose v. Mouth		Eyes > Nose > Mouth

* $p < .001$

The Gaze Assumption Hypothesis states that participants rely on the parts of a stimulus that are critical to the task instructions (Lansing & McConkie, 1999). During an emotion identification task, all participants elicited critical information from the eyes, with the nose and mouth regions providing additional information, when presented with congruent, basic emotions. The visual scanning patterns evidenced during the congruent task support and extend previous research utilizing static (Rutherford & Towns, 2008) or unimodal (Bal et al., 2010) stimuli, and

they counter the argument that individuals with ASD show reduced salience of important affective information provided by the eye region (Hernandez, 2009; Jones et al., 2008; Joseph and Tanaka, 2003; Klin et al., 2002; Norbury et al., 2009; Pelphrey et al., 2002; Rutherford & Towns, 2008). In the case of static displays of simple emotions, researchers have found that individuals with ASD demonstrate a similar rate of fixation on the eye region as TD individuals (Rutherford & Towns, 2008). In another study, individuals with ASD and TD individuals demonstrated comparable eye fixation duration when viewing dynamic (i.e., digitally morphing) stimuli over extended periods of time (Bal et al., 2010).

Congruent and Incongruent Movies. Looking at the difference between congruent and incongruent movies, it was predicted that the incongruent movies would be relatively more complex and less automatic, resulting in an increased reliance on the most affectively salient features of the face for TD individuals. Given pervasive difficulties in the processing of important social cues, it was expected that adolescents with ASD would evidence different processing under relatively more complex conditions (i.e., dynamic, brief presentation, incongruent stimuli). In the context of the relatively more complex, incongruent stimuli, as indicated by increased fixation duration for all individuals, group differences in facial fixation patterns emerged (Table 17).

Table 17

Congruent v. Incongruent Movies		
Regions	Result	Interpretation
Face v. Non-Face	Region***, Group*, Congruence** Region x Group** Region x Congruence *** Group x Congruence (ns) Region x Group x Congruence (ns)	Face > Non-Face TD > ASD Incongruent > Congruent Face v. Non-Face: TD > ASD, Incongruent > Congruent
Core v.	Region***, Group*, Congruence***	Core > Periphery

Periphery	Region x Group** Region x Congruence ^ Group x Congruence (ns) Region x Group x Congruence (ns)	TD > ASD Incongruent > Congruent Core v. Periphery: TD > ASD, Incongruent > Congruent
Eyes v. Nose v. Mouth	Region**, Group**, Congruence** Region x Congruence *** Region x Group (ns) Group x Congruence (ns) Region x Group x Congruence (ns)	Eyes > Nose > Mouth TD > ASD Incongruent > Congruent Eyes v. Nose: Incongruent > Congruent

Summary of fixation duration findings for congruent and incongruent movies of Task 2

* significant at $p < .05$ level

** significant at $p < .01$ level

*** significant at $p < .001$ level

^ approaches significance at $p = .055$

Across all regions of interest, TD individuals showed greater fixation duration for both congruent and incongruent stimuli. These findings provide support for previous research showing overall reduced fixation to the face, core, and eye regions in individuals with ASD (Hernandez, 2009; Jones et al., 2008; Joseph & Tanaka, 2003; Kirchner, Hatri, Heekeren, & Dziobek, 2011; Klin et al., 2002; Norbury et al., 2009; Pelphrey et al., 2002; Rutherford & Towns, 2008; Spezio et al., 2007). Despite the comparable distribution of fixations during Task 1 (congruent movies), groups did not demonstrate equivalent patterns of fixations to all critical regions during Task 2 (congruent + incongruent movies). Specific regions, or combinations of regions at broader levels of analysis, were more salient to TD individuals during both congruent and incongruent movies of Task 2, as evidenced by greater difference in fixation time to these regions (i.e., face, core, eyes) as compared to less salient regions (i.e., non-face, periphery, mouth). Despite the fact that some combinations of the critical regions (i.e., face, core) were particularly salient for the TD group during congruent and incongruent movies, both groups showed a comparable preference for the eye region. Contrary to the prediction, the eye region appeared to provide additional critical information during more complex affective judgments for

all individuals in this study. More precisely, these findings illustrate that the relative proportion of fixations between each of the core regions (eyes, nose, mouth) were analogous for both groups. However, a greater proportion of fixations were devoted to the core and facial regions for the TD group.

Evidence for a reduced affinity for critical information under complex conditions can be found through exploration of initial fixations (Table 18). The findings that overall number of initial fixations to the core regions (eyes, nose, or mouth) was greater for TD than ASD provides evidence that, during brief, relatively more complex emotion decisions, individuals with ASD do not as strong of an initial tendency toward the core regions. Consistent with some previous research, however, both groups were more likely to first fixate on the eyes among the three core regions (Williamson, & Young, 2012). In addition, both groups specifically increased fixations to the eye regions during incongruent movies (number of first fixations remained comparable for the nose and mouth regions).

Table 18

Summary of initial fixation findings for congruent and incongruent movies of Task 2

Regions	Congruent v. Incongruent Movies	
	Result	Interpretation
Eyes v. Nose v. Mouth	Region*** Group** Region x Congruence *	Eyes > Mouth TD > ASD Eyes v. Mouth: Incongruent > Congruent
	Congruence (ns) Region x Group (ns) Group x Congruence (ns) Region x Group x Congruence (ns)	

* significant at $p < .05$ level

** significant at $p < .01$ level

*** significant at $p < .001$ level

Looking versus Perceiving. Considering both fixation patterns between groups and performance on behavioral measures, particularly with regard to accuracy of emotion

identification, the question of whether fixations are directly related to effective processing emerges. If there is a causal relationship between visual attention and perception, then this might suggest that sufficient visual attention would lead to intact performance on an emotion identification task. As applied to the current study, this proposition would suggest that TD individuals demonstrated proficiency in emotion perception; they devoted a sufficient level of attention to salient facial regions, perceived critical affective information, and successfully identified emotional content across both tasks. In contrast, the fixation duration of ASD individuals was insufficient for processing and extracting critical affective information from the face, leading to their relatively reduced accuracy in emotion decisions. This effect was particularly evident during Task 1; both groups demonstrated comparable fixation patterns, but the ASD demonstrated reduced accuracy.

As previously discussed, the early integration and automatic processing of emotional information occurs during the perceptual stage of processing rather than the later response stage in TD individuals (e.g., Pourtois, Debatisse, Desplan, & de Gelder, 2002). Perhaps the slightly attenuated behavioral performance in the individuals with ASD results from a specific weakness or temporal delay during this perceptual stage of processing. Evidence for this comes from reports of abnormalities in early event-related potential (ERP) response to faces (Batty et al., 2011; O'Connor et al., 2005). When compared to age-matched controls, children with autism have produced delayed ERPs for two components (i.e., N170 and P1), as well as decreased amplitude for one component (i.e., P1), established as early measures of face processing (Batty et al., 2011). Authors also reported that ERP latency and amplitude decreased with age in both groups, indicating increased sensitivity to face stimuli. Interestingly, when children with autism were compared to verbal-ability matched controls (i.e., younger control children in this study),

only the difference in amplitude of the earlier P1 response component remained. Results suggest a developmental delay for face sensitivity in children with autism, but this delay appears to decrease with increasing age (Batty et al., 2011). Other researchers have also reported a developmental delay in toddler with ASD, as measured by ERPs (i.e., Nc, N290, P400) sensitive to attention to and familiarity of faces (Webb et al., 2011).

Weakness in Early Integration and Face Processing Approach. Given these apparent deficits in the early processing of facial stimuli for individuals with ASD, the subsequent perception of affect is not surprisingly impaired. Some research has indicated that TD participants process all of the features of the face in parallel, as opposed to representing these features individually (Bradshaw & Wallace, 1971; Farah et al., 1998). Preliminary findings utilizing DAVE stimuli suggested that TD participants perceived information from parts of the face on which they did not directly fixate, representing a holistic processing approach. Rather than only perceiving face regions within their foveal vision, it was hypothesized that participants were able to process facial information present in their peripheral vision in parallel (McManus et al., Unpublished Thesis). As children with autism have previously been shown to be less likely to use holistic information in the perception of faces, it seems plausible that individuals with ASD in this study did not process important facial information that was critical to accurate emotion identification (Behrmann et al., 2006; Davies, Bishop, Manstead, & Tantam, 1994; Hobson et al., 1988; Joseph & Tanaka, 2003; Karatekin et al., 2007; Schultz, 2005).

Recall that dynamic stimuli are believed to improve the perception of facial expressions by providing critical temporal information, in the form of changing facial configurations, which are believed to signal unique emotions. For example, two emotions for which identification relies on the shape of the eyes can generally be distinguished on the basis of unique eye configurations

(e.g., angry, fearful; Kohler et al., 2004). As a result, the ability to process dynamic changes perceived through peripheral vision may signal changes in configuration that are unique to a particular emotion (Nusseck, Cunningham, Wallraven, & Bulthoff, 2008). As it pertains to the current study, this suggests that weaknesses in the holistic processing of facial configurations contributed to perceptual deficits for the individuals with ASD.

Taken together, holistic processing occurs early and leads to the accurate perception of faces and facial affect. The question of how many fixations are required to achieve holistic processing then warrants exploration. Some researchers argue that facial identity can be recognized in a single fixation. In one particular study, initial fixations located at the nose occurred at rates above chance, lending support for the argument of the holistic perception of the face (Hsiao & Cottrell, 2008). In this same study, researchers demonstrated that second fixations also occurred at the nose, and that accuracy performance did not improve with additional fixations. The close relationship between facial and emotion recognition, both said to rely heavily on configural or holistic information, therefore suggests that a single fixation may be sufficient to accurately identify facial affect.

With regard to the current study, examination of first fixations revealed that individuals with ASD were less likely to initially fixate on the core regions of the face (eyes, nose, mouth), regions most relevant to emotion perception (e.g., Pelphrey et al., 2002; Rutherford & Towns, 2008). Following from the strong evidence for early, parallel processing of important facial configurations, this diminished tendency toward initial core fixation may account, in part, for the relatively reduced accuracy in emotion perception in the individuals with ASD. Recall that a crosshair was presented at the center of the computer screen prior to stimulus onset (i.e., approximately at the nose region), cuing all participants to begin processing the stimulus from

the same position. Among the core facial regions, TD individuals initially fixated on the nose and the eyes during congruent movies, with increased salience of the eyes during incongruent movies, suggesting that they were able to perceive sufficient affective information from this vantage point. As results indicated that individuals with ASD also showed a preference for the nose among the core regions, alternative explanations for the small effect of reduced accuracy must be considered. As previously noted, it is possible that the individuals with ASD are less efficient at the holistic processing of facial information, indicating that greater fixation times or additional fixations might lead to improved performance.

4.3 Social Cognition

Emotion Perception and Social Impairment. Across both groups, individuals who rated themselves as demonstrating a greater number of behaviors or preferences that are typically associated with ASDs showed less of a preference for the critical regions of the stimulus. Specifically, higher rates of ASD symptomatology were correlated with shorter fixation time to the face, core, and eye regions. This finding was consistent when examining the individuals with ASD alone, using a behavioral measure of social relatedness (ADOS); individuals with ASD who were rated as demonstrating greater deficits in reciprocal social communication showed less of a preference for the face, core, and eye regions. These findings are consistent with previous studies demonstrating that the processing of faces (Itier & Batty, 2009; Leopold & Rhodes, 2010) and facial emotion (Petroni et al., 2011) is associated with social cognition skills. Moreover, researchers have demonstrated an association between facial emotion recognition and social adaptive behavior in individuals with ASD (Garcia-Villamizar, Rojahn, Zaja, & Jodra, 2010).

In the current study, the relationship between facial perception strategies and ASD symptomatology, and more specifically social impairments, was relatively more pronounced under relatively more complex conditions (i.e., congruent and incongruent movies presented together). This suggests that deficits in social functioning were more detrimental to successful emotion perception under conditions that required higher-level cognitive processes. In accordance with this proposition, social impairments associated with ASDs have been shown to be related to deficits in executive functions, such as cognitive flexibility, impaired performance monitoring, and diminished error-correction abilities (Agam, Joseph, Barton, & Manoach, 2010). One particular aspect of executive functioning that seems particularly relevant to the current study is cognitive control, defined as the ability to monitor and regulate goal-directed behaviors by evaluating contextual information and using that information to appropriately adjust behavior (Botvinick, Carter, Braver, Barch, & Cohen, 2001). Given that participants in the current study were required to rapidly evaluate stimuli and detect affective conflict between modalities, research on cognitive control will be considered briefly in the context of its impact on subsequent emotion identification.

Cognitive Control and Conflict Resolution. As mentioned earlier, neuroimaging studies have identified a neural network involved in the processing of cognitive conflict, showing increased activation in the right dorsal anterior cingulate cortex (ACC), right superior temporal gyrus, and right superior temporal sulcus when listening to incongruent as compared to congruent prosody and semantic content (Wittfoth et al., 2010). In studies using different cognitive tasks (i.e., Stroop task, Eriksen flanker task), researchers have demonstrated a conflict adjustment or adaptation effect (Kim, Chung, & Kim, 2012; Larson, South, Clayson, & Clawson, 2012, respectively). The conflict adaptation effect is demonstrated when individuals respond

more slowly on incongruent trials preceded by congruent trials than on those preceded by incongruent trials. Researchers argue that the detection of incongruence on the initial trial results in the recruitment of additional cognitive resources (i.e., allocation of greater cognitive control) to improve subsequent performance (Clayson & Larson, 2011; Kim, Chung, & Kim, 2012; Larson et al., 2012). Specifically, researchers have proposed that conflict detection processes originate in the anterior cingulate cortex (ACC), which then recruits the dorsolateral prefrontal cortex (DLPFC) to minimize that conflict activation on subsequent trials (Kim, Chung, & Kim, 2012; Larson, South, Clayson, & Clawson, 2012).

Using the Eriksen flanker task to compare electrophysiological responses of TD individuals to individuals with ASD, Larson, South, Clayson, and Clawson (2012) found that TD demonstrated the expected electrophysiological evidence (i.e., ERPs) of conflict adaptation; the N2 amplitude, mediated by the ACC, was more negative for incongruent trials preceded by congruent trials than incongruent preceded by incongruent trials. In this study, electrophysiological evidence of conflict adaptation was absent in the ASD group. Furthermore, individuals with ASD rated higher on the ADOS scores evidenced weaker conflict adaptation effects (e.g., amplitude decrease was not attenuated for incongruent-incongruent trials). In another study that specifically examined cognitive conflict resolution with social stimuli (e.g., incongruent verbal and nonverbal information), participants engaged additional neural regions that have been implicated in monitoring response conflict across cognitive and affective domains when encountering incongruent social cues (i.e., dorsal anterior cingulate, posterior medial prefrontal, and ventrolateral prefrontal regions; Zaki, Hennigan, Weber, & Ochsner, 2010).

Based on this research, findings of the current study might suggest that the individuals with ASD were not as effective in the early detection of conflict and, therefore, their visual

perceptual strategies did not adapt to task demands in the same manner as did those of the TD individuals. More explicitly, it is argued that the rapid detection of affective cognitive conflict by the ACC results in the recruitment of additional neural regions involved in the perception of affective information from particular modalities. In support of this argument, Zaki, Hennigan, Weber, and Ochsner (2010) reported that, when participants were presented with incongruent social cues, activation patterns varied as a function of the modality bias evidenced in participant responses (e.g., modality with which their response was consistent). When participants relied more on nonverbal (i.e., facial expressions) cues than contextual cues (i.e., verbal descriptions of an event), activation in the ACC and DLPFC was coupled with activation in the premotor cortex, whereas reliance on contextual cues results in activation coupled with the left temporal cortex and angular gyrus. These regions have been implicated in the processing of domain specific information (i.e., face, hand, and leg motion, and semantic processing, respectively; Price, 2000; Wheaton et al., 2004)

Extrapolating from the above findings, the results of the current study suggest that the successfully early detection of conflict in the TD group and the resulting bias toward visual affective information, as indicated by behavioral responses, signaled for neural activation in regions associated with the processing of faces. It is argued that this bias concurrently resulted in increased fixations to the salient features for the perception of facial emotion (i.e., face, core, or eyes) for all stimuli during the more complex task (i.e., congruent and incongruent movies presented together). In contrast, perhaps the ASD group did not show the same early detection of conflict and, therefore, their visual strategies were not comparably adapted. Consideration of the decreased bias toward facial affective information and lesser bias toward salient facial information (i.e., face, core) provides evidence for this proposition.

4.4 Conclusions

The current study provides new insight into the field of emotion perception for individuals who are typically developing and those who are diagnosed with an autism spectrum disorder. Through the direct exploration of gaze behavior, information about the visual sampling of facial features was collected during the perception of novel dynamic audio-visual emotion (DAVE) stimuli. These stimuli were digitally altered to allow for the presentation of both congruent and incongruent audio-visual affect. These results replicate and expand previous knowledge of emotion perception in typically developing individuals, as well as afford researchers and clinicians a foundation from which to explore the social deficits found in individuals with ASDs.

Results indicated that all participants were able to accurately identify happy, angry, fearful, and neutral emotions. The increased task difficulty of the incongruent condition was evidenced by the increase in response times for both groups. Whereas TD individuals showed a bias toward the visual affective information when asked to determine the emotion portrayed, individuals with ASD did not show the same preference for visual affective information. Individuals with ASD showed a relative deficit in emotion identification, suggesting that they did not benefit from the provision of biological motion (i.e., dynamic stimuli) and bimodal affective information (i.e., audio-visual) to the same extent as did their TD peers when presented with congruent stimuli. The fact that a relative reduction in accuracy occurred for the ASD group despite comparable response times between groups, and that responses were not biased toward facial affect, suggests that disparities in emotion perception occurred early in the perceptual stage of processing.

The early, automatic integration and perception of affective cues appeared to rapidly influence visual scanning behavior. Participants demonstrated similar fixation patterns during the perception of congruent movies, in that both individuals with ASD and TD individuals showed a preference for the eye region. Under more complex conditions, requiring the detection of incongruence and subsequently making a decision regarding emotion identification, group differences emerged in visual approach. When TD participants perceive incongruence at the early stage of integration, their visual fixation patterns were efficiently altered as a function of the extraordinary nature of the stimuli. However, because of impairments in the automatic perception of emotional information in individuals with ASD, their visual fixation patterns were not as efficiently adapted to compensate for the increasing complexity of the task. Typically-developing individuals showed an increased preference for the most salient regions during the more complex task, increasing fixation duration to the face, and specifically to the core region, in order to maintain the same level of accuracy. In contrast, the individuals with ASD did not show the same increase to the face or to the core region. The exception to this is that the increase in and distribution among the core regions (i.e., eyes, nose, mouth) remained comparable for both groups during congruent and incongruent movies. Differences in the overall reliance on the core region likely influenced the subsequent emotion judgments, as TD individuals were more likely to respond with the facial emotion and showed a greater preference for fixation on the core region. It appears that this influence emerges particularly very early in perception, as the TD group also showed a specific increase in their initial fixations to the core region.

The perceptual differences that emerged and the difficulties in emotion perception, albeit of a small effect, demonstrated by the individuals with ASD in this controlled experimental study, have great implications for social skills intervention research. The presentation of

congruent and incongruent movies more closely simulates real-life social encounters, wherein individuals must rapidly discern the mental state and intentions of others amidst a variety of stimuli. Even when presented in brief segments, without the additional pressure of a face-to-face social interaction, differences in performance emerged for the ASD group in this study.

Furthermore, the visual scanning patterns were negatively associated with symptoms associated with ASDs, and specifically with level of social impairment. As individuals in this study were high-functioning individuals ASDs without cognitive impairments, it is logical to predict that more pronounced differences in visual scanning strategies would be found in individuals with more profound social impairments.

As such, the results of the current study highlight the importance of developing social skills interventions programs that target these areas of weakness in individuals with ASD. Specifically, intervention programs should utilize brief, dynamic stimuli that provide audio-visual information with the aim of improving efficiency in the rapid integration of information and the perception of configural changes that signal specific emotions. Moreover, individuals with ASD would be well served by intervention programs designed for practice in the detection of conflict between social cues (e.g., incongruent facial affect, prosody, semantic content). In addition, consideration of both response time and accuracy findings suggest that individuals with ASD in this study were both impacted by a speed-accuracy trade-off and that they were unaware of their inaccuracies. Individuals with ASD may therefore benefit from interventions that providing immediate feedback about performance and allow for the repetition of stimuli to support improved accuracy of responding. Finally, the evidence for particular difficulties that emerged during a relatively more complex task (i.e., both congruent and incongruent movies together) emphasizes the importance of designing social skills interventions that more closely

simulate real-world social interactions. Such interventions might require emotion or social judgments under a cognitive load or by otherwise increasing the complexity of stimulus context (e.g., distraction). One potential framework for such interventions might be that of programs aimed at supporting executive functioning skills (e.g., cognitive flexibility, divided or alternating attention tasks).

4.5 Limitations and Future Directions

There are some limitations to the current study that are important to consider and recommendations for future studies are provided. The sample included in this study consisted of a small, relatively homogenous group of individuals with ASD. The current study should be replicated in a larger sample and, to the extent possible while ensuring task understanding, with individuals demonstrating more profound social impairments. The clinical group in the current study consists of relatively high-functioning individuals with ASD, and the extent to which these findings can be extended to lower-functioning individuals with ASD is quite limited. In addition, the inclusion of a broader age range would allow for comparison across age sub-groups. Such comparison may provide information as to whether there is a developmental delay in emotion perception for the individuals with ASD, wherein they might be expected to improve to the level of TD individuals at an older age.

Findings from the current study provide important information about performance on a novel emotion perception task that includes brief stimulus duration, audio-visual affective information, and dynamic affective cues. Although these are three important aspects of emotion perception that extend the findings of available literature, future studies should attempt to separate the unique effect that each of these components has on individual performance. For example, the short duration of the movie clips also limited the amount of information that could

be gathered about fixation pattern that emerges over a period of time or possible improvements in accuracy following deliberation. Developing DAVE stimuli of a longer duration or allowing for the repetition of stimuli may address this limitation. In order to support the interpretation of the dynamic nature of the DAVE stimuli, a comparison of fixation patterns elicited during these novel dynamic stimuli should be made with previously validated static photos. Furthermore, examination of the DAVE stimuli as visual-only stimuli would provide additional understanding of how the visual scanning of dynamic faces differs from that of static faces without the added complexity of the auditory component; this will be important for teasing apart differences due to dynamic compared to static visual stimuli, and audio-visual compared to visual-only stimuli.

Individuals with ASD in this study demonstrated subtle, albeit significant, differences in accuracy of emotion identification and fixation patterns on a relatively more complex task. Given that these differences emerged during a controlled experimental task, these individuals would be expected to demonstrate difficulties in a real-world environment. However, the generalizability of these findings is somewhat limited. Specifically, the dimensions of the DAVE stimuli were relatively small during presentation in this study (i.e., limited field of view) and may have required minimal scanning of the stimulus. In addition, stimuli are limited to the actor's face and exclude additional cues from body language or environment. In order to support the comparison of this task to the abilities required in face-to-face social interactions, future studies should consider using stimuli of large size, presenting affective cues from additional sources, and including additional environmental stimuli.

Several inferences were made about the early integration of audio-visual information and the detection of conflict. These arguments could be further explored using neuroimaging or electrophysiological studies. Hypotheses regarding slowed processing of emotional information

in the ASD group should be explored further by assessing whether processing speed differences are limited to emotional tasks or they result from reduced processing speed more broadly. Future studies would benefit from the addition of both visuo-motor and cognitive processing speed measures to the neuropsychological battery. In addition, behavioral studies could evaluate participants' awareness of conflict (i.e., does the emotion in the face match the emotion in the voice?). Preliminary studies conducted with the DAVE stimuli and TD individuals also suggested that fixation patterns varied by facial emotion. Future studies should examine accuracy and face processing strategies separately for each of the four emotions presented in this study (i.e., happy, angry, fearful, neutral). It was also argued that early detection of conflict in the TD group resulted in increased attention to the most affectively salient modality (i.e., face) and region (i.e., core). Therefore, future studies should compare fixation patterns during trials in which there was a visual bias and those where there was an auditory bias. Such analyses would indicate whether visual fixation patterns differed according to modality bias.

REFERENCES

- Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience, 14*(8), 1264-1274.
- Adolphs, R., Tranel, D., & Damasio, A.R. (1998). The human amygdala in social judgment. *Nature, 393*(6683), 470-474.
- Allison, T., Puce, G., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trend in Cognitive Science, 4*, 267-278.
- Ambadar, Z., Schooler, J.W., & Cohn, J.F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16*(5), 403-410.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental Disorders* (4th edition, Text Revision). Washington, D.C.: Author.

- Ames, C.S., & Jarrold, C. (2007). The problem with using eye-gaze to infer desire: A deficit of cue inference in children with autism spectrum disorder? *Journal of Autism and Developmental Disorders*, *37*, 1761-1775.
- Applied Science Laboratories (2005). ASL Eye-Trac 6000 Manual. Applied Science Group: Bedford, MA.
- Aviezer, H., Hassin, R.R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies in the malleability of emotion perception. *Psychological Science*, *19*(7), 724-732.
- Aylward, E.H., Park, J.E., Field, K.M., Parsons, A.C., Richards, T.L., Cramer, S.C., & Meltzoff, A.N. (2005). Brain activation during face perception: evidence of developmental changes, *Journal of Cognitive Neuroscience*, *17*(2), 308-319.
- Bal, E., Harden, E., Lamb, D., Van Hecke, A.V., Denver, J.W., & Porges, S.W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders*, *40*(3), 358-370.
- Baldwin, D.A., & Moses, L.J. (1996). The ontogeny of social information gathering. *Child Development*, *67*(5), 1915-1939.
- Banks, M.S., McManus, S.M., Mumaw, M.A., King, T.Z., & Robins, D.L. (2008, February). Reaction time disparities between individuals with autism spectrum disorders and typically-developing controls in response to congruent and incongruent emotional movies. Poster presented at the annual meeting of the International Neuropsychological Society, Kailua-Kona, Hawaii.

- Bar-Haim, Y.C., Shulman, C., Lamy, D. & Reuveni, A. (2006). Attention to eyes and mouth in high-functioning children with autism. *Journal of Autism and Developmental Disorders*, *36*(1), 131-137.
- Baron-Cohen, S., Ring, H.A., Wheelwright, S., Bullmore, E.T., Brammer, M.J., Simmons, A., & Williams, S.C. (1999). Social intelligence in the normal and autistic brain: and fMRI study. *European Journal of Neuroscience*, *11*(4), 1891-1898.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome, high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5-17.
- Batty, M., Meaux, E., Wittmeyer, K., Roge, B., & Taylor, M.J. (2011). Early processing of emotional faces in children with autism: An event-related potential study. *Journal of Experimental Child Psychology*, *109*, 430-444.
- Bauminger, N., & Kasari, C. (2000). Loneliness and friendship in high-functioning children with autism. *Child Development*, *71*(2), 447-456.
- Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., & Minshew, N. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia*, *44*, 110-129.
- Benton, A. (1994). *Face Recognition*. Los Angeles: Western Psychological Services.
- Bonner, M.J., Hardy, K.K., Willard, V.W., Anthony, K.K., Hood, M., & Gururangan, S. (2008). Social functioning and facial expression recognition in survivors of pediatric brain tumors. *Journal of Pediatric Psychology*, 1-11.

- Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice Processing Abilities in Children with Autism, Children with Specific Language Impairments, and Young Typically Developing Children. *Journal of Child Psychology and Psychiatry*, *41*(7), 847-857.
- Bradshaw, J.L., & Wallace, G. (1971). Models for the processing and identification of faces. *Perception & Psychophysics*, *9*(5), 443-448.
- Bradshaw, J., Shic, F., & Chawarska, K. (2011). Face-specific recognition deficits in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*.
- Brunner, D.L., & Seung, H. (2009). Evaluation of the efficacy of communication-based treatments for autism spectrum disorders. *Communication Disorders Quarterly*, *31*(1), 15-41.
- Bryant, G.A., & Barrett, H.C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, *8*, 135-148.
- Bukach, C.M., Bub, D.N., Gauthier, I., & Tarr, M.J. (2006). Perceptual expertise effects are not all or none: spatially limited perceptual expertise for faces in a case of prosopagnosia. *Journal of Cognitive Neuroscience*, *18*(1), 48-63.
- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *TRENDS in Cognitive Sciences*, *10*(4), 159-166.
- Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of simple and complex emotions in non-retarded children with autism. *Journal of Child Psychology and Psychiatry*, *33*, 1169-1182.
- Carter, E.J., & Pelphrey, K.A. (2006). School-aged children exhibit domain-specific responses to biological motion. *Social Neuroscience*, *1*, 396-411.

- Centers for Disease Control and Prevention (2012). Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. *MMWR Surveillance Summary*, 61 (SS03), 1-19.
- Chamberlain, B., Kasari, C., & Rotheram-Fuller, E. (2007). Involvement or isolation? The social networks of children with autism in regular classrooms. *Journal of Autism and Developmental Disorders*, 37, 230-242.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ, Lawrence Erlbaum Associates.
- Collet, C., Vernet-Maury, E., Delhomme, G. & Dittmar, A. (1997). Autonomic nervous system response patterns specificity to basic emotions. *Journal of the Autonomic Nervous System*, 62(1-2), 45-57.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242, 126-135.
- Conturo, T.E., Williams, D.L., Smith, C.D., Gultepe, E., Akbudak, E., & Minshew, N.J. (2008). Neuronal fiber pathway abnormalities in autism: An initial MRI diffusion tensor tracking study of hippocampo-fusiform and amygdalo-fusiform pathways. *Journal of the International Neuropsychological Society*, 14, 933-946.
- Critchley, H.D., Daly, E.M., Bullmore, E.T., Williams, S.C., Van Amelsvoort, T., Robertson, D.M., et al. (2000). The functional neuroanatomy of social behaviour: Changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain: A Journal of Neurology*, 123, 2203-2212.

- Dalton, K., Nacewicz, B., Johnstone, T., Schaefer, H., Gernsbacher, M.A., Goldsmith, H.H., Alexander, A., & Davidson, R. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, 8, 519-526.
- Davies, S., Bishop, D. Manstead, A.S., & Tantan, D. (1994). Face perception in children with autism and Asperger's syndrome. *Journal of Child Psychology*, 35(6), 1033-1057.
- Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorders, *Development and Psychopathology*, 20, 775-803.
- Dawson, G., & Bernier, R. (2007). Social brain circuitry in autism. In D. Coch, G. Dawson, & K. Fisher (Eds.), *Human behavior and the developing brain, 2nd Ed., Atypical Development*, New York: Guilford Press.
- de Gelder, B., & Vroomen, J. (2000). The Perception of emotions by ear and by eye. *Cognition and Emotion*, 14(3), 289-311.
- de Gelder, B., Vroomen, J., & Bertelson, P. (1998). Upright but not inverted faces modify the perception of emotion in the voice. *Current Psychology of Cognition*, 17(4-5), 0121-1031.
- de Haan, M., & Nelson, C.A. (1997). Recognition of the mother's face by six-month-old infants: A neurobehavioral study. *Child Development*, 68(2), 187-210.
- Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115. 107-117.
- Dodge, K., Lansford, J., Burks, V., Bates, J., Pettit, G., Fontaine, R., et al. (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development*, 74, 374-393.

- Dziobek, I., Bahnemann, M., Convit, A. & Heekeren, H.R. (2010). The role of the fusiform-amygdala system in the pathophysiology of autism. *Archives of General Psychiatry*, 67(4), 397-405.
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99, 550-553.
- Ekman, P. & Friesen, W.V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Oxford: Prentice-Hall.
- Ethofer, T. Anders, S., Erb, M., Droll., C. Ryen., L. Sauer, R., et al. (2006). Impact of voice on emotional judgment of faces: An event-related fMRI study. *Human Brain Mapping*, 27(9), 707-714.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about Face perception? *Psychological Review*, 105 (3), 482-498.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Gauthier, I. & Tarr, M.J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, 37, 1673-1682.
- Gauthier, I., Tarr., M.J., Anderson, A.W., Skudlarski, P., & Gore, J.C. (1999). Activation of the middle fusiform ‘face area’ increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568-573.
- Gepner, B., deGelder, B. & deSchonen, S. (1996). Face processing in autistics: evidence for a generalized deficit? *Child Neuropsychology* 2(2), 123-139.
- Geschwind, D.H. & Levitt, P. (2007) Autism spectrum disorders: developmental disconnection syndromes. *Current Opinion in Neurobiology*, 17(1), 103-111.

- Grelotti, D.J., Gauthier, J., & Schultz, R.T. (2002). Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Developmental Psychobiology*, *40*, 213-225.
- Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of gaze behavior research* (pgs. 523-533). Amsterdam: Elsevier.
- Gross, T.F. (2005). Global-local precedence in the perception of facial age and emotional expression by children with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders*, *35*(6), 773-785.
- Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal Bias in Recognition of Facial Emotions in Children with Asperger Syndrome. *Journal of Child Psychology and Psychiatry*, *41*(3), 369-379.
- Hall, G.B., Szechtman, H., & Nahmias, C. (2003). Enhanced salience emotion recognition in autism: a PET study. *American Journal of Psychiatry*, *160*, 1439-1441.
- Haviland, J.M., & Lelwica, M. (1987). The induced affect response – 10-week-old infants responses to three emotion expressions. *Developmental Psychology*, *23*(1), 97-104.
- Haxby, J.V., Hoffman, E.A., & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, *4*, 223-233.
- Hecht, D., & Reiner, M. (2008). Sensory dominance in combinations of audio, visual, and haptic stimuli. *Experimental Brain Research*, *193*, 307-314.
- 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*, 1004-1012.

- Hesling, I., Dilharreguy, B., Peppe, S., Amirault, M., Bouvard, M., & Allard, M. (2010). The integration of prosodic speech in high functioning autism: a preliminary fMRI study, *PloS ONE*, 5(7), e11571.
- Hietanen, J., Leppanen, J.M., Illi, M., & Surakka, V., (2004). Evidence for the integration of audiovisual emotional information at the perceptual level of processing. *European Journal of Cognitive Psychology*, 16(6), 769-790.
- Hobson , R.P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: coordinating faces and voices. *Psychological Medicine*, 18, 911-923.
- Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli. *Memory and Cognition*, 4(2), 176-185.
- Horley, K., Williams, L.M., Gonsalvez, C., & Gordon, E. (2003). Social phobics do not see eye to eye: A visual scanpath study of emotional expression processing. *Journal of Anxiety Disorders*, 17(1), 33-44.
- Horley, K., Williams, L.M., Gonsalvez, C., & Gordon, E. (2004). Face to face: visual scanpath evidence for abnormal processing facial expressions in social phobia. *Psychiatry Research*, 127(1-2), 43-53.
- Howlin, P. (2005). The effectiveness of interventions for children with autism. *Neurodevelopmental Disorders*, 10, 101-119.
- Howlin, P., Magiati, & Charman (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities*, 114(1), 23-41.

- Humphreys, K., Minshew, N., Leonard, G.L. & Berhmann, M. (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia*, 45(4), 685-695.
- Itier, R.J., Alain, C., Sedore, K., & McIntosh, A.R. (2007). Early face processing specificity: It's in the eyes! *Journal of Cognitive Neuroscience*, 19(11), 1815-1826.
- Itier, R.J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Behavioral Reviews*, 33(6), 843-863.
- Izard, C.E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, 99(3), 561-565.
- Jarvis, B. (1999). DirectRT Precision Timing Software. Empirisoft Software.
- Jones, W., Carr, K., & Klin, A. (2009). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddler with autism spectrum disorder. *Archives of General Psychiatry*, 65(8), 946-954.
- Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry*, 44(4), 529-542.
- Kanwisher, N., McDermott, & Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11), 4302-4311.
- Kanwisher, N. & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B*, 361, 2109-2128.
- Karatekin (2007). Eye-Tracking studies of normative and atypical development. *Developmental Review*, 27, 283-348.

- Kleinhaus, N.M., Richards, T., Serling, L., Stegbauer, K.C., Mahurin, R., Johnson, L.C., et al. (2008). Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain, 131*, 1000-1012.
- Klin, A., Jones, W., Schultz, R. T., & Volkmar, F. R. (2005). The Enactive Mind – From Actions to Cognition: Lessons from Autism. In D. Cohen and F. Volkmar (Eds.), *Handbook of Autism and Pervasive Developmental Disorder* (3rd edition, pp. 682-703). NY: John Wiley and Sons, Inc.
- 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*(9), 809-816.
- Klennert, M.D. (1984). The regulation of infant behavior by maternal facial expression. *Infant Behavior & Development, 7*(4), 447-465.
- Koshino, H., Kana, R.K., Keller, T.A., Cherkassky, V.L., Minshew, N.J., & Just, M.A. (2008). fMRI investigation of working memory for faces in autism; Visual coding and underconnectivity with frontal areas. *Cerebral Cortex, 18*, 289-300.
- Kohler, C.G., Turner, T., Stolar, N.M., Bilker, W.B., Brensinger, C.M., Gur, R.E., & Gur, R.C. (2004). Differences in facial expression of four universal emotions. *Psychiatry Research, 128*, 235-244.
- Koppen, C., Alsius, A., & Spence, C. (2008). Semantic congruency and the Colavita visual dominance effect. *Experimental Brain Research, 184*(4), 533-546.
- Kreibig, S.D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*(3), 394-421.

- LaBar, K.S., Crupain, M.J., Voyvodic, J.T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, *13*(10), 1023-1033.
- Ladd, G.W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: predictors of children's early school-adjustment? *Child Development*, *61*, 1081-1100.
- Ladd, G. (2006). Peer Rejection, Aggressive or Withdrawn Behavior, and Psychological Maladjustment from Ages 5 to 12: An Examination of Four Predictive Models. *Child Development*, *77*, 822-846
- Ladd, G.W., Herald-Brown, S.L., & Resier M. (1997). Does chronic classroom peer rejection predict the development of children's classroom participation during the grade-school years? *Child Development*, *79*, 1001-1015.
- Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language, Hearing, and Resolution*, *42*, 526-539.
- Leppanen, J.M. & Hietanen, J.K. (2003). Emotion recognition and social adjustment in school-aged girls and boys. *Scandinavian Journal of Psychology*, *42*(5), 429-435.
- Lindner, J.L. & Rosen, L.A. (2006). Decoding of emotion through facial expression, prosody, and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *36*(6), 769-777.
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism Diagnostic Observation Schedule – WPS (ADOS – WPS)*. Los Angeles, CA: Western Psychological Services.

- Loveland, K. A., Tunali-Kotoski, B. Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A., & Gibbs, M. C. (1995). Emotion recognition in autism: Verbal and nonverbal information. *Development and Psychopathology, 9*, 579-593.
- Luria, S.M., & Strauss, M.S. (1978). Comparison of eye movements over faces in photographic positives and negatives. *Perception and Psychophysics, 2*, 547-552.
- Mackworth, N.H., & Bruner, J.S. (1970). How adults and children search and recognize pictures. *Human Development, 13*, 149-177.
- Manor, B.R., Gordon, E., Williams, L.M., Rennie, C.J., Bahramali, H., Latimer, R.J., & Meares, R.A. (1999). Eye movements reflect impaired face processing in patients with schizophrenia. *Society of Biological Psychiatry, 46*, 963-969.
- Massaro, D.W. (1998). *Perceiving talking faces: From speech perception to a behavior principle*. Cambridge, MA: MIT Press.
- Massaro, D.W., & Egan, P.B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review, 3*(2), 215-221.
- McManus, S.M., Banks, M.S., Mumaw, M.A., King, T.Z., & Robins, D.L. (February, 2008). Face bias during and audio-visual emotion perception task is absent in Autism Spectrum Disorders. Poster presented at the annual meeting of the International Neuropsychological Society, Waikoloa, HI.
- McManus, S.M., Robins, D.L., Tone, E.B., & Washburn, D.A. (2009). *Gaze fixations during the perception of visual and auditory affective cues*. Unpublished master's thesis, Georgia State University, Atlanta, GA.
- Mertens, I., Siegmund, H., & Grusser, O. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia, 31*, 989-998.

- Mondloch, C.J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*, 553-566.
- Mongillo, E.A., Irwin, J.R., Whalen, D.H., Klaiman, C., Carter, A.S., & Schultz, R.T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 1349-1358.
- Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 27*(5), 657-670.
- Munson, J., Dawson, G., Abbott, R., Faja, S., Webber, S.J., Friedman, S.D., Shaw, D., Artru, A., & Dager, S.R. (2006). Amygdalar volume and behavioral development in autism. *Archives of General Psychiatry, 63*(3), 686-693.
- Neuhaus, E., Beauchaine, T.P., & Bernier, R. (2010), Neurobiological correlates of social functioning in autism. *Clinical Psychology Review*, Article in Press.
- Norbury, C.F., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders, *Journal of Child Psychology and Psychiatry, 50*(7), 834-842.
- Noton, D. & Stark, L (1971). Scanpaths in eye movements during pattern perception. *Science, 171*, 308-311.
- Nowicki, S. Jr. (2004). A test and manual for the Diagnostic Analysis of Nonverbal Accuracy 2. Unpublished manuscript, Emory University, Atlanta, GA.
- Nowicki, S. & Carton, J. (1993). The measurement of emotional intensity from facial expressions. *The Journal of Social Psychology, 133*(5), 749-750.

- Nowicki, S. & Duke, M.P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of nonverbal Behavior*, 18(1), 9-35.
- Nusseck, M., Cunningham, D.W., Wallraven, C., & Bulthoff, H.H. (2008). The contribution of different facial regions to the recognition of conversational speech. *Journal of Vision*, 8(8), 1-23.
- O'Connor, K., Hamm, J.P., & Kirk, I.J. (2005). The neurophysiological correlates of face processing in adults and children with Asperger's syndrome. *Brain and Cognition*, 59, 82-95.
- Ortony, A. & Turner, T.J. (1990). What's basic about emotions? *Psychological Review*, 97, 315-331.
- Osterling, J.A., Dawson, G., & Munson, J.A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, 14(2), 239-251.
- Ozonoff, S., Pennington, B.F., & Rogers, S.J. (1990). Are there emotion perception deficits in young autistic children. *Journal of Child Psychology and Psychiatry and Allies Disciplines*, 31(3), 343-261.
- Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *The Behavioral and Brain Sciences*, 5, 407-467.
- Pell, M.D. (2005). Nonverbal emotion priming: Evidence from the 'facial affect decision task'. *Journal of Nonverbal Behavior*, 29(1), 45-73.
- Pelphrey, K., Adolphs, R., & Morris, J.P. (2004). Neuroanatomical substrates of social cognition dysfunction in autism. *Mental Retardation and Developmental Disabilities*, 10, 259-271.

- Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual Scanning of Faces in Autism. *Journal of Autism and Developmental Disorders*, 32(4), 249-261.
- Pourtois, G., de Gelder, B., & Vroomen, J. (2000). The time-course for intermodal binding between seeing and hearing affective information. *Neuroreport: For Rapid Communication of Neuroscience Research*, 11(6), 1329-1333.
- Price, C.J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, 197, 335-359.
- Price, J.L. (2006) Prefrontal cortex, In S. Moldin & J. Rubenstein (Eds.), *Understanding autism: from basic neuroscience to treatment (pp. 205-225)*, US: Taylor & Francis Group.
- Reynolds, C.R. & Kamphaus, R.W. (2004). BASC-2 Manual. AGS Publishing: Circle Pines, MN.
- Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinates of facial appearance. *Perception*, 17, 43-63.
- Rhodes, G., Brake, S., & Atkinson, A.P. (1993). What's lost in inverted faces? *Cognition*, 47, 25-57.
- Robins, D.L., Hunyadi, E., & Schultz, R.T., (2009). Superior temporal activation in response to dynamic audio-visual emotion cues. *Brain and Cognition*, 69, 269-278.
- Robins, D.L., & Schultz, R.T. (2004). Cross-modal integration of emotional information. Paper presented at the International Neuropsychological Society, February, 2004, Baltimore, MD.
- Rogers, S.J., & Vismara, L.A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology*, 3, (1), 8-38.

- Rump, K. M., Minshew, N. J., Giovannelli, J. L., & Strauss, M. S. (2009). The development of emotion recognition in individuals with autism. *Child Development, 80*, 1434-1447.
- 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.
- Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism Diagnostic Interview – Revised*. Los Angeles, CA: Western Psychological Services.
- Saracino, J., Noseworthy, J., Steiman, M., Reisinger, L., & Fombonne, E. (2010). Diagnostic and assessment issues in autism surveillance and prevalence. *Journal of Developmental and Physical Disabilities, 22*, 317-330.
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumara, M. (2004). Enhanced neural activity in response to dynamic expressions of emotions: An fMRI study. *Brain Research, 20*, 81-91.
- Schultz, R.T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*, 125-141.
- Schumann, C.M., Hamstra, J., Goodlin-Jones, B.L., Lotspeich, L.J., Kwon, H., Buonocore, M.H., Lammers, C.R. Reiss, A.L., & Amaral, D.G. (2004). The amygdala is enlarged in children but not adolescents with autism: the hippocampus is enlarged in all ages. *The Journal of Neuroscience, 24*(28), 6392-6401.
- Schwarzer, G., Huber, S., & Dummler, T. (2005). Gaze behavior in analytical and holistic face processing. *Memory & Cognition, 33*(2), 344-354.

- Seashore, C., Lewis, D., & Saetveit, J. (1960). *Seashore Rhythm Test*. In: Seashore, *Measures of Musical Talents, Revised Manual*. New York: The Psychological Corporation.
- Smith, M.L., Cottrell, G.C., Gosselin, F. & Shyns, P.G. (2005) Transmitting and decoding facial expressions. *Psychological Science, 16*(3), 184-189.
- Stacey, P. C., Walker, S. & Underwood, J.D. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology, 96*, 407-422.
- Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46*, 225-245.
- Tarr, M.J. & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level Visual processing automatized by expertise. *Nature Neuroscience, 3*(8), 764-769.
- Teunisse, J. P. & de Gelder, B. (2003). Face processing in adolescents with autistic disorder: The inversion and composite effects. *Brain and Cognition, 52*, 285-294.
- Thomas, L.A., De Bellis, M.D., Graham, R., & LaBar, K.S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science, 10*(5), 547-558.
- Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology, 17*, 257-279.
- Tracy, J.L. & Robins, R.W. (2008). The automaticity of emotion recognition. *Emotion, 8*(1), 81-95.
- Valentine, T. (1988) Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471-491.

- Vandell, D.L., & Hembree, S.E. (1994). Peer social status and friendship: Independent contributors to children's social and academic adjustment. *Merrill-Palmer Quarterly*, 40, 461-470.
- Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive, Affective, & Behavioral Neuroscience*, 1(4), 382-387.
- Vuilleumier, P. & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, 45, 174-194.
- Vuilleumier, P., Richardson, M.P., Armony, J.L., Dolan, R.J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 7(11), 1271-1278.
- Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, 6, 313-326.
- Wang, K., Hoosain, R., Lee, T.M., Fu, J., & Yang, Renmin (2006). Perception of six basic emotional expressions by the Chinese. *Journal of Cross-Cultural Psychology*, 37(6), 623-629.
- Webb, S.J., Jones, E.J. Merkle, K., Venema, K., Greenson, J., Murias, M., & Dawson, G. (2011). Developmental changes in the ERP responses to familiar faces in toddlers with autism spectrum disorders versus typical development. *Child Development*, 82(6), 1868-1886.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

- Whalen, P.J., Kagan, J., Cook, R.K., Davis, F.C., Kim, H., et al., (2004). Human amygdala responsivity to masked fearful eye whites. *Science*, *306*, 2061-2062.
- Wheaton, K.J., Thompson, J.C., Syngienotis, A., Abbott, D.F., & Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage*, *22*(1), 277-288.
- White, M. (1999). Representation of facial expressions of emotion. *American journal of Psychology*, *112*(3), 371-381.
- White, M. (2000). Parts and wholes in expression recognition. *Cognition and Emotion*, *14*, 39-60.
- Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141-145.
- Yin, R.K. (1970). Face recognition by brain-injured patients: A dissociable ability? *Neuropsychologia*, *8*, 395-402.
- Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, Y.S., & Boddaert, N. (2006) Autism, the superior temporal sulcus and social perception. *TRENDS in Neuroscience*, *29*(7), 359-366.

APPENDIX

The study will begin shortly... message will appear on the screen → Task Instructions

“I have placed this plastic/gel cover over the keyboard so that only four keys are exposed (*hold the keyboard up for the participant to see*) and those are the keys you will use to respond. After the ‘study...’ message goes away, a reminder of the directions for this task will appear on the screen. You are to pay attention to each video and answer the question, “How does this person feel?” Before each movie begins, a crosshair (plus sign) will appear in the middle of the screen (*point to the center of the screen as you say this*). That is a cue for you to look to that point before each video clip begins. Each video clip will play for approximately 2 seconds, after which four emotion choices will appear on the screen. The choices will be Happy, Angry, Fearful, Neutral” (*point to the positions on the screen as you say this*) and these choices will line up spatially (from left to right) with the keys on the keyboard. You can just rest four fingers on the exposed keyboard keys and then you won’t need to see the keyboard to make your response. As soon as you make your response, it will automatically advance to the next video clip. If you do not respond in the 5-second time allowed by the program, you will see a message that says ‘Please try and respond faster’. Do you have any questions?”

Practice Run:

“Because this is practice, you will only see pictures”

Task 2:

“The instructions for this next task will be exactly the same as before”