EXAMINING THE NEUROLOGICAL UNDERPINNINGS OF ATYPICAL MULTISENSORY PERCEPTION IN AUTISM SPECTRUM DISORDERS

STEPHANIE LAVOIE

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CLINICAL-DEVELOPMENTAL PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

September 2016

Abstract

Individuals with Autism Spectrum Disorder (ASD) have difficulties with multisensory perception, which would have a significant impact on their cognitive, sensory, language, and social development. The current collection of studies sought to understand the neurological mechanisms underlying this difficulty with multisensory perception across temporal synchrony variations using functional Magnetic Resonance Imaging (fMRI). The nature of these multisensory processing deficits was further explored by delineating social and linguistic processing. The overall goal was accomplished by examining multisensory processing in three studies. The first study evaluated and compared various methods of identifying brain regions responsible for multisensory integration in 17 young adults without ASD. This first study found that a newly proposed temporal synchrony method, which compares neural responses to temporally synchronous and asynchronous audiovisual stimuli, was more theoretically valid and more empirically tenable than other previously used methods. In the second study, the temporal synchrony method was used to compare responses to multisensory stimuli across social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic conditions in individuals with (*n* = 15) and without (*n* = 17) ASD using fMRI. The third study explored whole-brain patterns of activity involved in multisensory integration, using a multivariate fMRI analysis approach (partial least squares: PLS) with the same participant groups as study two. Taken together the results of studies two and three reveal that young adults with ASD do not process multisensory stimuli in the same way as young adults without ASD. When using targeted contrasts in study two, individuals with ASD displayed either a lack of multisensory integration or an opposite pattern of response to synchrony variants of multisensory information relative to the group without ASD. Further, the results of study three suggested that when examining synchronous and asynchronous multisensory stimuli, individuals with ASD do not engage the same social- and language-specific networks that were engaged by individuals without ASD. Based on the results of these studies, a novel hypothesis was proposed to explain the differential response profiles for individuals with and without ASD: the multisensory catalyst nodes hypothesis.

Dedication

To my loving and supportive family, your encouragement and understanding has helped me reach this goal. In particular, to my mother, Lynne Mckay, who has always been my number one fan and cheerleader.

To my husband, Kevin Lavoie, your unwavering support, love, and motivation has kept me going. Thank you for believing in me.

Acknowledgements

I would like to express my deepest appreciation to my supervisors, Dr. W. Dale Stevens and Dr. James Bebko, for their time, guidance, patience, and dedication. Thank you for encouraging and challenging me.

My sincerest gratitude to Dr. Gary Turner and Dr. Christine Till for their thoughtful comments and feedback during the revision process. Thank you to Dr. Mary Desrocher, Dr. Lauren Sergio, and Dr. Terry Bennett for their insightful questions that pushed me to deepen my knowledge and widen my perspective.

To my colleagues, collaborators, and members of the Bebko Lab and Cognitive Aging Neuroscience & Neuro-Intervention Group, thank you for your assistance and support. In particular, thank you to Dr. Ryan Stevens, Magali Segers, Naail Khan, Joseph Viviano, and Areeba Adnan. I would also like to thank Dr. Carly McMorris, Lisa Alli, Dr. Michelle Fardella, and my cohort for their friendship and encouragement throughout the completion of this project.

Most importantly, thank you to the individuals and families who dedicated their time to this study.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

A - Audio

ADI-R - Autism Diagnostic Interview-Revised

ADOS - Autism Diagnostic Observation Scheduled

AES - Anterior Ectosylvian

AFNI - Analysis of Functional Neuroimages

ASD - Autism Spectrum Disorder

BOLD - Blood Oxygen-Level Dependent

DSM - Diagnostic and Statistical Manual of Mental Disorders

EPI - Echo-Planar Images

ERP - Event-Related Potentials

fMRI - Functional Magnetic Resonance Imaging

M - Multisensory

MEG - Magnetoencephalography

PET - Positron Emission Tomography

PLS - Partial Least Squares

pSTS - Posterior Superior Temporal Sulcus

rLS - Rostral Lateral Suprasylvian

ROI - Region of Interest

SC - Superior Colliculus

SNR - Signal to Noise Ratio

STG - Superior Temporal Gyrus

STS - Superior Temporal Sulcus

URPP - Undergraduate Research Participation Pool

V - Visual

WASI - Wechsler Abbreviated Scales of Intelligence

Examining the Neurological Underpinning of Atypical Multisensory Perception in Autism Spectrum Disorders

Our environment is comprised of an abundance of multisensory information, meaning that we are constantly exposed to multiple simultaneous sensory experiences. In order to make sense of our environment, our perceptual systems are tasked with combining and differentiating these sensory experiences into distinct unitary events. Multisensory perception is the process of combining these multiple sensory experiences into a coherent perception of a unitary event and discriminating these combined perceptions from other events (Bahrick & Lickliter, 2000; Bahrick & Todd, 2012).

1.1. Multisensory Perception

Multisensory perception starts to form early in infancy and is integral to development. The integration of sensory events is an essential element and foundation for many important cognitive and perceptual processes. Particularly relevant to the current thesis and for understanding atypical development, multisensory perception has been found to be a key element in language development and social development (Bahrick, 2010; Bahrick & Lickliter, 2002; Bahrick & Todd, 2012; Edelman, 1992; Gibson, 1969; Gogate & Bahrick, 1998; Lewkowicz, 2000; Thelen & Smith, 1994). Further, multisensory perception remains an important foundation for cognition beyond infant and child development. Compared to unisensory experiences, multisensory stimuli have been found to lead to faster processing speed (Calvert, Campbell, & Brammer, 2000; Frens & Van Opstal, 1995; Hershenson, 1962; Morrell, 1968), guide selective attention (Bahrick & Lickliter, 2002; Bahrick & Todd, 2012), lead to enhancement of degraded signals (Reisber, McLean, & Goldfield, 1987; Rosenblum, Johnson, & Saldana, 1996; Sumby &

Pollack, 1954), and help one understand social interactions (Bahrick & Todd, 2012). Given the importance of multisensory perception on human development, cognition, and perception, the consequences of atypical multisensory perception could have wide ranging implications on human functioning.

Through a review of the development, importance, and neurological underpinnings of multisensory perception in typical development, I will lay the foundation for understanding the impact that atypical multisensory perception could have in cognitive, language, and social development. The central focus of the current set of studies is to explore how altered neurological systems for multisensory perception may be a foundational element of autism spectrum disorder (ASD).

Multiple terms have been used to describe the process of combining multiple sensory experiences into a coherent perception of a unitary event and discriminating it from other events, such as crossmodal, multimodal, polysensory, multisensory, intersensory, heteromodal, intermodal, and supramodal (Calvert, 2001). Each of these terms may offer varying connotations for differing research fields. In order to reduce ambiguity and/or overlap of multiple meanings across research fields, the current paper will use the term “multisensory perception” to refer to the phenomenological experience of perceiving causally related input from multiple sensory systems into a unitary precept, and the term “multisensory integration” to refer to the neural processes that integrate the incoming neural signals from multiple sensory systems, which underlie this form of perception.

1.1.1. Development of Multisensory Perception

The process of differentiation for explaining the development of multisensory perception has received much support. The central tenant of differentiation theory is that infants are born with perceptual unity and, through development, learn to discriminate increasingly complex forms of relations (Gibson, 1969).

The key to integrating information from multiple modalities into one event and differentiating it from unrelated events is through redundant information (Walker-Andrews, Bahrick, Raglione, & Diaz, 1991). This redundant information occurs through amodal properties, which are attributes that are common across several sensory modalities, such as tempo, duration, and intensity (Bahrick, Lickliter, & Flom, 2004). For example, when someone is bouncing a ball, we see the visual stimulus and hear the auditory stimulus of the ball being bounced, and the visual and auditory stimuli would match in tempo and duration. Differentiation theory posits that we integrate information by using amodal properties that exist across modalities to discriminate multisensory events that co-occur (Gibson, 1966; Gibson, 1969, Mendelson, 1979).

The development of responsiveness to amodal properties has been tracked using preferential looking paradigms. Typical preferential looking paradigms present two visual displays simultaneously to infants to examine if they look more than at chance levels (greater than 50%) to one of the two displays, thus displaying a preference for that particular visual display. In multisensory perception experiments, the integration of audiovisual information has been examined by presenting two visual displays simultaneously with an auditory track that matches only one of the visual displays. The infants' looking behaviour to the two screens is examined for evidence of a preference for one screen over another, as demonstrated through nonrandom looking patterns. If infants prefer one screen (usually the audio-visual-matched screen), it is interpreted to mean that the infants' looking behaviour is influenced by the integration of the auditory information with the visual displays (Bebko, Weiss, Demark, & Gomez, 2006).

The use of amodal properties for integration of sensory events develops in a hierarchical fashion, with infants developing the ability to integrate based on increasingly complex forms of amodal information. Temporal synchrony is the most global of amodal properties (Bahrick, 2001) and is thought to be one of the first properties used by infants to integrate their sensory experiences (Bahrick, 1987, 1988; Lewkowicz, 1999; 2000). Using a preferential looking paradigm with temporally synchronous and asynchronous audiovisual presentations of a woman speaking a nursery rhyme, Dodd (1979) found that 10-16-week-old infants reliably prefer the temporally synchronous display. Lewkowicz demonstrated that by one-month of age infants are responsive to temporal synchrony (i.e. onset and offset). Lewkowicz (2000) further extended the theory of increasing sensitivity of temporal discrimination by demonstrating the development of responsiveness to increasingly complex temporal amodal properties. Lewkowicz provided a good illustration of these differing temporal properties with the example of a person playing a violin: "As the violinist draws the bow across the strings, an observer can see and hear that the visible actions of the arm are temporally contiguous with the heard actions of the arm and that each discrete up-and-down movement of the arm has a specific duration. In addition, as the violinist repeatedly moves the bow up and down, the observer can see and hear that the action occurs at a certain rate over time and that it has a specific rhythmic quality to it" (p. 286). Infants are found to be responsive to changes in temporal duration between 2-6-months of age, rate by 10-months, and rhythm prior to 12-months (Lewkowicz, 2000). At six-months of age, infants are found to be responsive to temporal microstructure (i.e. that the sound of one large marble hitting the floor is different from multiple small marbles, even if they occur at the same rate and rhythm; Bahrick, 1987). Infants as young as five-months of age will display susceptibility to the McGurk effect, an auditory visual effect where the presentation of incongruent auditory and visual stimuli leads individuals to perceive an auditory sound that matches neither the actual visual nor auditory stimuli (McGurk & MacDonald, 1976; Rosenblum, Schmuckler, & Johnson, 1997). For example, the visual stimulus of a person saying “ga” paired with the auditory stimulus of “ba” will lead an individual to report hearing the sound “da”. The observation of five-month-old infants displaying susceptibility to the McGurk effect demonstrates that the infants are integrating auditory and visual linguistic-based stimuli.

Other forms of more complex amodal properties for discrimination also develop following temporal synchrony. A developmental shift in the perception of speech spectral information in voice and face synchrony seems to occur between three- and four-months of age. Four-month-olds, but not three-month-olds, display a preference for a video of their mother talking (relative to a video of their father talking), while a simultaneous auditory track of the mother’s speech is played (Spelke & Owsley, 1979). In addition, four-month-old infants, but not three-month-olds, will display a preference for a video matching the auditory track of vowel sounds (Kuhl & Meltzoff, 1982, 1984). By four-months of age, infants attend to co-location of audiovisual information, with increasing specificity of co-location by six-months of age (Fenwick & Morrongiello, 1998). Finally, prior to one-year of age, infants are found to be responsive to gender and age components of speech stimuli (Bahrick, Netto, & Hernandez-Keif, 1998; Lewkowicz, 2000; Walker-Andrews et al., 1991).

Supporting this hierarchical development of amodal discrimination, Lewkowicz (2000) found that infants did not display an ability to integrate audiovisual information based on duration and rate if the development of integration based on temporal synchrony was not formed. Thus, temporal synchrony is thought to be the most important element for integration of sensory events and the foundation for the development of multisensory perception (Lekowicz, 1999; 2000; Radeau, 1994; Stein & Meredith, 1993; Welch & Warren, 1980).

1.1.2. Multisensory Perception and Language Development and Comprehension

Through the development of increasing specificity of integration for multisensory experiences, infants start to learn about their world and use increasingly complex forms of differentiation to recognize invariant and arbitrary relationships, such as in language.

The Intersensory Redundancy Hypothesis explains how the detection of amodal information guides learning and attention. Infants have been found to detect amodal properties (such as temporal synchrony) prior to unisensory properties (e.g. colour, shape, pitch, etc.) and to selectively attend to events with amodal information (Bahrick, 1994; Bahrick & Todd, 2012; Bahrick & Lickliter, 2002). In addition, multisensory stimuli that are matched based on amodal information provide more information than unisensory stimuli (Bahrick & Lickliter, 2000). Thus, the recognition of, and attention to, amodal properties plays an important role in regulating perceptual development by drawing infants' attention to global information over local information (Bahrick & Todd, 2012).

Through selectively and preferentially attending to amodal information, infants start to learn and recognize invariant arbitrary relationships, which helps children learn about and coordinate responses to their environment (Bahrick, 1994; Bahrick & Todd, 2012; Iarocci &

McDonald, 2006). Arbitrary relationships between multiple modalities are unpredictable, and provide no prior reasoning for the matching between the sensory stimuli, such as why a dog makes a barking noise.

One of the most relevant examples of invariant arbitrary relationships is language. Words are matched with objects and actions based on arbitrary rules. Learning a new word requires infants and children to match complex arbitrary audio information with a visual stimulus (Gogate, Walker-Andrews, & Bahrick, 2001). Using multisensory audiovisual teaching, parents draw their children's attention to these arbitrary word-object relations. Parents have been found to use temporal synchrony to highlight the novel words they were teaching to five- to eight-month-old pre-lexical infants, a time period when researchers have found temporal synchrony to be particularly important for language development (Gogate & Bahrick, 1998; Gogate, Bahrick, & Watson, 2000). Three to four-month-old infants have been found to make vocal imitations and babbling noises when watching temporally and spectrally synchronous audiovisual information, but not temporally asynchronous and non-voice based spectral auditory information (e.g. pure tones) (Legerstee, 1990; Kuhl & Meltzoff, 1982).

Beyond language development, multisensory perception continues to be important for understanding and coordinating responses to the environment. The visual information provided by facial and lip movements enhances the comprehension of speech when the auditory signal is degraded by background noise (Rosenblum et al., 1996), when the speaker has a foreign accent, and when the speech content is complicated (Reisber, et al., 1987). In fact, as the signal of the auditory speech is degraded, the visual input becomes increasingly more important in the detection of the meaning of speech, and can increase intelligibility by up to 80% (Sumby &

Pollack, 1954). Campbell and Dodd (1980) proposed that individuals are able to use stored auditory and visual memory traces based on previous acoustic experiences in order to help them speech read (i.e. lip read).

While synchronous auditory and visual information leads to a better understanding of the environment, incongruent information leads to slower response times and misperceptions or illusions (Calvert et al., 2000; McGurk & Macdonald, 1976; Sekuler, Sekuler, Lau, 1997; Stein, Meredith, Huneycutt, & McDale, 1989). Therefore, if the audio and the visual components do not match, it can severely decrease the intelligibility of the speech.

1.1.3. Neurological Foundations of Multisensory Integration

1.1.3.1. Animal research. Lewkowicz (2000) proposed a reciprocal relationship between experience with the multisensory environment and the structural and functional organization of the neural system. Much of the evidence for this reciprocal relationship is based in animal research. Animal studies have revealed multisensory specific neurons, which uniquely respond to simultaneous sensory events (Calvert, 2001; Wolf, Gales, Shane, & Shane, 2001). These multisensory specific neurons have specific properties for perceptual integration. In particular the spatial, temporal, and inverse effectiveness rules are important for understanding the specifications guiding integration. The *spatial rule* dictates that sensory stimuli will be integrated more successfully when they originate from the same spatial location (Meredith & Stein, 1986, 1996). The *temporal rule* dictates that sensory stimuli will also be integrated more successfully when they are temporally synchronous (Miller & D'Esposito, 2005; Stein & Meredith, 1993). The final rule guiding neuronal responses to multisensory stimuli is the *rule of inverse effectiveness*. This rule states that when the unisensory components of a multisensory experience are degraded or weak (but not asynchronous) greater multisensory enhancement occurs, and thus, the sensory stimuli are more effectively integrated (Calvert & Thesen, 2004; Meredith & Stein, 1983; Stevenson & James, 2009). When these rules or conditions have been met, multisensory enhancement is observed. Multisensory enhancement is reflected by a significant increase in neuronal activation and firing rate of multisensory neurons in response to multisensory stimuli, which is greater than the neuronal activation and firing rate of these multisensory neurons in response to unisensory stimuli. Some of these multisensory neurons will display a significant increase in neuronal activation and firing rate beyond what would be expected by the summation of stimulation to each sensory modality separately, referred to as a *super-additive* (also called a supra-additive) response (Stein & Meredith, 1993). In contrast, when sensory stimuli are temporally or spatially asynchronous, *response depression* is observed (Calvert et al., 2000; Calvert & Thesen, 2004; Polley, Hillock, Spankovich, Popescu, Royal, & Wallace, 2008).

These neurological findings are analogous to behavioural findings with respect to reactions to multisensory and unisensory stimuli. In particular, the faster response times towards multisensory stimuli parallels the multisensory enhancement of responses found at the neuronal level. Compared to unisensory stimuli, multisensory stimuli have lower thresholds for activation and reduced reaction times (Calvert et al., 2000; Frens & Van Opstal, 1995; Hershenson, 1962; Morrell, 1968).

Many of these properties have been most clearly demonstrated in the superior colliculus (SC), a subcortical structure that integrates sensory information and directs attention (Calvert et al., 2000; Stein & Meredith, 1993). The SC receives visual, auditory, somatosensory, vestibular, and proprioceptive inputs and has cells that project to motor and premotor areas that control the orientation of the eyes (Meredith & Stein, 1986; Wallace & Stein, 1994). Importantly, the SC has neurons specifically designated for the encoding of amodal multisensory information. Further, when the multisensory-specific neurons in the SC are activated, they have been found to elicit orientation and attention behaviours (Stein et al., 1989).

Examination of cortical connections to the SC in cats has revealed that the SC is under the control of two cortical areas: the anterior ectosylvian (AES) and rostral lateral suprasylvian (rLS) sulci. These areas are devoted to core visual, auditory, and somatosensory domains, and project to the multisensory neurons in the SC (Royal, Carriere, & Wallace, 2009). Importantly, research has demonstrated that these two cortical areas play a key role in the development of multisensory enhancement in the SC. When the AES and rLS have been removed in neonatal cats, multisensory enhancement to multisensory stimuli is lost, yet no impact on unisensory stimuli is observed. Furthermore, a disruption of receptive fields that are typically responsive to multisensory stimuli is observed in the SC. Relevant for the current study, this disruption of receptive fields is particularly noticeable for auditory and visual stimuli (Jiang, Jiang, & Stein, 2002). When either the AES or the rLS are removed in adult cats, a large portion of multisensory neurons in the SC fail to demonstrate multisensory enhancement (Jiang et al., 2001). However, the brain seems to be capable of reorganization when either the AES or rLS are removed neonatally; as such, researchers examining the removal of either the AES or rLS have demonstrated that the multisensory enhancement in the SC is largely equivalent to that in intact brains (Jiang et al., 2006; Wilkinson, Meredith, & Stein, 1996). These findings support the idea that neurological development influences the perception of the multisensory environment.

In further support of Lewkowicz's theory that the relationship between the postnatal brain development and multisensory integration is reciprocal, studies have also found that environmental exposure influences neurological development. King and Carlile (1993) demonstrated that ferrets deprived of visual stimulation show anatomical abnormalities in the auditory spatial maps of the SC. Cats reared in environments where auditory and visual stimuli are displayed consistently in a fixed spatial disparity demonstrate marked reorganization of the receptive fields of audiovisual multisensory neurons, and that maximal multisensory enhancement is found when audiovisual stimuli are presented in the same spatial disparity (Wallace & Stein, 2007). Similar findings have also been found for animals reared in environments where the multisensory stimuli are consistently presented with the same temporal disparity. However, this adaptation to the environment is limited to a narrow temporal window. If the temporal disparity between auditory and visual information is pushed beyond 250 milliseconds, then no multisensory enhancement is observed to any stimuli (Polly et al., 2008).

1.1.3.2. Human research. Parallels can be made between animal and human research regarding the neurological mechanisms of multisensory integration. In particular, the suprasylvian cortex in cats is thought to be an analogue to the superior temporal sulcus (STS) in humans. Further, in a closer parallel to humans, audiovisual integration has also been implicated in the STS of rhesus monkeys (Ghazanfar, Chandrasekaran, & Logothetis, 2008). Therefore it is not surprising that in humans, the STS has also been largely implicated in audiovisual integration. While the STS contains the primary auditory cortex, and has thus been identified as critical for auditory processing, it has also been found to respond to silent speech reading (Calvert et al., 1997). The STS has also been identified as an important area for social processing and has been implicated in theory of mind (Pelphrey, Morris, & McCarthy, 2004; Saxe, 2006; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004; Zilbovicious et al., 2006), biological motion (Materna, Dicke, & Thier, 2008; Redcay, 2008), face processing (Haxby, Hoffman, & Gobbini, 2000), gaze direction (Calder et al., 2007), and audiovisual integration (Amedi, von Kriegstein, van Atteveldt, Beauchamp, & Naumer, 2005; Beauchamp, 2005, Calvert, 2001). The posterior portion of the STS (pSTS) in particular has been implicated in audiovisual integration (Hein & Knight, 2008).

While the pSTS is the most implicated structure globally in multisensory integration, including visual-tactile (Banati et al., 2000) and audiovisual integration, Calvert (2001) proposed that linguistic and nonlinguistic audiovisual perception involve different structures and circuitry (this exact circuitry has not been identified, however the different structures that have been identified for linguistic and non-linguistic stimuli are described below). Thus, when examining audiovisual multisensory integration in humans, linguistic and non-linguistic information must be reviewed separately.

1.1.3.2.1. Linguistic stimuli. Studies examining linguistic multisensory integration can further be subdivided into studies that have examined arbitrary integration (e.g. the visual form of a letter matched with the auditory sound of the letter) and invariant amodal integration (e.g. audio of speech matched with the visual mouth movements of speech). With respect to arbitrary integration, Raij, Uutela, and Hari (2000) found activation predominantly in the right temporal-occipital-parietal junction, and bilateral pSTS. However, such studies are limited, as they do not contribute to the understanding of how humans process simultaneous sensory events in the natural environment.

Through examination of invariant amodal aspects of linguistic multisensory integration, researchers have been able to demonstrate that some areas of the human brain demonstrate the same properties of multisensory integration as responses that have been found at the neuronal level in animals to other (nonlinguistic) stimuli. Callan, Callan, Kroos, and Vatikiotis-Bateson (2001) demonstrated the inverse effectiveness principal by showing that the greatest response to multisensory stimuli was observed when the auditory stimulus was degraded by noise.

Researchers have also demonstrated multisensory enhancement to synchronous stimuli and depression to asynchronous stimuli. Calvert et al., (2000) demonstrated that humans display similar super-additive responses as found in the animal studies. Calvert and colleagues compared the audiovisual presentation of a person reading a story to auditory only and visual only presentations. They also used functional magnetic resonance imaging (fMRI) to compare congruent and incongruent presentations of audiovisual information in order to examine areas of the brain that demonstrated supra-additive and sub-additive responses (response depression). The blood oxygen-level dependent (BOLD) responses were found to be supra-additive (but not subadditive) in the middle occipital gyri, occipito-temporal junction, bilateral pSTS, primary auditory cortex localized along Heschl's gyrus, left middle frontal gyrus, and the right inferior parietal lobule. The BOLD response to the multisensory stimuli was actually found to be 30-80% greater than that obtained by the summation of auditory and visual responses alone in the ventral bank of the left STS. Further incongruent audiovisual input reduced BOLD responses in the left pSTS to less than 50% of the unisensory responses, consistent with sub-additive response. Other areas that demonstrated sub-additive responses (but not supra-additive) were in bilateral inferior frontal regions, premotor cortex, right superior temporal gyrus, and anterior cingulate gyrus.

In addition to reports of sub-additive responses to asynchronous audiovisual information, conflicting findings have been reported as to whether or not specific areas of the brain are distinctly activated in response to asynchronous information. Using positron emission tomography (PET), Macaluso, George, Dolan, Spence, and Driver (2004) found that the bilateral fusiform gyrus, right medial lingual gyrus, left STS, and bilateral dorsal occipital cortex responded preferentially to synchronous over asynchronous stimuli, but they did not identify any regions that showed greater activation to asynchronous stimuli. However, using fMRI, Miller and D'Esposito (2005) found that the SC, anterior insula, and anterior intraparietal sulcus (IPS) responded only to asynchronous stimuli and not synchronous stimuli.

Although there is some inconsistency with respect to asynchronous activation, the pSTS seems to be the most consistently involved area in linguistic synchronous multisensory integration. In fact, the pSTS BOLD activation has been found to distinguish individuals who are susceptible from those who are not susceptible to the McGurk effect (Nath & Beauchamp, 2012). Further, transcranial magnetic stimulation, which can temporarily inactivate a brain region, applied to the pSTS turns those who are susceptible to the effect into non-perceivers (Beauchamp, Nath, & Pasalar, 2010), indicating that they are no longer integrating auditory and visual information. Thus, the pSTS must play a key role in the integration of auditory and visual information for linguistic stimuli. It has even been proposed that the pSTS may be uniquely involved in linguistic multisensory integration, as it responds to audiovisual speech, but not to non-intelligible speech-like sounds (Scott, Blank, Rosen & Wise, 2000). However, as explained below, other studies have identified that the pSTS has been involved in non-linguistic multisensory integration.

1.1.3.2.2. ***Non-linguistic stimuli.*** The findings with respect to non-linguistic multisensory integration are less consistent. When using a tone paired with a visual stimulus of a circle, the pSTS was not indicated; instead the right insula, posterior parietal, and prefrontal regions were found to be responsive (Bushara, Grafman, & Hallett, 2001). However, white noise pulses paired with an alternating checkerboard pattern lead to activation in the SC when the auditory and visual stimuli were synchronous; the SC demonstrated the most suppression when the auditory and visual stimuli were temporally asynchronous. While the SC demonstrated both the strongest enhancement to synchronous and depression to asynchronous stimuli, other areas that were implicated included the left pSTS, insula/claustrum bilaterally, right IPS, and frontal regions (Calvert et al., 2001). However, these studies are limited in their generalizability and interpretation given that the stimuli involved arbitrary pairings of sights and sounds that the participants had not previously seen. When using more naturalistic and higher level non-linguistic stimuli, such as a telephone ringing, the pSTS was found to display a supra-additive BOLD response (Beauchamp, Lee, Argall, & Martin, 2004). Stevenson, Geoghegan, and James (2007) proposed that non-linguistic stimuli, such as objects, might elicit super-additive responses only when they are highly degraded. Thus, a super-additive response would be found only when the signal to noise ratio (SNR) is low, such that the visual information is obscured or the auditory signal is degraded by background noise. It is possible that given the significance of linguistic information, activation (BOLD response) may be naturally greater to linguistic than to non-linguistic stimuli. Therefore, it is possible that in order to display multisensory enhancement with non-linguistic stimuli, a low SNR would be needed to cause greater BOLD responses as explained by the inverse effectiveness principal.

A few studies have examined linguistic and non-linguistic processing together. Through this direct comparison with similar methodology, a better understanding of the brain networks or regions involved in both linguistic and non-linguistic multisensory integration can be obtained. Stevenson and James (2009) presented participants with tool (non-linguistic) and speech (linguistic) stimuli in auditory only, visual only, and audiovisual conditions with varying SNRs. For both linguistic and non-linguistic stimuli, as the SNR decreased, the effect size, statistical significance, and multisensory enhancement of the BOLD response in the pSTS increased. Further, at high SNR, multisensory enhancement was not observed for either the linguistic or non-linguistic stimuli. Using the criterion of the magnitude of the audiovisual response being greater than the sum of the magnitude of responses to audio only and visual only stimuli, the speech audiovisual region of interest (ROI) was slightly anterior to the tool ROI in the pSTS. Watson, Latinus, Charest, Crabbe, & Belin (2014) compared audio only, visual only, and audiovisual stimuli using speech and non-speech stimuli (e.g. yawning, humming, words, etc.) and objects (e.g. a bouncing ball or toy car). While Watson and colleagues were more interested in examining face selective regions, the Superior Temporal Gyrus (STG)/STS and bilateral thalami were activated more to audiovisual stimuli, relative to the auditory only and visual only stimuli, regardless of condition. Watson and colleagues also noted that the right STG and pSTS responses were greater for the face audiovisual stimuli than the object audiovisual stimuli.

1.2. ASD and Multisensory Perception

Diagnostically, ASD are characterized by two primary characteristics: social communication impairments and restricted repetitive patterns of behaviour, interests or activities (DSM-5; American Psychiatric Association, 2013). Although sensory impairments are not included as diagnostic requirements, many individuals with ASD have been found to have atypical sensory processing (Iarocci & McDonald, 2006; Tomchek & Dunn, 2007). Dawson and Watling (2000) found that estimates of abnormal sensory behaviours in children with ASD range form 30-100%. These abnormal sensory behaviours are often one of the first prominent early diagnostic symptoms exhibited by children with ASD (Dahlgren & Gillberg, 1989). In fact, recent theories propose that sensory atypicalities are core deficits of individuals with ASD (Grandin, 1995; Klinger & Dawson, 1996; Ornitz, 1989). Iarocci and McDonald (2006) proposed that these "perceptual atypicalities may arise from the integration of specific processes rather than solely from impairments in the different components" (p.85). Thus, the atypical sensory processing of children with ASD may be due to difficulties with multisensory perception rather than problems with unisensory perception. Social interactions and communication are highly dependent on multisensory perception, and thus it is not surprising that difficulties with multisensory processing have been found to correlate with ASD symptomatology (Donohue, Darling, & Mitroff, 2012). It is likely that atypicalities in multisensory perception are at the core of some of the symptoms that are characteristic of ASD.

1.2.1. Low-Level Multisensory Perception

The nature of multisensory perception deficits in individuals with ASD remains equivocal. At a low-level, meaning more simplistic stimuli, children with ASD have been found to be equally as susceptible to audiovisual illusions in a flash-beep paradigm (van der Smagt, van Engeland & Kemmer, 2007), indicating an ability to integrate visual and auditory stimuli that happen in close temporal proximity. However, for individuals with ASD, this susceptibility to low-level illusions has been found to occur over a wider temporal window (Foss-Feig et al., 2010). This finding of a wider temporal window of audiovisual integration has been supported by studies using temporal order judgment tasks of the flash-beep paradigm (de Boer-Schellekens, Eussen, & Vroomen, 2013; Kwakye et al., 2011). Further, de Boer-Schellekens and colleagues found that this wider temporal window for simple low-level information was present for social (hand-clap), linguistic (syllables), and non-linguistic (flash-beep) stimuli. Taken together, these findings suggest that at an early low-level of processing, individuals with ASD are able to integrate auditory and visual information, but that this integration occurs over a wider window, suggesting that sensory experiences that occur temporally further apart are being perceived as a unitary event. However, these low-level studies lack ecological validity, as light flashes and auditory beeps are not commonly experienced nor do they offer important information for understanding and responding to the environment.

1.2.2. Higher-Level Multisensory Perception

Recent evidence suggests that low-level sensory processing may impact higher-order multisensory perception, such as language and communication (Stevenson et al., 2014). At a higher-order and more natural level of processing, multisensory perception differences may be limited to, or most evident for, language related information (Bryson, 1972). Compared to typically developing children, children with ASD do not benefit from the addition of visual information when identifying aurally presented words (Smith & Bennetto, 2007). Yet, children with ASD were found to be able to integrate the visual and auditory signals of a bouncing ball at equivalent rates to their typically developing counterparts (Mongillo et al., 2008), which suggests intact multisensory perception of nonsocial-nonlinguistic based stimuli. However, others have found contradictory, and even opposite patterns of results (Williams, Massaro, Peel, Bosseler, & Suddendorf, 2004)

In support of the linguistic-specific processing deficits hypothesis, individuals with ASD have been found to not be as susceptible to the McGurk effect (de Gelder, Vroomen, & van der Heide, 1991). Further, studies using preferential looking paradigms have shown that children with ASD display linguistic-specific deficits relative to cognitive and age matched controls without ASD. Bebko and colleagues (2006) found that while children without ASD displayed preferential looking to the temporally synchronous audiovisual screen in linguistic and non-linguistic conditions, children with ASD only displayed preferential looking to the temporally synchronous audiovisual screen in the non-linguistic condition. Thus, the children with ASD did not show a deficit in discriminating temporal synchrony for non-language related stimuli. Lavoie, Hancock, and Bebko (in preparation) found that children without ASD (both with and without intellectual disabilities) demonstrated greater preferential looking to the temporally synchronous audiovisual linguistic stimuli than to the non-linguistic stimuli, whereas the children with ASD (regardless of intellectual level) did not display this increase in preferential looking to the temporally synchronous audiovisual linguistic stimuli. These results suggest that individuals without ASD were better able to match auditory and visual information when the content of the information was linguistic in nature, whereas individuals with ASD did not display this linguistic gain.

Although the findings from Lavoie and colleagues (in preparation) and Bebko and colleagues (2006) differ, in both studies, the children with ASD displayed differential processing between linguistic and non-linguistic stimuli that differed from children without ASD. Bebko and colleagues (2006) found that children without ASD show similar performance across content types regardless of the nature of the stimuli, but that the rate of preferential looking for children with ASD drops in the linguistic task. In the Lavoie and colleagues (in preparation) study, the children with ASD did not drop in the linguistic task; rather they did not show an increase in preferential looking as the other groups did. The differing findings could be caused by differences in the stimuli used for each study. The non-linguistic stimuli in the Bebko et al. study may have been more compelling (a marble going through a game of mouse trap), causing an increase in looking behaviour for all children, and thus greater preferential looking to those stimuli. If the non-linguistic stimuli in the Lavoie and colleagues study were less compelling, the participants may have shown lower levels of looking, but the children without ASD returned to their higher levels of preferential looking in the linguistic condition. Therefore, the linguistic responses in both studies were similar in that they both demonstrated unique response differences for children with ASD, whereby their preferential looking in the linguistic condition is below that of children without ASD. This lack of increase in preferential looking may be related to the language-specific delays found within the ASD population. However, what these studies have failed to address is whether the multi-sensory deficits are specific to linguistic information, or also apply to social information, as the linguistic stimuli (e.g. a woman telling a story as used the Bebko and colleagues and Lavoie and colleagues studies) were inherently social.

1.2.3. ASD and Neurological Underpinnings of Multisensory Integration

One study using event-related potentials (ERP) revealed altered neural responses to audiovisual speech stimuli in individuals with ASD relative to individuals without ASD (Megnin, Flitton, Jones, de Haan, Baldewag, & Charman, 2011). However, ERPs offer relatively poor spatial resolution for determining the specific regions responsible for multisensory integration. Thus, the use of neuroimaging techniques to explore the critical brain regions and their functional network connectivity underlying multisensory integration is critically needed.

As indicated previously, the pSTS is thought to be a key region for multisensory integration. While there is limited research directly examining the neural correlates of multisensory integration in individuals with ASD, there is considerable evidence to suggest that the structure and functioning of the pSTS in it's role within the social brain is atypical for individuals with ASD (Castelli Frith, Happé, & Frith,, 2002; Hadjikhani, Joseph, Snyder, & Tager-Flusberg, 2007; Pinkham, Hopfinger, Pelphrey, Piven, & Penn, 2008). The social brain is composed of connections between cortical and sub-cortical networks (Gotts et al., 2012), which allow one to make sense of social interactions, understand other's intentions, predict what is going to happen next, and plan a response to the social exchange (Frith, 2007; Frith & Frith, 2007). The pSTS in particular, is thought to have a role in understanding and interpreting biological movement (Morris, Pelphry, & McCarthy, 2008; Pelphry, Morris, & McCarthy, 2004). In individuals with ASD, many areas of the social brain have been found to display atypical functioning, including the amygdala (Di Martino, Ross, Uddin, Sklar, Castellanos, & Milham, 2009; Hadjikhani et al., 2007; Kleinhans, et al., 2010; Pinkham et al., 2008), somatosensory and premotor cortex (Hadjikhani, et al., 2007), anterior cingulate cortex (Thakkar, et al., 2008), posterior cingulate cortex (Di Martino, et al., 2009) and superior colliculi (Kleinhans et al., 2010). Most importantly, one of the areas involved in the social brain that is found to respond atypically in individuals with ASD is the pSTS (Castelli, et al., 2002; Hadjikhani, et al., 2007; Pinkham, et al., 2008). Given that the pSTS is thought to play such a critical role in social processing and multisensory integration, it is particularly relevant for the current study to delineate if multisensory perception deficits found in individuals with ASD are limited to language information or for social information more broadly, which would include language processing.

1.3.1. Summary of Previous Research

As demonstrated through this review, multisensory perception is integral to cognitive, language, and social development and functioning (Bahrick, 2010; Bahrick & Lickliter, 2002; Bahrick & Todd, 2012; Edelman, 1992; Gibson, 1969; Gogate & Bahrick, 1998; Lewkowicz, 2000; Thelen & Smith, 1994) and individuals with ASD have been found to demonstrate atypical multisensory perception. Given the importance of multisensory perception in the development and functioning of higher-order cognitive processes, underlying atypical multisensory perception may be a core symptom of ASD that has downstream consequences for the development of language and social-cognition. However, previous research has been equivocal with the respect to the nature of the atypical multisensory perception identified in individuals with ASD. There is evidence to suggest that the deficit in multisensory perception of individuals with ASD may be limited to language related audio-visual integration (Bryson, 1972; Bebko et al., 2006; Lavoie et al., in preparation). However, previous research has failed to control the social components of the linguistic stimuli. Therefore, these proposed linguistic-specific multisensory perception deficits might be subsumed under a social multisensory perception deficit more broadly.

In typically developing populations (individuals without ASD), the pSTS has been the most consistently implicated cortical structure involved in multisensory integration (Beauchamp et al., 2004; Beauchamp et al., 2010; Calvert et al., 2000; Macaluso et al., 2004; Nath &

Beauchamp, 2012; Raij et al., 2000; Stevenson & James, 2009). Importantly, there is evidence to suggest that the structure and functioning of the pSTS is atypical for individuals with ASD (Castelli, et al., 2002; Hadjikhani, et al., 2007; Pinkham, et al., 2008). Therefore, examining the functioning of the pSTS in response to linguistic, social, and nonsocial-nonlinguistic stimuli will be essential in further understanding the nature of atypical multisensory integration in ASD.

1.4. Current Study

The overall goal of the current study was to explore the neurological basis of deficits in multisensory integration in individuals with ASD. Previous studies have failed to determine whether the ostensibly linguistic-specific deficit observed in ASD is actually due to the linguistic nature of the stimuli, or to the social nature of the stimuli that is inherent in audiovisual presentations of a person talking. Thus, the current study delineated linguistic and social processing using two distinct categories of stimuli that differed on the basis of linguistic content: *social-linguistic* and *social-nonlinguistic* stimuli. A third stimulus category was also included, *nonsocial-nonlinguistic* (i.e. stimuli that contain neither social nor linguistic information).

Because temporal synchrony is thought to be the basis for more complex forms of integration, it is important to start the exploration of multisensory integration abnormalities with temporal variations of the stimuli. Therefore, in the current study, the stimuli varied on three dimensions: (1) content (social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic); (2) synchrony (temporally synchronous and temporally asynchronous); and (3) modality (audio only, visual only, and audiovisual).

The overall goal was accomplished by examining multisensory integration in three studies. As the pSTS has been largely implicated as a primary area for multisensory integration, the current studies examined the activation and functional connectivity of the pSTS in order to fully understand altered neurological processing of multisensory stimuli in ASD. The goal of the first study was to evaluate and compare various methods of identifying areas that demonstrate clear multisensory enhancement in a neurotypical population. In study two, the regions that demonstrate multisensory enhancement, including the pSTS in particular, served as ROIs to compare multisensory integration in all four content conditions, and both synchrony conditions, between individuals with ASD and those without ASD. However, demonstrating that the pSTS is involved in multisensory integration does not necessarily justify the conclusion that it is the region where multisensory integration occurs. Rather, the pSTS may be part of a broader network of regions engaged by multisensory integration. Thus, the goal of the third study was to explore whole-brain patterns of covariance involved in multisensory integration, rather than the activity of individual brain regions in isolation. The third study used a multivariate fMRI analysis approach (partial least squares: PLS) to examine whole-brain patterns of covariance related to multisensory integration.

Study One:

Comparisons between Methodological Approaches for Determining Multisensory Integration Using Functional Magnetic Resonance Imaging

Multisensory perception is integral to forming a coherent perception and understanding of the environment. In addition to offering a more concise and integrated view of the environment, multisensory perception leads to enhanced detection, faster processing speed and response times (Calvert, et al., 2000; Frens & Van Opstal, 1995; Hershenson, 1962; Hughes et al, 1994; Morrell, 1968; Perrott, Saberi, Brown,& Strybel, 1990; Stein et al., 1989;), guides selective attention (Bahrick & Lickliter, 2002; Bahrick & Todd, 2012), enhances degraded signals (Reisber et al., 1987; Rosenblum et al., 1996; Sumby & Pollack, 1954), and facilitates the understanding of social interactions (Bahrick & Todd, 2012). Further, multisensory perception of sensory experiences typically provides information that is not available from unisensory experiences (O'Hare, 1991) (e.g., the influence of smell over the taste of food).

Researchers have gained an understanding of the importance of multisensory perception for development through behavioural studies with well-developed methodology. However, the methodology utilized to explore the neurological mechanisms of multisensory integration is still an area of controversy and in need of further development.

2.1. Single Cell Recordings in Animals

Much of our understanding of the location and nature of the neuronal basis of multisensory integration comes from animal research using single cell recordings. In these designs, researchers examine the activation (or firing) of neurons in response to environmental stimuli. Through these single cell recording designs, unisensory (or unimodal) and bimodal neurons have been identified. Unisensory neurons are activated/respond to one modality of sensory stimuli and this response is not altered by the addition of another sensory modality. Bimodal neurons respond to two sensory modalities, such as a neuron that fires in response to both visual and auditory stimuli (James, Stevenson, & Kim, 2012; Meredith & Stein, 1983; Stein & Stanford, 2008). These bimodal neurons can be further subdivided into those that demonstrate multisensory enhancement (i.e., the multisensory activation is greater than the activation of the audio and the visual stimuli) and those that do not. These bimodal neurons that demonstrate multisensory enhancement can then be further subdivided again into those that display super-additivity and those that do not. Super-additivity occurs when the response of the neuron to multisensory stimuli is greater than what would be expected from the summation of the unisensory responses (Stein & Meredith, 1993) (see figure 1). This super-additive response is dependent on the stimuli satisfying the temporal and spatial properties of integration (i.e. that the sensory cues are in close temporal and/or spatial proximity) (Meredith & Stein, 1986, 1996; Miller & D'Esposito, 2005).

In animal research, many of the properties of multisensory neurons (e.g. super-additivity) have been demonstrated at the subcortical level in the SC of macaques and cats (Calvert et al., 2000; Stein & Meredith, 1993), at the cortical level in the AES and rLS sulci in cats (Royal et al., 2009), and in the pSTS in rhesus monkeys (Ghazanfar et al., 2008). These single-cell recording studies have largely focused on the development of multisensory integration and its importance in other cognitive processes through the examination of when super-additivity occurs and when super-additivity fails to occur.

In humans, the pSTS has been one of the most consistently implicated regions for audiovisual integration (Beauchamp, Lee, Argall, & Martin, 2004; Beauchamp et al., 2010; Calvert et al., 2001; Nath & Beauchamp, 2012; Stevenson & James, 2009; Watson et al., 2014). While there is evidence, as described above, for multisensory integration in the cortex, less is known about subcortical areas that demonstrate these multisensory properties. In humans, the SC, the suprageniculate and medial pulvinar nuclei of the thalamus, the amygdaloid complex including rhinal cortex, and the hippocampus have been implicated, but have not been as extensively explored as other cortical areas, such as the pSTS in particular (Fries, 1984; Mesulam & Mufson, 1982; Mesulam, Mufson, Levey, & Wainer, 1984; Pearson, Brodal, Gatter, & Powell, 1982).

2.2 Human Methodology

Past studies have attempted to extend findings from animal research to the identification of brain structures in humans that are responsible for multisensory integration (Laurienti et al., 2005). Specifically, researchers have attempted to use the property of super-additivity in order to identify regions of the human brain that display a greater response to synchronous multisensory stimuli than to the sum of the unisensory stimuli. However, there are controversies over the applicability of animal methodology (i.e. single-cell recording) to human research, which relies exclusively on non-invasive techniques. Specifically, extensions from single cell recordings to methods typically used with humans, including fMRI, positron emission tomography (PET), magnetoencephalography (MEG), or ERP, are problematic, as these techniques involve averaging over a large population of cells (Laurienti et al., 2005).

Functional MRI (fMRI) has been used to determine the location of multisensory integration by examining BOLD contrasts across tasks in humans. BOLD fMRI reflects activation across a large population of thousands of neurons. Techniques that measure activation across a large population of neurons are examining a heterogeneous population of neurons that may respond differentially to stimuli. For example, in the SC and in the pSTS, only 25 – 60% of the total population of neurons is estimated to be multisensory (Beauchamp et al., 2004; Laurienti et al., 2005). Similarly, only 25% of the neurons in the AES of cats have been found to be multisensory neurons (Wallace, Meredith, & Stein, 1992). Beauchamp and colleagues (2004) identified a patchy distribution of clusters of neurons that responded to auditory, visual, and multisensory (audiovisual) stimuli in the pSTS. As Goebel and van Atteveldt (2009) postulate, this may indicate a structure that is organized in cortical columns (i.e. approximately a hundred thousand neurons with similar response specificity) similar to Seltzer and colleagues' (1996) findings in rhesus monkeys. However, fMRI is an indirect measure of the mean activity of many thousands of neurons contained within voxels – 3-dimensional cubes of cortical tissue on the order of several millimetres cubed. Thus, the spatial resolution of fMRI is too low to parse this heterogeneous organization of unisensory and multisensory neuronal populations (Goebel & van Atteveldt, 2009). Therefore, the nature of multisensory response is ambiguous using fMRI as it pools the response of a large number of cells.

Given these concerns regarding the applicability of single cell recording methods, such as super-additivity, to large neuronal populations, a number of different methods and analytic strategies have been developed for the identification of brain areas responsible for multisensory integration in humans. The least commonly used methods are intersection and conjunction techniques, which involve looking for overlap in response to multiple unisensory stimuli in the same brain regions within or across studies (see Calvert, 2001 and Laurienti, et. al., 2005 for a review of these methods). One large problem with these methods is that activation in response to different unisensory conditions may simply indicate a response from two distinct sets of unisensory neurons in the same voxel. A type I error is made if an area is identified as multisensory when in fact it consists of co-localized populations of unisensory neurons (Calvert, 2001). Other researchers have suggested identifying regions that respond to only multisensory stimuli and not to unisensory stimuli (see Calvert, 2001 and Laurienti et al., 2005); however, finding a region that fits this criterion is unlikely given that researchers have found that most brain regions that contain multisensory neurons also contain a large population of unisensory neurons. Further, multisensory neurons may also respond weakly to unisensory stimuli. The most commonly used techniques for the identification of multisensory regions, other than the super-additive method, are the mean and max methods.

2.2.1 Mean Method

The mean method, which is the least conservative method, classifies multisensory regions based on the BOLD response to the multisensory stimuli being greater than the mean of the two unisensory responses. The formula for identifying audiovisual multisensory integration (where M represents multisensory audiovisual, A represent unisensory audio, and V represents unisensory visual) would be $M > \text{mean}(A, V)$ (Beauchamp, 2005; Love, Pollick, & Latinus, 2011). Using this technique to compare multisensory and unisensory with nonsocial-nonlinguistic stimuli (e.g. tools), Beauchamp (2004) identified regions bilaterally in the pSTS that were specific to multisensory integration.

However, this method is both theoretically and empirically problematic. This criterion is very liberal and thus may actually reflect an increased BOLD response due to activation of both unisensory sets of neurons, rather than true multisensory neurons. Therefore, the mean technique is likely to cause an inflated type I error rate, particularly in the situation of one unisensory stimulus causing suppression and the other activation (Laurienti et al., 2005). For example, if a voxel displayed activation in response to an auditory unisensory stimulus (for the purposes of this example the beta value for the activation displayed by the voxel in response to the audio unisensory stimulus will be represented by "x") and suppression to a visual unisensory stimulus (for the purposes of this example the beta value for the activation displayed by the voxel in response to the visual unisensory stimulus will be represented by "y"), the resulting calculated average of the two conditions (mean (x, y)) would be lower than the auditory unisensory stimuli activation alone (e.g. $[x + y]/2$ = a value that is $< x$). Further, the activation seen in the voxel in response to a multisensory stimulus (for the purposes of this example the beta value for the activation displayed by the voxel in response to the multisensory stimulus will be represented by "z") could be simply the activation of the auditory neurons in that voxel (such that $z = x$). Therefore, in this example the result would be that the voxel is incorrectly identified as multisensory when really the response of the multisensory stimuli was driven entirely by the auditory component of the stimuli.

2.2.2. Max Method

Others have suggested looking for areas where the multisensory stimuli elicit a greater response than the largest unisensory response (Calvert et al., 1999; Hadjikhani & Roland, 1998; Love, et al, 2011). Therefore the formula for determining audiovisual multisensory regions would be $(M > V) \cap (M > A)$. Using the max criterion to examine audiovisual integration of social-linguistic information, the bilateral pSTS has consistently been implicated as an area involved in/critical for multisensory integration (Kreifelts, Ethofer, Grodd, & Wildgruber, 2007, Szycik, Tausche, & Munte, 2008; Wright, Pelphrey, Allison, McKeown, & McCarthy, 2003).

The max method is more conservative than the mean criterion and less conservative than the super-additive criterion. The max criterion has been proposed as the most suitable for a wide range of voxels with a variety of unisensory response profiles (e.g. weak unisensory BOLD response, saturated BOLD, suppression response to one unisensory stimulus and activation in the other). However, while the max criterion has been proposed to be a more empirically tenable method, the theoretical possibility of false positives is still a concern, particularly when the max criterion is met in brain areas not thought to be multisensory (James et al., 2012). Therefore this criterion may still be susceptible to type I errors, as the multisensory response may be a linear sum indicative of the existence of both unisensory auditory and visual neurons, rather than the presence of multisensory neurons (Calvert et al., 2001; James et al., 2012; Love et al., 2011).

2.2.3. Super-Additive Method

As indicated above, the super-additive method is directly comparable to the single cell recording method, whereby in order to determine multisensory integration, the BOLD response to the multisensory stimuli must be greater than the sum of both unisensory responses. In the example of audiovisual integration the equation would be $M > A + V$. Using the super-additive method, Joassin and colleagues (2011) and Calvert and colleagues (2000, 2001) demonstrated significant activation to multisensory audiovisual social-linguistic information in the STS bilaterally. However, others have not found these same regions using the same technique

(Beauchamp, 2005; Beauchamp et al., 2004; Laurienti et al., 2005; Stevenson et al., 2007). Therefore, while the super-additive method may be theoretically valid, it may not be empirically tenable.

A super-additive BOLD response is less prone to false inferences than the other methods (Calvert, 2001), but is less likely to be observed because of the heterogeneity of cells at the voxel level (Goebel & van Atteveldt, 2009), increasing the type II error as voxel size increases. Thus the super-additive criterion is at risk of being too conservative. In a particular brain region, it is possible that only 25% of the neurons are multisensory, that not all multisensory neurons will display a super-additive response, and that the unisensory neurons may also respond to the multisensory stimuli (Laurienti et al., 2005; Wallace et al., 1992). Thus, if there is a greater proportion of unisensory neurons, it is possible that the response of the unisensory neurons may be greater than the multisensory response (James et al., 2012).

An additional concern with the super-additivity method is that it is particularly impacted by different neuronal response profiles. Super-additivity is prone to false negatives (i.e. type II error) when there is BOLD saturation (high activation to both unisensory conditions) (Calvert, 2001). It has been recommended to use "weak" stimuli to prevent BOLD saturation (Calvert, 2001; Stevenson et al., 2009) and to enhance inverse effectiveness (Stein & Meredith, 1993; Goebel & van Atteveldt, 2009). Further, when one unisensory stimulus produces a positive BOLD response and the other unisensory stimulus produces a depressed BOLD response, the super-additive method may also increase the type I error (i.e. inferring that activation occurs from multisensory neurons, when in fact it is from unisensory neurons). In this case of differing responses to each unisensory stimulus, a super-additive effect would incorrectly be found due to the summation of positive and depression responses (Calvert, 2001). Therefore, Calvert (2001) suggested that while it is not necessary for the unisensory conditions to reach significance, it is necessary to determine that neither of the unisensory responses are significantly below baseline. Calvert further suggested that a more valid method for testing the super-additive method is through the use of a rest condition, thus the formula would become (M – rest) > [(A – rest) + (V – rest)].

2.2.4. Comparison of Techniques

There have been a few fMRI studies that have directly compared some of the methods described above. Love and colleagues (2001) compared all three methods using social-linguistic stimuli. They found that the super-additive method only identified regions where there was deactivation to one of the unimodal stimuli and a positive response to the other unimodal stimuli. This finding indicated that there were no regions that responded positively to both auditory and visual unimodal stimuli and where the combined audiovisual presentation displayed greater activation than the sum of the unimodal conditions. In contrast, they found that the mean method identified only regions that responded to unimodal stimuli (i.e. the “multisensory areas” defined by the mean method overlapped completely with the areas identified by unisensory activation), and thus the mean method appeared to be too liberal. Finally, the max method (i.e. multisensory audio-visual activation greater than both the auditory and visual unisensory responses individually) identified only the left hippocampus when using a more liberal threshold. Love and colleagues (2001) recommended a combination of the above methods, with the addition of manipulation of congruency or stimuli effectiveness (e.g. inverse effectiveness).

Beauchamp (2005) compared the mean, max, and super-additive methods with nonsocial-nonlinguistic stimuli. In contrast to Calvert's (2001) findings, Beauchamp's study did not detect any brain regions that could be identified as multisensory using the super-additive method, but detected the pSTS using the mean and max methods. Therefore, based on Love and colleagues (2001) and Beauchamp's (2005) findings, the super-additive method appeared to be too conservative and the mean method appeared to be too liberal. While the max method identified regions in both studies, there are theoretical limitations to the understanding of what the max method is actually identifying.

James and colleagues (2012) provided a theoretical comparison between the max and super-additive methods. They characterized the debate between the super-additive and max criteria as the difference between theoretical and empirical approaches. James and colleagues stated that the super-additive method is a theoretical approach that best captures the hypothetical underlying neuronal activity in response to multisensory stimuli (Beauchamp, 2005; Calvert et al., 2000; Laurienti et al., 2005; Stevenson et al., 2009). In comparison, the max criterion is an empirical approach and has been found to most consistently result in the identification of brain regions thought to be involved in multisensory integration (Beauchamp et al., 2004; Doehrmann et al., 2008; Hein et al., 2007, van Atteveldt et al., 2007). However, this logic underlying the max method amounts to circular reasoning, such that the justification of the max method as being valid for identifying multisensory regions is because it has identified regions presumed to be multisensory.

2.3. Temporal Synchrony Method

As the current methods for identifying multisensory regions have been shown to have empirical and/or theoretical limitations, it is imperative to the study of multisensory perception that an alternate method that is both empirically tenable for fMRI research and theoretically valid be developed. The current study proposes a new method based on the properties of perceptual integration identified at the single-cell level, in particular, the temporal property of integration. This fundamental temporal property is that sensory stimuli will be integrated more successfully when they are temporally synchronous (Miller & D'Esposito, 2005; Stein & Meredith, 1993). Temporal synchrony of sensory stimuli is required in order for multisensory super-additivity to be seen at the single-cell level (Stein & Meredith, 1993). The new temporal synchrony criteria proposed for fMRI research will identify brain regions that demonstrate multisensory enhancement to temporally congruent audiovisual stimuli and suppression, or significantly less activation, to temporally incongruent audiovisual stimuli. The formula for this method would be synchronous M > asynchronous M.

Few studies have explored the use of temporally synchronous and asynchronous stimuli with human brain imaging. Marchant, Ruff, and Driver (2012) and Noesselt et al. (2012) used arbitrary combinations of audiovisual stimuli (e.g. flashing checkerboard paired with a pure tone) and found greater activation in the bilateral pSTS when the flash-tone pairings were synchronous versus asynchronous. With stimuli that were more naturalistic and applicable to human sensory processing, Macaluso and colleagues (2004) used PET with eight participants to examine activation in response to temporally and/or spatially congruent (i.e. the visual stimulus and the auditory stimulus originate from the same spatial location) and incongruent stimuli of a person saying highly familiar words. They reported that the left pSTS and right inferior parietal lobule displayed greater activation for temporally synchronous relative to temporally asynchronous audiovisual stimuli, regardless of the spatial congruity. The lateral and superior occipital gyri were also found to display greater activation for temporally synchronous audiovisual information when it was also spatially congruent. However, these studies are limited with respect to the type of stimuli examined, as they did not compare language and non-language based stimuli; and most importantly, have failed to compare this method to previously used methods, such as the mean, max, and super-additive methods.

I propose that the temporal synchrony method is a more theoretically valid and empirically tenable method for fMRI research than the previously used methods reviewed above. Specifically, unlike the mean and max criteria, the temporal synchrony method is based on properties known to be demonstrated by multisensory neurons; therefore, the temporal synchrony method is more theoretically valid than both the mean and max methods. Further, the temporal synchrony method is more likely to be empirically practical for fMRI research than the super-additive method. The super-additive method is not likely to be observed because of the heterogeneity of cells at the voxel level (Goebel & van Atteveldt, 2009). In contrast to the super-additive method, where the BOLD response of a multisensory condition is compared to the BOLD response of two unisensory conditions, the temporal synchrony method employs a method in which the BOLD response to two conditions that both contain auditory and visual stimuli is compared (i.e. both conditions are multisensory). Therefore, the information provided in the temporal synchrony method is equivalent in both conditions, with the exception of the temporal synchrony of the stimuli. Thus the heterogenous population of neurons will respond equivalently to both the synchronous and asynchronous conditions, with the exception of the multisensory neurons that will have a greater BOLD response to the temporally synchronous condition. As a result, the temporal synchrony method will be able to effectively identify regions that have a large population of multisensory neurons.

Furthermore, because the mean, max, and super-additive methods are based on comparing the multisensory response to some combination of the two unisensory responses, the calculations would be impacted by any unisensory stimuli that elicit a depressed BOLD response (e.g. the summation of a positive BOLD response to audio and a negative BOLD response to visual would result in an incorrectly identified super-additive region). As a result, the mean, max, and super-additive criteria require that only areas that demonstrate unisensory activation (with the removal of suppression relative to baseline) are used in the calculations for identifying areas that demonstrate multisensory integration. However, finding areas that demonstrate greater BOLD response to temporally synchronous audiovisual stimuli than to temporally asynchronous audiovisual stimuli does not necessitate the use of only positive BOLD activation.

2.4. Current Study

As reviewed above, previous attempts to identify multisensory regions using fMRI have relied on either the use of a theoretically based method adapted from single cell recording studies (super-additive method) that is not likely to be seen in fMRI research because of heterogeneity at the voxel level, or on non-theoretically based methods (mean and max methods) because they are believed to work empirically. This new temporal synchrony method (i.e. adapted for fMRI) had never been compared to the commonly used methods (mean, max, and super-additive) for identifying multisensory regions. Therefore, the current study was designed to examine the efficacy of the temporal synchrony technique for fMRI research relative to previously used methods. It was hypothesized that the temporal synchrony method, which was grounded in theory, would be a more empirically tenable method for fMRI than the other theoretically based method (super-additive) and equivalent in empirical utility to the non-theoretically based methods (mean and max).

The current study also extended the literature by applying the temporal synchrony method to linguistic and nonlinguistic stimuli. Specifically, the mean, max, and super-additive techniques were compared to the temporal synchrony method across social-linguistic, nonsocial-nonlinguistic, and social-nonlinguistic stimuli. Further, in order to examine the empirical utility of the methods, it was imperative to ensure that the location of regions of activation found to display multisensory integration in each of the three methods also be empirically tenable. Because the pSTS is the most consistently implicated region for multisensory integration in humans, the current study focused on the applicability of the above methods to identify clusters of activation in the pSTS bilaterally.

2.5. Methods

2. 5.1. Participants

Seventeen young adults (4 females), between the ages of 18-29 (*M* = 22.89, *SD* = 3.56), participated in the current study. All participants had average to above average intellectual abilities, as measured by the Wechsler Abbreviated Scales of Intelligence two-item screener (WASI; Wechsler, 1999). The participants' full scale IQ standard scores ranged from 96 (39th percentile) to 137 (99th percentile) (*M* = 108.38, *SD* = 11.74). Participants were required to speak English as their first language, have normal hearing, normal or corrected to normal vision, and be neurologically healthy (i.e. free of brain injury, tumors, seizures, and ASD). Participants were recruited through posters, advertisements, and the York University Undergraduate Research Participation Pool (URPP). Participants recruited through the URPP received 2 credits towards course completion. All other participants were given a $10 gift card for their time and $10 for travel expenses.

2.5.2. Materials

2.5.2.1 Measures.

2.5.2.1.1. Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). The WASI is a brief two to four subtest individually administered test of intelligence for children and adults. The two-subtest form was used in the current study, which consists of the Vocabulary and Matrix Reasoning subtests. Administration of the two-subtest form takes approximately 15 minutes and provides estimates of full scale intelligence. The WASI has demonstrated adequate reliability, including split-half reliability and test-retest reliability between 2 to 12 weeks. It also correlates well with other measures of intelligence, including the Wechsler Adult Intelligence Scale (Wechsler, 1999).

2.5.2.2. Stimuli. Videos were composed of three different content conditions that were originally developed for a previous study (Lavoie et al., in preparation): social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic. The *linguistic* stimulus displayed a woman telling a story, with only the head and neck of the woman visible, and her hair pulled back, to minimize distractions. The *social-nonlinguistic* stimulus involved the same woman making different sounds that did not have language content, such as tongue clicks and humming. Two *nonsocial-nonlinguistic* stimuli were used in the current study. The first nonsocial-nonlinguistic stimulus was composed of a bird's-eye view of a hand playing a children's song on a piano (referred to for the remainder of the paper as music nonsocial-nonlinguistic). As music processing may be a distinct and dissociable neurological process from that of other stimuli (see Peretz & Zatorre, 2005 for a review of brain organization for music processing), an additional nonsocial-nonlinguistic stimulus of a marble going through a marble run was created. This stimulus involved a series of tubes and steps that make different noises as the marble rolls (referred to as a mousetrap nonsocial-nonlinguistic condition). For each content condition, there were four synchrony conditions: auditory only, visual only, audiovisual synchronous, and audiovisual asynchronous versions. In the asynchronous conditions, the visual content was 500 milliseconds temporally ahead of the auditory content. All videos were 12 seconds long.

2.5.2.3. Apparatus. Participants were scanned using a Siemens 3T Magnetom Tim Trio MRI scanner and a 32-channel head coil at York University. The task was presented on a screen outside of the scanner controlled by an Apple laptop computer and presented using E-prime software. Participants viewed the screen via a mirror system attached to the head coil. An Avotec RE-5721 Dual Channel eyetracker was attached to the head coil using the SensoMotoric Instruments iView X software. Participants' eye gaze was monitored by the experimenters in order to ensure the participants were attending to the videos. Due to technical issues with the system, quantitative metrics of eye gaze were not analyzed.

2.5.3. Experimental Design

Anatomical scans were collected first and acquired using a T1-weighted magnetization prepared rapid gradient echo (MPRAGE; TR = 1900, TE= 2.52, TI= 900, flip angle= 9 degrees, 1mm isotropic voxels). Next, task-evoked brain activity was measured using fMRI during multiple runs of the task-based functional localizer (TR = 2000 ms, TE = 30 ms, flip angle = 90 degrees, 32 interleaved contiguous axial slices per volume, slice thickness = 3.5 mm, FOV = 240 mm, acquisition matrix = 96 x 96, single-voxel volume = 2.5mm x 2.5mm x 3.5mm). All participants completed 6 runs of a standard block design task presented in a different randomized order. Four of the runs contained one block of each synchrony condition (audiovisual synchronous, audiovisual asynchronous, visual only, and auditory only) for three of the four content conditions (social-linguistic, social-nonlinguistic, music nonsocial-nonlinguistic) for a total of 12 stimulus blocks. Two additional runs contained two presentations of each synchrony condition for the alternate mousetrap nonsocial-nonlinguistic stimulus for a total of 8 stimulus blocks. The acquisition of the data from the social-linguistic, social-nonlinguistic, and music nonsocial-nonlinguistic conditions from the Lavoie et al. study was prioritized in order to allow comparisons with previously collected behavioural data. The additional mousetrap nonsocial-nonlinguistic data were acquired thereafter to serve as a better control condition than the music stimuli, as music may be processed in a unique and distinct fashion from other multisensory stimuli (see Peretz & Zatorre, 2005). However, this prioritization could potentially create some order effects, because the mousetrap nonsocial-nonlinguistic stimuli were not presented intermixed with the other conditions, but rather, in separate runs. Therefore, after participants completed the 6 runs described above, extra runs were displayed to increase power and control for order effects. These extra runs included a presentation of each synchrony condition by content condition combination in a randomized order, including both nonsocial-nonlinguistic conditions. Participants completed as many additional runs as possible, taking into consideration participant alertness, comfort, and fatigue. In all runs (the original 6 runs and additional combined runs), task blocks (12 s) were interleaved with fixation blocks (12 s), with an additional fixation block at the start and end of the run. For all participants, only runs with less than 2 mm of movement were included. The average number of runs per participant was 8.22 runs (*SD* = 1.22; range 6-10 runs).

2.5.4. Procedure

All participants gave written informed consent prior to participation and completed a screening form to ensure they were eligible to enter the MRI environment. Participants were in the scanner for between 40-90 minutes, depending on the number of runs they completed. After completing the fMRI task, participants were administered the WASI. All participants were then thanked for their time and provided with their compensation for participation (e.g. course credit or $10 gift card).

2.5.5. Data Processing

Echo-planar images (EPI) were preprocessed using the Analysis of Functional Neuroimages (AFNI) software package (Cox, 1996). For each subject, the anatomical scan was segmented into tissue compartments using Freesurfer (Fischl et al., 2002). As the scanner removes the initial pre-steady state volumes automatically, no volumes were removed from the start of the runs. Large transients in the time-series were removed through interpolation (3dDespike). Volumes were slice time corrected in ascending interleaved order, deobliqued, and motion corrected to the 9th volume from the first run. A whole-brain mask was generated from the first run and each run was scaled to have a mode value of 1000. Linear transformations were created between single-subject EPI and T1 space, using the mean EPI image generated earlier. T1 segmentations were registered to EPI space using nearest neighbor interpolation. Ventricle, white-matter, and draining-vessel masks were created based on the Freesurfer segmentations and eroded (by 1 voxel around all edges) to prevent partial volume effects with gray matter. The masks were then applied to the volume-registered EPI data yielding pure nuisance time-series for the ventricles and draining-vessels, as well as local estimates of the white-matter signal averaged within a 15-mm radius sphere. To summarize, nuisance variables for each voxel's time-series included: an average ventricle time-series, an average draining-vessel time-series, a local average white-matter time-series, six head motion parameter estimates and the temporal derivative of each, and nine physiological signal regressors from Retroicor and RVT. All of the above nuisance time-series were detrended with fourth-order polynomials. Least-squares model fitted time-series of these nuisance variables were then subtracted from the voxel time-series, yielding a residual time-series that was used in all subsequent statistical analyses.

For cortical surface-based analyses, subject-specific surface models were created from each participant's anatomical scan using Freesurfer. Standard-mesh surfaces of 141, 000 nodes per hemisphere were created using AFNI Surface Mapper (SUMA, Saad et al., 2004) to produce node-to-node anatomical correspondence across surfaces for all participants. The denoised residual time-series described previously for the volume based analyses were mapped onto the cortical surfaces (3dVol2Surf), with a mean kernel of 10 sampling points uniformly distributed along a line between smooth white matter and pial surfaces, extending 80% of the distance between corresponding nodes on the two surfaces. Spatial smoothing was performed on the surface-mapped functional data (SurfSmooth) with a heat kernel resulting in a 6 mm full-width-at-half-maximum noise spatial correlation structure along the white matter surface.

2.6. Data Analysis

To derive the BOLD response magnitudes for each of the conditions of interest at the individual subject level, the task runs were modeled with a boxcar function. The onset and offset points coincided with the beginning and end of each stimulus, respectively. Runs were convolved with a canonical hemodynamic response function and deconvolved using AFNI (3dDeconvolve – block). In addition to nuisance regressors (12 regressors for the motion parameters and a third-order polynomial regressor to account for very low-frequency MRI signal drift), two models were created. One model collapsed across content conditions, and thus, included 4 regressors corresponding to the 4 synchrony conditions (audio only, visual only, audiovisual synchronous, audiovisual asynchronous). The second model included 16 regressors corresponding to the 16 stimulus categories (one social-linguistic, one social-nonlinguistic, and two nonsocial-nonlinguistic each presented in audio only, visual only, synchronous, and asynchronous audiovisual conditions). Any runs that had absolute movement greater than 2mm in any direction were removed from the analyses.

The inclusion of deactivation in the calculation of the mean, max, and super-additive methods can inflate type 1 error and lead to the inaccurate identification of multisensory regions (e.g. leads to the misidentification of regions that show unisensory activation as showing multisensory activation). Therefore, as suggested by Beauchamp (2005) and Calvert (2001), task-related deactivations were removed for the audio only, visual only, and audiovisual synchronous conditions prior to creating the contrasts for the mean, max, and super-additive criteria. In addition, for the max criterion, a conjunction map was created showing regions where activation was greater in the synchronous audiovisual condition than both the audio-only condition and the visual-only condition. The threshold was set at $p < 0.03$ for each of the individual contrast maps included in the conjunction map for the max method, thus, the significance threshold of the resulting conjunction map was $p < 0.001$ (which approximately provides a q (false discovery rate) < 0.05).

The pSTS has been found to respond to both linguistic and non-linguistic content in both hemispheres. It is therefore, a critical hypothesis of the current paper that the pSTS is a region responsible for multisensory integration regardless of content type. Therefore, ROIs were identified by collapsing across content conditions. Individual-level analyses were performed to try to identify an ROI in the left and right pSTS of each individual for each method. Only one ROI was identified per hemisphere per participant. Further, in order to constrain the identification of the ROIs to the posterior portion of the STS, all ROIs were posterior to the central sulcus. If there were two possible ROIs in the pSTS for a participant, the one that was more posterior and with a higher peak voxel value was used. ROI peak activation locations were defined in surface-based space ($p \leq 0.01$). Masks were created on the cortical surface using SUMA (ROIgrow) with a surface node coordinate distance calculated along the surface's mesh (lim) that was set at an approximate radius of 6mm.

Group-level analyses were completed using the individually defined ROIs. In order to further demonstrate that the pSTS activation was driven by multisensory content in general, rather than specifically by the linguistic or non-linguistic multisensory content, the individual content conditions were also examined within each ROI. For each participant, the mean activation for each content by synchrony condition, created from the second deconvolved model, and the collapsed synchrony and asynchrony activation, from the first deconvolved model, was extracted from each ROI.

For the whole-brain analyses task-related deactivations for the audio-only, visual-only, audiovisual synchronous, and audiovisual asynchronous maps were removed at the group level, after running 3dANOVA2. The contrasts of interest in order to create the mean, max, and super-additive method were run using the same strategies as discussed for the individual level conditions, but collapsed across conditions. The temporal synchrony method was created using a contrast run during 3dANOVA2.

2.7. Results

2.7.1. Individual Level Analyses

2.7.1.1. Collapsed Across Conditions. Collapsing across content conditions, individual level analyses were performed to identify a region in the left and right pSTS that demonstrated multisensory integration as defined by each method (e.g. for the mean method, a region in the pSTS that demonstrated M > mean A and V). The number of participants for which a ROI could be identified in the left and right pSTS collapsed across content conditions is demonstrated in table 1. As shown in the table, the temporal synchrony method resulted in the identification of more ROIs than any other method. It is important to note that when it was difficult to identify an ROI for the super-additive and max method, it was because the methods resulted in little activation overall. An example of the activation for each method in one participant is demonstrated in Figure 2. Figure 3 shows an example of the resulting ROI for each of the methods, and where they overlapped, in two individual participants. As can be seen in figure 3 there is some variability between individual participants, which is consistent with previous research that has demonstrated that there is substantial individual variability when localizing functional regions that are not transformed into standard space (Glezer & Riesenbuber, 2013; Stevens, Tessler, Peng, & Martin, 2015).

2.7.1.2. Content Condition Analysis. Using the ROIs created from the collapsed conditions, the mean activation within the ROI for each content condition (social-linguistic, social-nonlinguistic, mousetrap, and music) was extracted. For the ROIs created using the mean, max, and super-additive methods, any deactivation was removed, then, mean activations for the audio only, visual only, and audiovisual synchronous conditions were extracted. For the ROIs created using the temporal synchrony method, the mean activations were extracted for the synchronous audiovisual and asynchronous audiovisual conditions. For each method, an overall repeated measures ANOVA comparing the method (e.g. for the mean method comparing the synchronous audiovisual condition to the mean of the two unisensory conditions) across hemispheres for each content condition was completed. Also, as the repeated measures ANOVA necessitated the inclusion of only participants that had ROIs in both the left and right hemispheres, which lowered the statistical power, individual 2-tailed t-tests for each content condition in each hemisphere were also completed for each method. As the examination of each method in each hemisphere was determined a priori, multiple comparisons corrections were not applied.

2.7.1.2.1.Mean Method. Using the data extracted from the ROIs defined by the mean method (table 2, figures 4 and 5), a 3-way (2×4×2) repeated measures ANOVA with modality (Synchronous audiovisual, mean of visual only and audio only), content condition (social-linguistic, social-nonlinguistic, and both nonsocial-nonlinguistic conditions), and hemisphere (left vs. right) as factors was conducted on the 8 participants that had both left and right ROIs. The overall 3-way interaction (modality × condition × hemisphere) was not significant ($F_{(3,\ 18)}$ = 0.66, p = 0.59, η^2 = .10). Further, there was no significant interaction for either content condition or modality by hemisphere ($F_{(3,\ 18)}$ = 0.96, p = .43, η^2 = 0.14 and $F_{(1,\ 6)}$ = 1.83, p = 0.22, η^2 = 0.23, respectively). There was also no significant interaction between modality by content condition ($F_{(3,\ 18)}$ = 0.34, p = 0.79, η^2 = 0.05). Finally, there was no main effect of content condition or hemisphere ($F_{(3,\ 18)}$ = 0.86, p = 0.48, η^2 = 0.13 and $F_{(1,\ 6)}$ = 3.43, p = 0.11, η^2 = 0.36, respectively). There was a significant main effect of modality ($F_{(1,\ 6)}$ = 8.80, p = 0.02, η^2 = 0.59) where the synchronous audiovisual activation (M = 0.40, SE = 0.09) was greater than the mean activation of the audio only and visual only conditions (M = 0.18, SE = 0.03).

Further, in order to examine each content condition and hemisphere individually (as the power for overall differences in the ANOVA was reduced), eight t-tests were performed to examine if the synchronous audiovisual condition was greater than the mean of the audio only and visual only conditions for each content condition in each hemisphere. As can be seen in table 2, the synchronous audiovisual activation was greater than the mean of the audio only and visual only conditions in the right hemisphere in the mousetrap condition, the right and left hemispheres for the music condition, and in the left hemisphere for the social-nonlinguistic condition. The comparison of the synchronous activation to the mean of the audio only and visual only conditions approached significance in the left hemisphere for the social-linguistic condition. All other comparisons were not significant.

2.7.1.2.2. Max Method. Using the data extracted from the ROIs defined by the max method (table 3 figure 6), a 2 x 4 repeated measures ANOVA with modality (synchronous audiovisual, max activation of audio only and visual only combined) and content condition (social-linguistic, social-nonlinguistic, and both nonsocial-nonlinguistic conditions) was performed with the three participants in which an ROI could be identified in the left hemisphere. No ROIs were identified for any of the participants in the right hemisphere in the collapsed condition; therefore there were no data to analyze for the right hemisphere and the ANOVA was only completed with the left hemisphere. The interaction between the modality and the content conditions was not significant ($F_{(3, 6)} = 1.46$, $p = .32$, $\eta^2 = 0.42$). The main effects of content condition ($F_{(3, 6)} = 1.89$, $p = .23$, $\eta^2 = 0.48$) and modality ($F_{(1, 2)} = 0.95$, $p = .43$, $\eta^2 = 0.32$) were also not significant.

Four two-tailed t-tests were performed to determine if activation in the synchronous audiovisual condition was higher than either the audio only or visual only conditions for each content condition in the left hemisphere. As shown in table 3, the synchronous audiovisual activation was not greater than the max of either the audio only or visual only conditions in any of the content conditions. However, given that only three participants were found to display a Max ROI, there was likely not enough power to identify any significant results at the individual condition level.

2.7.1.2.3. Super-Additive Method. Using the data extracted from the ROIs defined by the super-additive method (table 4, figures 7 and 8), a repeated measures (2 x 4 x 2) ANOVA with the factors of modality (synchronous audiovisual activation and the sum of the audio only and visual only conditions) for each content condition (social-linguistic, social-nonlinguistic, and both nonsocial-nonlinguistic conditions) in each hemisphere (left vs. right) was conducted with the seven participants that had both left and right ROIs. The overall 3-way interaction (modality

× condition × hemisphere) was not significant ($F_{(3, 18)} = 0.28$, $p = 0.84$, $\eta^2 = 0.04$). Further, there was no significant interaction for either the hemisphere by content condition ($F_{(3, 18)} = 1.01$, $p = 0.38$, $\eta^2 = 0.15$) or modality by content condition ($F_{(3, 18)} = 0.58$, $p = 0.64$, $\eta^2 = 0.09$). There was also no interaction between super-additive method and hemisphere ($F_{(1, 6)} = 0.08$, $p = 0.79$, $\eta^2 = 0.01$). Finally, there were no main effects of content condition ($F_{(3, 18)} = 1.8$, $p = 0.18$, $\eta^2 = 0.23$), super-additive method ($F_{(1, 6)} = 0.14$, $p = 0.72$, $\eta^2 = 0.02$), or hemisphere ($F_{(1, 6)} = 0.12$, $p = 0.74$, $\eta^2 = 0.02$).

To examine the contribution of each condition individually, eight two tailed t-tests were performed to determine if activation in the synchronous audiovisual condition was greater than the sum of the audio only and visual only conditions for each content condition in each hemisphere. As shown in table 4, the synchronous audiovisual activation was not greater than the sum of the audio only and visual only conditions in the left or right hemisphere in any of the content conditions. It is possible that, given the low number of participants identified with an ROI using the super-additive method, there was not enough power to detect differences in the individual conditions.

2.7.1.2.4. ***Temporal Synchrony Method.*** Using the data extracted from the ROIs defined by the temporal synchrony method (table 5, figures 9 and 10), a repeated measures (2 x 4 x 2) ANOVA comparing modality (synchronous audiovisual activation to asynchronous audiovisual activation) in each content condition (social-linguistic, social-nonlinguistic, and both nonsocial nonlinguistic conditions) across hemispheres (left vs right) was conducted with the 11 participants that had bilateral ROIs. The overall 3-way interaction (modality × condition × hemisphere) was not significant ($F_{(3, 30)} = 0.65$, $p = 0.58$, $\eta^2 = 0.06$). The hemisphere by temporal synchrony method interaction was also not significant ($F_{(1, 10)} = 2.12$, $p = 0.17$, $\eta^2 = 0.17$). The hemisphere by content condition interaction approached significance ($F_{(3, 30)} = 2.73$, $p = 0.06$, $\eta^2 = 0.21$). The main effect of hemisphere was significant ($F_{(1, 10)} = 5.09$, $p = .05$, $\eta^2 = 0.34$) whereby the left hemisphere (*M* = -0.02, *SE* = 0.03) displayed less activation overall than the right hemisphere (*M* = 0.08, *SE* = 0.05). However, given that the hemisphere by content condition approached significance and that the difference between synchronous and asynchronous is of relevance to the temporal synchrony method, this main effect using the mean of synchrony and asynchrony activation must be interpreted with caution. The synchrony by content condition interaction was significant ($F_{(3, 30)} = 2.96$, $p = .04$, $\eta^2 = 0.23$). Follow up pairwise comparisons (using the LSD correction) collapsed across hemispheres comparing the difference between synchronous and asynchronous activation between the content conditions revealed that the difference value was greater for the mousetrap (*M* = 0.32, *SE* = 0.07), social-nonlinguistic (*M* = 0.33, *SE* = 0.11), and social-linguistic (*M* = 0.40, *SE* = 0.13) conditions relative to the music condition (*M* = 0.01, *SE* = 0.09), $p = 0.03$, $p = 0.03$, $p = 0.05$, respectively.

In addition, eight two-tailed t-tests were performed to examine if the synchronous audiovisual condition was greater than the asynchronous audiovisual condition within each of the content conditions for each hemisphere. As shown in table 5, the synchronous audiovisual activation was greater than the asynchronous audiovisual condition in both the left and right hemispheres in the mousetrap condition, the social-linguistic, and the social-nonlinguistic condition, but not in either hemisphere for the music condition.

Given that the temporal synchrony method identified left and right ROIs in the greatest number of participants when collapsed across content conditions, and was upheld in the majority of the individual content conditions, additional analyses were undertaken to explore this new method. Two independent one-tailed t-tests were conducted to determine (1) if activation in the collapsed synchronous conditions was significantly above baseline (i.e., inter-stimulus intervals when the participants were staring at the fixation cross) and (2) if the activation in the collapsed asynchronous conditions was significantly below baseline. The synchronous audiovisual condition was not significantly greater than baseline in the left hemisphere ($t_{(12)} = -.09$, $p = .93$, Cohen's d = 0.03), but approached significance in the right hemisphere ($t_{(11)} = 2.11$, $p = .06$, Cohen's d = 0.74). The asynchronous audiovisual condition was significantly below baseline in the left hemisphere ($t_{(12)} = -4.46$, $p = .001$, Cohen's d = 1.24) and approached significance in the right hemisphere, ($t_{(11)} = -2.11$, $p = .06$, Cohen's d = 0.62). These findings indicate that the relationship between synchronous and asynchronous conditions must be considered, rather than each condition separately. Given this finding, the ANOVA that was run above was re-run using the difference values of synchronous and asynchronous (e.g., synchronous - asynchronous) in order to examine the difference between hemispheres and content conditions. The 2 (hemisphere) by 4 (content condition) ANOVA did not reveal any significant interactions between hemisphere and content condition ($F_{(3, 30)} = 0.66$, $p = 0.58$, $\eta^2 = 0.62$), nor was there a significant main effect of hemisphere ($F_{(1, 10)} = 2.11$, $p = 0.18$, $\eta^2 = 0.017$. The significant main effect of content condition was upheld ($F_{(3, 30)} = 2.~97$, $p = 0.05$, $\eta^2 = 0.23$), with the pairwise comparisons for the difference values being the same as above (e.g. the social-linguistic ($M = 0.40$, $SE = 0.13$), social-nonlinguistic ($M = 0.33$, $SE = 0.11$), and mousetrap ($M = 0.32$, $SE = 0.07$) all being greater than the difference value for the music condition ($p = 0.05$, $p = 0.03$, $p = 0.03$, respectively).

2.7.2. Whole-Brain Analyses

Whole-brain analyses were also completed in volume space in order to examine the subcortical structures that may also be involved in multisensory integration. The voxel-wise results of the group level analyses were corrected for multiple comparisons using 3dclustsim with an uncorrected p-value of .01 and a cluster size threshold of 20 voxels to obtain a corrected alpha values of less than .05. At the group level collapsed across conditions, the mean method (figure 11) revealed multiple large clusters in the bilateral thalamus, bilateral superior temporal gyrus, and the bilateral visual cortex. In addition, a smaller cluster was identified in the left fusiform gyrus. As can been seen in figure 11, the mean condition identified several areas with large clusters, but did not provide much specificity. Both the max and super-additive conditions did not reveal any clusters of activation (even when they were explored with significantly lower thresholds). Finally, the temporal synchrony method (figure 12) identified clusters bilaterally in the parahippocampal gyrus and posterior cingulate. Clusters were also identified in the left precuneus and the right superior frontal gyrus.

2.8. Discussion

As researchers have tried to extrapolate methods of identifying multisensory neurons from single-cell recordings to the non-invasive techniques used with humans, the research methodology has been fraught with limitations. In particular, inherent in the methodology used with humans is that the activation of neurons is studied in large neuronal populations. In one voxel, the large heterogeneous neuronal population may be composed of cells that respond differently to external stimuli. Closer examination of regions in the human brain that demonstrate multisensory integration has revealed a patchy distribution of neurons, only some of which respond to multisensory stimuli (Beauchamp et al., 2004). As a result, a voxel will not demonstrate the same response properties as single-cells. In essence, this precludes the applicability of the widely used single cell method of super-additivity in humans. Due to these methodological limitations, researchers have developed other methods that are less conservative and more likely to demonstrate a significant response than the super-additive method. The most common of these methods are the mean and max methods. However, as discussed above, these methods have theoretical limitations as they are not based on any known properties of multisensory neurons.

Given the empirical and theoretical limitations of previous research, the temporal synchrony method was proposed as an alternative method that would be more likely to identify multisensory regions than the super-additive method. Further, because the temporal synchrony method was based on the property of temporal synchrony identified at the single-cell level, it was also proposed to be more theoretically valid than the mean and max methods. As the temporal synchrony method has never been examined in the same study as the other methods, the purpose of the current study was to evaluate the three most common methods of identifying multisensory regions relative to the new temporal synchrony method. In order to evaluate and compare these methods, activation in the pSTS specifically was explored, as it is the most commonly identified area of multisensory integration in humans (Beauchamp et al., 2010; Beauchamp et al., 2004; Calvert et al., 2001; Nath & Beauchamp, 2012; Stevenson & James, 2009; Watson et al., 2014). The max, mean, super-additive, and temporal synchrony methods were examined at both the individual-level, using individually-defined ROIs, and the group-level using whole-brain analyses.

2.8.1. Group Level Whole-Brain Analyses

Based on the group-level whole-brain analyses collapsed across content conditions, it appears that both the super-additive and max methods were too conservative for the identification of any areas of activation. Similar to Calvert and colleagues' (2000) findings, the mean method appeared to identify structures responsible for unisensory processing (e.g. large clusters of activation in the auditory and visual cortices), not specifically multisensory integration (i.e. the mean method identified clusters in the visual and auditory cortices). At the whole-brain level, the temporal synchrony method did not identify the pSTS. Rather, areas involved in unimodal perception, contextual associations (Aminoff, Kveraga, & Bar, 2013) and attention regulation (Leech & Sharp, 2014) were identified. Multisensory stimuli have been found to play an important role in attention regulation and thus it is possible that the temporally synchronous audiovisual activation facilitated control of attention directed to the stimuli presentation. However, additional research is needed in order to further explore how these areas may be involved in networks related to multisensory integration.

2.8.2. Individually Defined ROI Analyses

The group analyses offered an overall examination of the methods across all participants after their scans had been transformed into standard brain space. However, each individual brain is unique and the transformation may reduce specificity and thus the critical effects (e.g., see Glezer & Riesenhuber, 2013; Stevens, Kahn, Wig, & Schacter, 2012; Stevens et al., 2015). For a more precise examination of the methods, the individual-level analyses that were completed in surface space must be considered in more detail. For the individual-level analyses, the methods were first examined collapsed across content conditions. As can be seen in the collapsed analyses, the temporal synchrony method identified a multisensory region in the pSTS in the largest number of participants bilaterally and in at least one hemisphere, with 82% of participants having at least one hemisphere identified. The super-additive method identified the second largest percentage of participants in at least one hemisphere (65%) and the mean method identified the third largest percentage of participants, with 59% of participants. While the mean method is thought to be the least conservative and would be assumed to identify regions in all participants, this overly liberal technique resulted in large areas of activation with peak activation falling outside of the pSTS. The current finding that the method is too liberal, which impairs the practicality of the method, is consistent with Love and colleagues (2011) and Beauchamp's (2005) findings. Finally, the max method identified the lowest percentage of participants in at least one hemisphere (18%). It is likely that the calculations required to create the max condition resulted in it being the most conservative. In order to create the max conditions, all negative activation had to be removed, and the multisensory activation had to be significantly greater than each of the unisensory conditions. These calculations resulted in creating a conjunction of two brain maps that were each thresholded at a $p < 0.03$: (1) where the multisensory response was significantly greater than the visual only response, and (2) where the multisensory response was significantly greater than the auditory only response. Applying these thresholds in addition to the removal of any negative activation likely resulted in small clusters that, when combined in the conjunction map, had few overlapping areas for both unisensory stimuli. Previous research using the max technique has used the minimum statistic method for conjunction analyses (e.g., Beauchamp, 2005), which has been proposed to be fraught with theoretical and methodological concerns (see Nichols, Brett, Andersson, Wagner, & Poline, 2004 for a review).

Calvert (2001) proposed that linguistic and nonlinguistic multisensory integration involve different structures and circuitry, and along with this reasoning, previous research has been equivocal with respect to finding multisensory activation in the pSTS in response to non-linguistic stimuli (Bushara et al., 2001; Calvert et al., 2000; Macaulso et al., 2004; Raj et al., 2000). Thus, it was important to examine the applicability of each method in the different content conditions. Three content conditions were used in the current study in order to examine potential differential findings, consistent with previous research using social-linguistic and nonsocial-nonlinguistic (tool) stimuli. In order to differentiate the linguistic and social components typically confounded in previous research, the three content conditions of particular relevance were social-linguistic, social-nonlinguistic, and the mousetrap nonsocial-nonlinguistic. An additional music nonsocial-nonlinguistic condition was also included, as it was relevant to analyses outside the scope of this study.

Using the ROIs created at the collapsed level for each participant, the activation within the ROIs for each of the content conditions was extracted. The calculation of the methods was then applied with the extracted values for each of the content conditions in each hemisphere (e.g. for the mean method, the synchronous audio-visual activation for social-linguistic stimuli was significantly greater than the mean of the audio only and visual only conditions for the social-linguistic stimuli). Using the mean method, activation in the synchronous audiovisual condition was found to be significantly greater than the mean of the audio-only and visual-only conditions in about half of the hemispheres per content condition. Therefore, although the mean method is thought to be the most liberal, greater activation for synchronous audiovisual than the mean of audio and visual only was not observed in each hemisphere for each condition. Further, as the max method only identified ROIs in three participants, it is not surprising that the max methodology was not upheld in the individual content conditions, given the lack of statistical power. Interestingly, the super-additive method was also not supported in any of the individual content conditions. This finding suggests that there may not have been enough power (e.g. too few trials) to find the super-additive effect once the trials were no longer combined. However, the temporal synchrony method was supported in three out of the four conditions, as indicated by the synchronous audiovisual activation being significantly greater than the asynchronous audiovisual activation in the social-linguistic, social-nonlinguistic, and mousetrap nonsocial-nonlinguistic conditions.

James and colleagues (2102) previously characterized this comparison of methods as a debate between theoretical and empirical approaches. Only the super-additive and temporal synchrony methods can be considered to be theoretically valid, as they are both based on known properties of multisensory neurons. The results of the analyses across collapsed conditions demonstrated that the temporal synchrony method was the most empirically practical, identifying at least one multisensory integration region in the greatest number of participants. Therefore, in contrast with previous studies (Beauchamp, 2005), the max method was not found to be empirically tenable when analyses collapsed across conditions, nor in individual content conditions. The super-additive method was found in the current study to be empirically practical at the overall collapsed level, but was not supported in the individual content conditions. The mean method was potentially empirically practical at the overall collapsed condition, but was not consistently supported in the individual content conditions. Only the temporal synchrony method was identified as theoretically valid and empirically practical at the collapsed and individual content condition levels. The only content condition that was not upheld at the individual level using the temporal synchrony method was the music nonsocial-nonlinguistic condition. However, this is not surprising, as at the neural level, music processing is thought to be dissociable from processing of speech and other activities (see Peretz & Zatorre, 2005), and thus may not be integrated in the same brain networks as non-music stimuli.

It is worthwhile to further compare the results of the super-additive and temporal synchrony methods, as they were previously identified to be the only methods involved in this study that were based on known properties of multisensory neurons. In fact, at the single-cell level, the super-additive response is dependent on the stimuli satisfying the temporal properties of integration (i.e. that the sensory cues are in close temporal proximity) (Meredith & Stein, 1986, 1996; Miller & D’Esposito, 2005). It is interesting that, given the close association between these methods, only the temporal synchrony method was upheld in the individual content conditions. The differing findings likely occurred for a few reasons. One of the most important reasons is that the underlying formula of the super-additive method requires that the multisensory response be greater than the sum of the unisensory responses. Due to the heterogeneity of cells at the voxel-level, super-additivity is unlikely to be found, as the unisensory neurons would degrade this calculation. Importantly, it appears that many trials need to be used in order to have enough power to detect this super-additive response. The temporal synchrony method is not as impacted by the heterogeneity of cells at the voxel level, as the non-multisensory neurons in the voxels will respond equivalently to all conditions. Thus, by subtracting the asynchronous condition from the synchronous condition, the noise of the non-multisensory neurons is removed. The fact that the temporal synchrony method identified multisensory regions across different stimulus types with fewer trials is an important finding that supports its use in future research. This finding makes the temporal synchrony method not only empirically tenable, but also more useful for participant populations that may fatigue quickly or have a difficult time staying in the MRI environment for extended periods of time.

Overall, the current study demonstrated that the temporal synchrony method was both theoretically valid and empirically tenable. There are some important considerations to note when using the temporal synchrony method. Unlike the previous methods that require the removal of suppression relative to baseline, the temporal synchrony method uses a comparison of activation of stimuli that differ only in one property. The examination of the collapsed synchrony activation relative to baseline and collapsed asynchrony activation relative to baseline revealed that the synchrony and asynchrony condition should be considered together as a relative difference value, rather than separately. It is the relationship between the synchronous and asynchronous conditions that is most relevant. In particular, unlike the other methods examined in this study, the temporal synchrony method compares two sets of stimuli that both contain the identical audio and visual information, and the stimuli only differ with respect to temporal synchrony. Theoretically, the use of the temporal synchrony method would result in both the synchronous and asynchronous conditions activating neurons that respond to auditory stimuli only, neurons that respond to visual stimuli only, and multisensory neurons. However, only the synchronous conditions would cause greater activation of the multisensory neurons than the asynchronous conditions. Thus, the comparison between synchronous and asynchronous stimuli is necessary to find areas of the human brain that are responsible for multisensory integration.

2.8.3. Conclusions

Previous research examining the neurological underpinnings of multisensory integration in human populations suffered from both empirical and theoretical limitations. Therefore the temporal synchrony method was proposed, and was found to be both theoretically and empirically tenable for social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic stimuli. Overall, given the theoretical and empirical support behind the temporal synchrony method, future research examining human development or atypical development of multisensory perception should consider the application of this method.

Study Two:

Posterior Superior Temporal Sulcus Activation During Perception of Multisensory Stimuli in Autism Spectrum Disorders

It has been proposed that sensory atypicalities are a hallmark deficit of individuals with Autism Spectrum Disorders (ASDs; Grandin, 1995; Klinger & Dawson, 1996; Ornitz, 1989). In particular, is has been proposed that these sensory atypicalities are caused by differences in multisensory perception (Iarocci & McDonald, 2006).

Multisensory perception can be subdivided into many aspects of sensory integration, including high-level (more complex stimuli), low-level (less complex, e.g. flash-beeps tasks), social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic. Results are equivocal with respect to the nature of multisensory perception deficits in individuals with ASD. In particular, it appears that the multisensory perception abnormalities are either limited to, or most evident for, linguistic related information (Bebko, et al., 2006; Bryson, 1972; deGelder et al., 1991; Smith & Bennetto, 2007). However, these studies failed to clarify if the multisensory deficits are unique to linguistic information or social information, as the linguistic stimuli used in previous studies are inherently social, such as a person telling a story or making vowel sounds. To date there has been no research that has parsed out the role that the social aspects in language play in multisensory perception.

3.1. Role of the STS

While the superior temporal gyrus contains the primary auditory cortex and has thus been identified as responsible for auditory processing, it has also been found to activate for silent speech reading (Calvert et al., 1997). The rostral lateral suprasylvian sulcus (rLS) in cats is thought to be an analogue to the STS in humans. The rLS has strong cortical connections with the SC, a structure that in cats and in humans integrates sensory information and directs attention (Calvert, et al., 2000; Stein & Meredith, 1993) and has consistently been implicated in multisensory integration (Wallace & Stein, 1994; Meredith & Stein, 1986). Importantly, research has demonstrated that these cortical areas play a key role in the development of multisensory enhancement in the SC (Wilkinson et al., 1996; Jiang, Jiang, & Stein, 2006).

In a closer parallel to humans, audiovisual integration has also been associated with the STS in rhesus monkeys (Ghazanfar, Chandrasekaran, & Logothotis, 2008). Therefore, it is not surprising that in humans, the STS has been largely implicated in audiovisual integration. While the STS has been proposed to be involved in many social-cognitive processes, the pSTS in particular has been implicated in audiovisual integration (Pelphrey et al., 2004).

The pSTS is the most consistently implicated structure in social-linguistic multisensory integration (Beauchamp, et al., 2010; Calvert et al., 2000; Macaluso et al., 2004; Nath & Beauchamp, 2012; Raij et al., 2001; Scott et al., 2000). The role of the pSTS in social-nonlinguistic and nonsocial-nonlinguistic processing is less well examined and equivocal. There is little research examining social-nonlinguistic processing as a separate stimulus category from social-linguistic processing. This author knows of only one study that has examined non-intelligible speech-like sounds; and interestingly, no significant pSTS activation was found (Scott et al., 2000). It is possible that social-nonlinguistic speech may only elicit activation of the pSTS when the stimuli are degraded causing inverse effectiveness (Stevenson, et al., 2007). Stevenson and James (2009) found that for both linguistic and non-linguistic stimuli, as the signal to noise ratio (SNR) of the stimuli relative to background noise decreased, the effect size, statistical significance, and multisensory enhancement of the BOLD response in the pSTS increased. However, Calvert and colleagues (2001), Beauchamp and colleagues (2004) found pSTS activation for nonsocial-nonlinguistic stimuli without degrading the signal. Therefore, it is possible that rather than the SNR of the stimuli (e.g. level of degradation of the signal) causing the differential findings, the methods used to identify the multisensory regions (e.g. mean, max, super-additive) and stimuli used to identify these regions contributed to equivocal findings. Importantly, selection of the pSTS region examined in the current study was informed by the results of study one, which demonstrated that the temporal synchrony method for the identification of regions responsible for multisensory integration was theoretically valid and empirically tenable in individuals without ASD.

3.2. Posterior STS and ASD

The pSTS has been implicated in many processes that have been found to be impaired in ASD, including theory of mind (Gallager & Frith, 2003; Pelphrey et al., 2004; Saxe, 2006; Saxe et al., 2004; Zilbovicious et al., 2006), face processing (Haxby, Hoffman, & Gobbini, 2000), gaze direction (Calder et al., 2007), and audiovisual integration (Amedi et al., 2005; Beauchamp, 2005; Calvert, 2001).

Structural abnormalities of the STS have also been identified in individuals with ASD. A thinner cerebral cortex (Hadjikhani, Joseph, Snyder, & Tager-Flusber, 2006) and reductions in grey matter (Boddart et al., 2004; McAlonan et al., 2002) have been identified in the STS of individuals with ASD compared to individuals without ASD. These reductions in volume have been found to be correlated with ASD symptomatology (Hadjikhani et al., 2006). Further, anterior shifting of the STS was found when comparing the cortical surface maps of individuals with ASD to those without ASD (Levitt et al., 2003).

Importantly, the pSTS has also been found to display atypical functional responses in individuals with ASD (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, 2007; Boddaert et al.,, 2004; Castelli et al., 2002; Redcay, 2008; von dem Hagen et al., 2011). However, there has been little direct testing of the neuronal basis of audiovisual integration in individuals with ASD and researchers have not elucidated the specific multisensory difficulties or neural correlates associated with the deficit in multisensory integration found in individuals with ASD. Therefore, the use of fMRI to explore the pSTS and multisensory integration in ASD is needed in order to further understand the neurological underpinnings of altered multisensory integration in ASD.

3.3. Current Study

The current study used the temporal synchrony method that was validated in study one to identify the multisensory region of the pSTS in young adults both with and without ASD. Using this multisensory ROI, the BOLD response to synchronous and asynchronous stimuli was compared between young adults with and without ASD, for social-linguistic, nonsocial-nonlinguistic, and social-nonlinguistic stimuli. It was expected that the young adults with and without ASD would show a similar BOLD response for nonsocial-nonlinguistic stimuli, with both groups showing a greater response in the pSTS to synchronous stimuli, and suppression of activity in the pSTS to asynchronous stimuli (i.e. temporal synchrony enhancement). In contrast, it was expected that the young adults with ASD would differ from those without ASD on social-linguistic processing. Specifically, young adults without ASD relative to those with ASD were expected to show enhanced activation in the pSTS to synchronous audiovisual social-linguistic stimuli, and suppressed activation to the same stimuli presented in an asynchronous format. In contrast, the participants with ASD were expected to show equivalent responses to the synchronous and asynchronous stimuli for the social-linguistic stimuli. The findings from the social-nonlinguistic condition were exploratory, in order to illuminate whether differences in multisensory integration in individuals with ASD were specific to linguistic-based information, or social information per se.

3.4. Methods

3.4.1. Participants

The current study involved two groups of participants: the seventeen high functioning young adults without ASD (13 males and 4 females) used in study one and 15 young adults with ASD (13 males and 2 females) matched on mean chronological age and full scale IQ (see table 6). All participants were between the ages of 18-29 years and had IQs above a standard score of 80, as measured by the Wechsler Abbreviated Scales of Intelligence two-item screener (WASI; Wechsler, 1999). Participants were required to be right handed, speak English as their first language, have normal hearing, normal or corrected to normal vision, and be neurologically healthy (i.e. free of severe or repeated concussions, brain injury, tumors, and seizures).

All adults with ASD had a previous diagnosis of an Autism Spectrum Disorder as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR or DSM-V) criteria. Diagnoses for each participant were confirmed through the use of the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter & Le Couteur, 1994) or the Autism Diagnostic Observation Scheduled (ADOS; Lord, Rutter, DiLavore, & Risi, 2002). Participants with ASD were recruited through community and private agencies and the Autism Spectrum Disorders-Canadian-American Research Consortium (ASD-CARC) participant database.

The young adults without ASD were recruited through posters, advertisements, and the York University Undergraduate Research Participation Pool (URPP). Participants recruited through the URPP received 2 credits towards course completion. All other participants with and without ASD were given a $10 gift card for their time and $10 for travel expenses.

3.4.2. Materials

3.4.2.1. Measures.

3.4.2.1.1. ***Autism Diagnostic Interview-Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003).*** The ADI-R is a semi-structured caregiver interview based on the DSM-IV and ICD-10 criteria for autism and pervasive developmental disorders. The ADI-R is composed of 93-items used to assess current and past behaviours related to ASDs such as an individual's quality of social interaction, communication and language, and repetitive, restricted and stereotyped interests and behaviours. The caregiver responses are coded and scored using an algorithm that provides scores for current and past behaviours, and a cut-off score for ASD.

The ADI-R has demonstrated sound reliability and validity. Inter-rater reliability for the domains, subdomains, and individual items were found to be high (majority with $\kappa > .7$; Lord et al., 1994; Poustka et al., 1996). Interclass correlational tests have indicated strong inter-rater reliability (majority ranging from $r = .82$-$.97$; Chakrabarti & Fombonne, 2001; Lord et al., 1994; Poustka et al., 1996). In addition, two studies have demonstrated strong test-retest reliability across two to five months ($r = .77$-$.97$; Hill et al., 2001; Lord et al., 1994). The ADI-R also demonstrates a strong sensitivity and specificity for differentiating individuals with ASD from typically developing individuals (Lecavalier et al., 2006).

3.4.2.1.2. Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 2002). The ADOS is a standardized semi-structured observation schedule utilized to assess and diagnose ASD across ages, developmental levels, and language skills. Through a series of semi-structured activities, the ADOS is used to assess social-interaction skills, socio-communication and language abilities, repetitive, restricted and stereotyped interests and behaviors, and imagination. The ADOS includes four modules, each requiring about 30-45 minutes to administer. The module is selected based on the participant's age and language level. Module 4 is designed for use with fluent adolescents and adults, and thus was used with all participants with ASD in the current study.

Previous research has consistently found the ADOS to have strong psychometric properties, including internal consistency, inter-rater reliability, and test-retest reliability (Lord, Risi, Lambrecht, Cook, Leventhal, DiLavre, et al, 2000). The ADOS also demonstrates strong sensitivity and specificity for differentiating individuals with ASD from typically developing individuals (Lord, et al., 2000).

Both the ADI-R and the ADOS have been frequently used in previous research as tools to confirm ASD diagnoses and as an outcome measure (Ozonoff, Goodlin-Jones, & Solomon, 2005).

3.4.2.1.3. Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). The WASI is a brief two to four subtest individually administered test of intelligence for children and adults. The two-subtest form was used in the current study, which consists of the Vocabulary and

Matrix Reasoning subtests. The two-subtest form takes approximately 15 minutes to administer. The WASI two-subtest form provides estimates of full scale, verbal, and performance intelligence. The WASI has demonstrated adequate reliability, including split-half reliability and test-retest reliability between 2 to 12 weeks. It also correlates well with other measures of intelligence, including the Wechsler Adult Intelligence Scale (Wechsler, 1999).

3.4.2.2. Stimuli.

In order to be comparable to previous behavioural research for analyses, several videos were used that were composed of three different content conditions from Lavoie et al. (in preparation): linguistic, social-nonlinguistic, and nonsocial-nonlinguistic. The linguistic stimulus displayed a woman telling a story, with only the head and neck of the woman visible, and her hair pulled back, to minimize distractions. The social-nonlinguistic stimulus involved the same woman making various sounds that do not have language content, such as tongue clicks and humming. Two nonsocial-nonlinguistic stimuli were used in the current study. The first nonsocial-nonlinguistic stimulus was composed of a bird's-eye view of a hand playing a children's song on a piano; this stimulus condition was included to facilitate comparisons with a previous study (Lavoie et al., in preparation). As music processing may be a discrete and dissociable process from other stimuli (see Peretz & Zatorre, 2005 for a review of brain organization for music processing), an additional nonsocial-nonlinguistic video was created for the current study that involved a marble going through a marble run (“mousetrap”), i.e., a series of tubes and steps that make various noises as the marble rolls. Each content condition was created with an auditory only, visual only, audiovisual synchronous, and audiovisual asynchronous condition. In the asynchronous conditions, the visual content was 500 miliseconds temporally ahead of the auditory content. All videos were 12 seconds long.

3.4.3. Experimental Design and Data Analysis

The current study used the same procedures for data collection, data processing, and creation of ROIs as used in study one.

3.4.4. Procedure

All participants were given written informed consent prior to participation and completed a screening form to ensure they were eligible to enter the MRI environment. Participants were in the scanner for between 40-90 minutes, depending on the number of runs they completed for participants with ASD (*M* = 8, *SD* = 1.16, range 6 - 10 runs) and without ASD (*M* = 8.22, *SD* = 1.22, range 6 - 10 runs). After completing the fMRI task, participants were administered the WASI. The participants with ASD were also administered the ADOS if a previous ADI-R had not been completed. All participants were then thanked for their time and provided with their compensation for participation (e.g. course credit or $10 gift card).

3.5. Results

3.5.1. Individual-Level ROI Analyses

The temporal synchrony method validated in study one was used to identify a multisensory region in the pSTS in each hemisphere in participants with ASD. Similar to study one with the participants without ASD, the region was identified by collapsing across all content conditions. Figure 13 displays the mean beta values (representing the amplitude of the BOLD response) for the ROIs collapsed across content conditions for each group for those who had an identifiable ROI using the temporal synchrony method. Table 7 displays the number of participants in each group for whom a pSTS ROI was identified in each hemisphere. The temporal synchrony method identified approximately the same number of participants showing these pSTS ROIs in the participants with and without ASD.

Using the ROIs identified at the individual level, the mean beta values were extracted for the synchronous and asynchronous conditions, both when collapsed across content conditions, and for the separate content conditions individually (depicted in Figures 13, 14, and 15). As was demonstrated in study one, the difference between synchronous and asynchronous conditions must be considered, rather than each synchrony condition separately. Therefore, a 2 (hemisphere: left vs. right) by 2 (group: ASD vs. control) factorial ANOVA was conducted with the synchronous/asynchronous difference value collapsed across content conditions. The overall hemisphere by group interaction was not significant ($F_{(1, 20)} = 0.63$, $p = 0.44$), nor were the main effects of hemisphere ($F_{(1, 20)} = 0.06$, $p = 0.80$) and group ($F_{(1, 20)} = 0.09$, $p = 0.76$).

The ROIs were also examined within each of the content conditions using a 4 (content condition) by 2 (hemisphere) by 2 (group) ANOVA. The overall hemisphere by content condition by group interaction was not significant ($F_{(3, 60)} = 0.87$, $p = 0.46$). There were also no significant interactions between hemisphere and condition ($F_{(3, 60)} = 0.45$, $p = 0.72$); hemisphere and group ($F_{(1, 20)} = 1.01$, $p = 0.33$); or condition and group ($F_{(3, 60)} = 1.17$, $p = 0.33$). Finally, there were no significant main effects of condition ($F_{(3, 60)} = 1.05$, $p = 0.37$), hemisphere ($F_{(1, 20)} = 2.06$, $p = 0.16$), or group ($F_{(1, 20)} = 0.11$, $p = 0.74$).

The analyses completed above depend on having a large sample size with both left and right ROIs identified. However, as shown in table 7, the requirement of having bilateral ROIs resulted in only 11 participants in each group, which may not provide enough power to detect any group, hemisphere, or content condition differences or interactions. When the temporal synchrony enhancement (i.e., synchronous > asynchronous) was examined for each individual content condition, the synchronous audiovisual BOLD response was found to be greater than the asynchronous audiovisual response in both the social-nonlinguistic and the mousetrap conditions in both hemispheres for both groups (Table 8; Figures 14, 15). Neither group showed temporal synchrony enhancement for the music condition in the left hemisphere, although it approached significance in the right hemisphere for the ASD group. However, while the group without ASD showed temporal synchrony enhancement for the social-linguistic condition in both the left and right hemispheres, the group with ASD did not show this in either hemisphere (Table 14; Figures 14, 15). Follow-up t-tests comparing the difference values (synchronous-asynchronous) between the groups for each content condition failed to identify statistically significant differences between groups (all *p*s > 0.05).

As it is possible that the analyses using bilateral ROIs and the analyses using separate hemispheres did not provide a large enough sample size to detect group differences, the results above were completed again collapsing across hemispheres. In participants that had bilateral ROIs, the peak voxel value for the collapsed ROI was used to select the hemisphere with the higher peak voxel value (see table 9 to view the number of participants in which the left or right hemisphere was included in the collapsed across hemisphere analyses using the peak voxel value). An independent two-way t-test comparing the synchronous/asynchronous difference value collapsed across content conditions between groups was performed, and no significant difference between groups was found ($t_{(25)} = -1.24$, $p = 0.23$). The ROIs were also examined within each of the content conditions using a 4 (content condition) by 2 (group) ANOVA. The content condition by group interaction was not significant ($F_{(3, 75)} = 1.32, p = 0.28$). There were also no significant main effects of condition ($F_{(3, 75)} = 1.59, p = 0.20$) or group ($F_{(1, 25)} = 0.02, p = 0.97$). The analyses were repeated again by using the max BOLD beta value to select the hemisphere that would be included for participants that had bilateral ROIs in each condition (see table 10 to view the number of participants for each condition in which the left or right hemisphere was included in the collapsed across hemisphere analyses using the max BOLD beta value). Again, the independent two-way t-test comparing the synchronous/asynchronous difference value collapsed across content conditions between groups found no significant difference between groups ($t_{(25)} = -1.30, p = 0.20$). The ROIs were also examined within each of the content conditions using a 4 (content condition) by 2 (group) ANOVA. The content condition by group interaction was not significant ($F_{(3, 75)} = 1.74, p = 0.16$). There were also no significant main effects of condition ($F_{(3, 75)} = 0.18, p = 0.91$) or group ($F_{(1, 25)} = 0.01, p = 0.93$).

In order to further explore these differing response profiles between groups in different conditions, each participant was examined for each condition in order to determine the frequency of individuals in each group that displayed a greater response to synchronous audiovisual stimuli relative to asynchronous audiovisual stimuli. Each participant was coded for each condition as either displaying the temporal synchrony enhancement (synchronous > asynchronous) or not displaying this enhancement (see table 11 to view the number of participants in each group who were classified as displaying temporal synchrony enhancement for each condition). There was no difference between the groups for the social-linguistic and social-nonlinguistic conditions in either hemisphere (social-linguistic left hemisphere- $\chi^2 (1, N = 24) = 0.99, p = 0.41$, social-linguistic right hemisphere- $\chi^2 (1, N = 24) = 1.20, p = 0.59$, social-nonlinguistic left hemisphere

χ^2 (1, N = 24) = 0.16, p = 1.0, social-nonlinguistic right hemisphere- χ^2 (1, N = 24) = 0.00, p = 1.0). Similarly, results of the analysis were not significant in the right hemisphere for the mousetrap nonsocial-nonlinguistic condition, χ^2 (1, N = 24) = 1.81, p = 3.71, or the right hemisphere for the music condition, χ^2 (1, N = 24) = 1.60, p = 0.40. The chi-squared analyses approached significance in the left hemisphere for the mousetrap condition, with the trend towards the group without ASD being more likely to display temporal synchrony enhancement relative to the group with ASD, χ^2 (1, N = 24) = 3.94, p = 0.07. The group with ASD was found to be 6.75 times more likely to be classified as displaying temporal synchrony enhancement in the left hemisphere for the music condition compared to the group without ASD, χ^2 (1, N = 24) = 4.89, p = 0.05.

3.5.2. Whole-Brain Group Analyses

Whole-brain analyses were also performed to determine if different regions were activated across the two groups. Thus, a whole-brain analysis was conducted at the group-level collapsed across conditions. The analysis compared the difference between synchronous audiovisual and asynchronous audiovisual conditions between the participants with and without ASD. The difference values were created by subtracting the BOLD response to asynchronous audiovisual stimuli from the BOLD response to synchronous audiovisual stimuli (i.e. synchronous AV - asynchronous AV). Therefore, higher positive difference values indicate that the BOLD response to synchronous was greater than to asynchronous conditions and lower difference values indicate that the synchronous and asynchronous conditions were more similar. Further, negative difference values indicate that the BOLD response to the asynchronous condition was higher than the BOLD response to the synchronous condition. The voxelwise results of the group-level analyses were corrected for multiple comparisons using 3dclustsim with a voxelwise threshold of $p < 0.01$ and a minimum cluster size threshold of 20 voxels to obtain corrected alpha values of less than 0.05. There were several clusters where there was a larger positive difference for the group without ASD relative to the group with ASD, including within the bilateral posterior and anterior cingulate (more notably left than right), bilateral parahippocampal gyrus (more notably right than left), left precuneus, bilateral precentral gyrus, left fusiform gyrus, left middle temporal sulcus, and left middle occipital sulcus (Figure 16). Most notable for the current study given the emphasis on examining multisensory integration in the pSTS, was a large cluster in left pSTS (Figure 17). It is also important to note that there were no areas where the group with ASD displayed a significantly greater temporal synchrony enhancement than the group without ASD.

Using the cluster identified in the pSTS at the whole-brain level, an ROI was created around the peak value (x = -57, y = 42, z = 27), with a radius of 10mm and including only voxels that were above a threshold of $p < 0.01$. Using this ROI, the means of the beta values for the synchronous and asynchronous conditions were extracted (Figure 18 demonstrates the difference between synchronous and asynchronous beta values). A 4 (condition) by 2 (group) factorial ANOVA comparing the difference value for synchrony (synchronous AV - Asynchronous AV) was performed. The condition by group interaction was not significant ($F_{(3, 90)} = 0.22$, $p = 0.88$). There was a significant main effect of condition ($F_{(3, 90)} = 3.62$, $p = 0.02$, $\eta^2 = 0.11$), whereby the social-linguistic ($M = -0.002$, $SD = 0.005$), social-nonlinguistic ($M = -0.001$, $SD = 0.003$), and the music nonsocial-nonlinguistic ($M = -0.001$, $SD = 0.005$) conditions all differed from the mousetrap nonsocial-nonlinguistic ($M = 0.001$, $SD = 0.003$) condition (all ps < 0.05). There was also a significant main effect of group, whereby the group without ASD ($M = 0.001$, $SE = 0.000$) displayed a greater difference value than the group with ASD, ($M = -0.002$, $SE = 0.00$), $F_{(1, 30)} = 14.82$, $p = 0.001$, $\eta^2 = 0.33$.

The evaluation of the group means for each individual condition revealed that while there was no overall interaction between group and condition, the groups themselves did not have the same pattern of responses to each of the conditions. Likely due to the low power as a result of low sample size and potentially too few trials per condition, differential group profiles did not emerge. T-tests comparing the synchronous and asynchronous audiovisual stimuli revealed that the difference value was significantly greater for the group without ASD relative to the group with ASD for the mousetrap condition, $t(30) = 2.45$, $p = 0.02$ and the social-linguistic condition, $t(30) = 3.25$, $p = 0.003$. There were no significant differences between the groups for the music condition, $t\ (30) = 1.69$, $p = 0.10$ or the social-nonlinguistic condition, $t(\ 30) = 1.36$, $p = 0.18$. However, these analyses revealed a trend towards group differences for the latter two conditions, particularly for the music condition.

3.6. Discussion

The purpose of the current study was to use the temporal synchrony method validated in study one to evaluate the neurological process of multisensory integration in individuals with ASD across several conditions that varied in terms of social and linguistic content. It was proposed that, relative to the group without ASD, the group with ASD would display similar multisensory integration related activation for nonsocial-nonlinguistic stimuli, but would not display temporal synchrony enhancement (synchronous audiovisual BOLD response greater than asynchronous audiovisual BOLD response) for social-linguistic stimuli. The social-nonlinguistic stimuli were also included to explore whether or not this specific deficit for social-linguistic multisensory perception in ASD, which has been observed at the behavioural level, would extend to social processing more broadly. The results were explored in two ways: with ROIs identified at the individual-level, and with group-level whole-brain analyses.

3.6.1. Individual-Level Analyses

At the individually defined ROI level, the temporal synchrony method collapsed across content conditions identified an ROI in the same proportion of participants in both groups. Analysis of the relationship between temporally synchronous and temporally asynchronous audiovisual stimuli in each of the content conditions revealed that there was significant temporal synchrony enhancement in the group without ASD for the social-linguistic, social-nonlinguistic, and mousetrap nonsocial-nonlinguistic conditions. While the group with ASD displayed similar significant temporal synchrony enhancement in the social-nonlinguistic and mousetrap nonsocial-nonlinguistic conditions, they did not display a significant temporal synchrony enhancement for the social-linguistic condition. However, there were no significant differences between the groups when directly comparing the difference values (synchronous - asynchronous) for each content condition (i.e., no significant group by synchrony interactions), and so these findings must be interpreted with caution.

3.6.2. Group-Level Analyses

Examining the whole brain at the group-level, a significantly different pattern of multisensory integration was found in the left pSTS between the group without ASD and the group with ASD. Collapsed across content conditions, the group without ASD displayed significantly greater temporal synchrony enhancement in the pSTS relative to the group with

ASD. This difference between the synchrony conditions was characterized by a greater BOLD response to synchronous audiovisual stimuli than asynchronous audiovisual stimuli in the group without ASD only.

In order to understand what was driving this difference between groups, further analyses were conducted by creating an ROI in the area of the left pSTS that was identified as displaying a significant difference between groups when collapsed across conditions. This ROI was then used to explore differences between synchronous and asynchronous audiovisual activation between the groups across the various content conditions. As shown in Figure 18, the group without ASD displayed a pattern of activation where the BOLD response for the temporally synchronous audiovisual stimuli was greater than the BOLD response for the temporally asynchronous audiovisual stimuli in the nonsocial-nonlinguistic and social-linguistic conditions. In contrast, the group with ASD displayed no difference in BOLD response to synchronous and asynchronous audiovisual stimuli for the mousetrap nonsocial-nonlinguistic condition, and in fact displayed the reverse effect, with greater activation to the temporally asynchronous than synchronous stimuli in the music nonsocial-nonlinguistic and social-linguistic conditions. In other words, the group with ASD displayed either a lack of multisensory enhancement to temporally synchronous stimuli or a completely reversed response. These differential patterns of response to synchronous vs. asynchronous stimuli were significantly different between the groups for the social-linguistic and the mousetrap nonsocial-nonlinguistic stimuli. In both groups, there was a greater BOLD response to asynchronous audiovisual stimuli compared to the synchronous audiovisual stimuli in the social-nonlinguistic condition. An explanation for why the individual level results may have differed from the group-level results is provided later with respect to the discussion of the temporal synchrony method.

3.6.3. Atypical Multisensory Integration in ASD

The finding that individuals with ASD failed to display temporal synchrony enhancement (synchronous > asynchronous) at the neuronal level for both social-linguistic and nonsocial-nonlinguistic stimuli suggests that at an underlying neurological level, individuals with ASD are not effectively integrating audio and visual information. It may be that individuals with ASD are integrating information over a wider temporal window, meaning that they may be perceiving discrete sensory experiences as one event, or they may be perceiving single sensory experiences as discrete and separate events. Evidence has been found for individuals with ASD binding simple sensory stimuli (i.e. flash-beep tasks) over a wider temporal window in behavioural studies (Foss-Feig et al., 2000). In contrast, the lack of susceptibility of individuals with ASD to the McGurk effect (an auditory visual illusion that relies on audio-visual multisensory integration) (de Gelder et. al., 1991), may suggest that auditory and visual stimuli are perceived as discrete events. The current study used a temporal asynchrony of 500 millseconds, which is at the end of the temporal window that has been tested in the flash-beep tasks, which have used 25-500 ms delays in auditory-visual stimuli pairing. If the theory of the wider temporal window is correct, the current results would indicate that individuals with ASD may be ineffectively integrating stimuli over 500 milliseconds apart. If the theory of a lack of integration (e.g. discrete sensory events) is correct, then it would suggest individuals with ASD were not integrating the sensory stimuli in the current study. The current study did not vary the timing of the temporal asynchrony and can therefore not determine if it is a lack of integration or a wider temporal window of integration. However, regardless of the reasons for why the individuals with ASD displayed less effective integration, atypical integration would have a significant impact on their perception of their environment.

The atypical temporal multisensory integration demonstrated by individuals with ASD in the current study is an area of particular concern based on what is known about the development of multisensory integration abilities. The ability to effectively integrate and differentiate perceptual events in the environment develops in a hierarchical fashion and temporal synchrony is thought to be the foundation of this development (Lekowicz, 1999; 2000; Radeau, 1994; Stein & Meredith, 1993; Welch & Warren, 1980). Therefore, the fact that individuals with ASD are not showing the same neurological underpinnings for the temporal integration of audiovisual information may have significant consequences for the development of the ability to integrate sensory experiences with more complex forms of amodal information. If individuals with ASD are not using amodal properties such as temporal synchrony to differentiate and integrate sensory experiences (especially those that are social-linguistic in nature), there would be profound consequences for many aspects of cognitive, language, and social development and functioning. Thus, it is not surprising that difficulties with multisensory perception have been found to correlate with ASD symptomatology (Donohue et al., 2012).

In typically developing populations, the ability to integrate and differentiate sensory experiences has been demonstrated with 10- to 16-week-old infants (Dodd, 1979). At a behavioural level, children with ASD have been found to show multisensory perception deficits at 12 years of age (Loveland et al., 1995). The current study demonstrated that in early adulthood individuals with ASD are continuing to display atypical multisensory integration at a neurological level. Future research should explore from a developmental perspective if these differences in multisensory integration emerge in infancy alongside the development of multisensory perception in infancy, or if they become more discrepant from individuals without ASD as they age. Given that typical development is based on a model of increasing specificity, whereby infants learn to discriminate using more complex forms of amodal information (Bahrick, 2001), it would be important to understand when this process of multisensory integration either fails to happen or reverses (e.g. becomes less specific) for individuals with ASD.

It is also important to note that the findings of individuals with ASD displaying atypical integration of both social-linguistic and nonsocial-nonlinguistic stimuli do not support the hypothesis of linguistic-specific multisensory perception deficits (Bryson, 1972; Bebko et al., 2006). Interestingly, using similar stimuli to the current study, Lavoie and colleagues (in preparation), found that children with ASD did not integrate linguistic multisensory information in the same way as a typically developing group without ASD. The current study did not have a behavioural component that measured if participants could perceptually discriminate between synchronous and asynchronous stimuli. It is possible that individuals with ASD are not showing an automatic response at the neuronal level to integrate temporally synchronous information, but that they are able to use other networks and experience to compensate. This would mean that individuals with ASD may be able to use non-automatic compensatory strategies in behavioural studies, such as the conscious direction to attend to visual cues that would assist in matching auditory and visual stimuli based on temporal synchrony. Given the social deficits of individuals with ASD, and the finding that they process important emotional and social information from faces differently (Adolphs, Sears, & Piven, 2001), it is also possible that these compensatory strategies are not strong enough to deal with the complexity of human social interactions and language processing.

Alternatively, the role of atypically multisensory integration in the pSTS may disproportionally impact other networks responsible for higher-order processing. This study demonstrated that response to temporally synchronous and asynchronous multisensory stimuli was atypical for individuals with ASD in one specific area (the pSTS) that is thought to play an important role in the integration of sensory information. It is unclear from this study how this region interacts with other networks to impact what is observed at the behavioural level, and therefore further analyses examining whole-brain patterns of covariance are essential. Calvert (2001) proposed that there is different multisensory circuitry for linguistic and non-linguistic stimuli. In the current study the pSTS appeared to be implicated in both linguistic and non-linguistic stimuli. It will be important to examine how the pSTS works with other brain regions in a network to potentially differentially processes linguistic and non-linguistic stimuli. If one brain region is failing to develop typically (either structurally or functionally), it would likely impact the development of other brain regions that are connected to it (Perlman, Hudac, Pegors, Minshew, & Pelphrey, 2010). The impairment of multisensory integration in the pSTS could disproportionally affect linguistic multisensory networks. This would lead to the behavioural findings of atypical multisensory integration for linguistic information.

In addition to multisensory integration, the pSTS has been more broadly implicated in many processes underlying deficits of individuals with ASD, including theory of mind (Gallager & Frith, 2003; Pelphrey, et al., 2004, Saxe, 2006; Saxe et al., 2004; Zibovivious et al., 2006), face processing (Haxby et al., 2000) and gaze direction (Calder, 2007). There are several different explanations for why the pSTS may be implicated in these different processes. It is possible that these different deficits of individuals with ASD are each related to distinct areas within the pSTS. Alternately, the same area of the pSTS may be critically involved in each of these processes by acting as a hub in different neural networks that underlie each of these processes. Further research examining the functional circuits and the unique impact of impaired multisensory integration on the development of other social-cognitive processes is required in order to understand the connection between altered pSTS processing and this cluster of unique ASD-specific deficits.

3.6.4 Multisensory Processing and Social-Nonlinguistic Stimuli

The nonsocial-nonlinguistic stimuli were included in the current study to investigate whether the linguistic-specific deficits in multisensory integration that is seen at the behavioural level for individuals with ASD was due to the linguistic content, which is inherently social, or social multisensory information more broadly. It is interesting that both the participants with and without ASD did not display synchronous multisensory enhancement for the social-nonlinguistic stimuli in the left pSTS. While this may be due to the fact that the stimuli lacked language information, and thus may have been more effectively integrated for both groups in the right pSTS, this does not explain why the left pSTS displayed multisensory enhancement for the mousetrap nonsocial-nonlinguistic stimuli for the group without ASD. Therefore, it appears that there is something unique about the social-nonlinguistic stimuli, such that the pSTS is not involved in the processing of this type of stimuli. Interestingly, Scott and colleagues (2000) also found no activation of the pSTS to non-intelligible speech-like sounds. While, it has been proposed that the SNR of nonlinguistic stimuli must be low in order to elicit greater activation (i.e. inverse effectiveness; Stevenson, 2009); this hypothesis is not supported by the finding that nonsocial-nonlinguistic stimuli activated the pSTS in this study and in previous research (Calvert et al., 2001; Beauchamp et al., 2004; Watson, 2004).

The differential response profile of both groups to the social-nonlinguistic stimuli, relative to the other content conditions, suggests that another area of the brain may be responsible for processing this type of information. As the pSTS has been proposed to have many segments that are involved in many different processes, it is possible that another area of the pSTS may be more involved in multisensory integration of social information that does not contain language. Given that the whole-brain analysis was conducted collapsed across conditions, this potential social-nonlinguistic specific area of the pSTS would not have been identified. Further, this differential response profile to the social-nonlinguistic stimuli indicates that examining coordinated activation across networks, rather than specific cortical areas in isolation, is an important next step. Further research is needed in order to determine if social information that does not contain language is processed differently from social-linguistic information, or if there was something unique about the stimuli used in the current and previous studies (e.g. Scott et al., 2004).

3.6.5. Method of Identifying Multisensory Integration

The group differences and findings in the current study differed depending on the method used to examine multisensory integration. While there were no differences between the groups within the individually defined ROIs, there were differences in the group-level whole-brain analyses. Although it may appear that these two sets of results are incongruent, the method used to identify the individual-level ROIs must be examined further to understand these differential findings. The individually defined ROIs were created by looking for areas that displayed greater activation to synchronous stimuli relative to asynchronous stimuli and the whole-brain group-level analyses demonstrated that the individuals with ASD actually display an inverted response to temporally synchronous vs. asynchronous audiovisual stimuli. Therefore, as a result of the way the ROIs were identified at the individual-level, this seemingly paradoxical response pattern would not have been identified in the individual level analyses. This indicates that the results are not necessarily incongruent, but rather, that the methods used in the individual analyses did not allow for the method to identify this paradoxical response. Further, this suggests that the temporal synchrony method calculation used to identify the individual level ROIs needs to be adapted for use with atypical populations in order to allow for the identification of these paradoxical response patterns. Future research should define ROIs at the individual-level by looking at absolute relative activation differences for synchronous vs. asynchronous stimuli in the pSTS, which would include identifying ROIs for individuals who display this inverse response to temporally synchronous and asynchronous information in the pSTS.

3.6.6. Conclusions

Individuals with ASD appear to have an atypical response profile to the temporal synchrony of both social-linguistic and nonsocial-nonlinguistic multisensory stimuli in the pSTS. Given that integration of sensory information based on temporal synchrony is the foundation for multisensory integration, the finding that individuals with ASD failed to display temporal synchrony enhancement may have significant implications for other more complex forms of multisensory integration. Further, the atypical multisensory integration at the neural level of individuals with ASD has significant consequences for the development of higher order social, cognitive, and language processing. The development of these differential response profiles to synchronous and asynchronous stimuli needs to be more fully understood. Importantly, in order to understand these profiles, we first need to understand the underpinnings of social-nonlinguistic processing in typically developing populations.

Further, the results of the current study are at odds with the hypothesis of a linguistic-specific multisensory perception deficit in individuals with ASD. However, it is possible that the pSTS may be critically involved in neural networks that significantly impact linguistic processing, which would explain the previously observed linguistic-specific atypical multisensory perception identified at the behavioural level in individuals with ASD. Further exploration of the possibility that the pSTS acts as a hub or critical relay in networks responsible for processing multisensory events, and how these networks may be altered in individuals with ASD is needed.

Study Three:

Exploring Whole-Brain Covariance to Investigate Multisensory Perception in ASD

In the quest to understand the neurological mechanisms that underlie multisensory perception, researchers have narrowed their focus to specific brain regions. In particular, the pSTS has been largely implicated and explored as a critical region for multisensory integration (Beauchamp, et al., 2010; Calvert et al., 2000; Macaluso et al., 2004; Nath & Beauchamp, 2012; Raij et al., 2001; Scott et al., 2000). However, human cognition and perception are complex processes that likely do not rely on individual brain regions operating in isolation. McIntosh (2004) argues that in order to understand the connection between neural activation and mental functions, the neural context has to be considered. The neural context hypothesis states that involvement of a specific brain region in a cognitive process is dependent on the concurrent activity of other regions with which it is connected. Thus, the same region can be involved in different cognitive tasks depending on which other regions are co-activated (McIntosh, 2004). Exploring the neural networks that underlie cognitive functions becomes particularly important when examining ASD, as it has been characterized as a disorder of abnormal connectivity (Courchesne & Pierce, 2005; Maximo, Cadena, & Kana, 2014, Uddin, Supekar, & Menon, 2013).

4.1. Connectivity and Cognitive Functions

McIntosh (2004) proposed that there are two important aspects of the neural context: anatomical connectivity and response plasticity. With respect to anatomical connectivity, researchers have found that the neurotypical brain tends to have dense local connections and sparse distal connections (Felleman & Van Essen, 1991; Scannell, Burns, Hilgetag, O'Neil, &

Young, 1999; Sporns, Tononi, & Edelman, 2000; Stephan et al., 2000). This configuration creates regions that have unique inputs and outputs and also creates parallel routes of information transfer between areas that have similar connections, which allows the neural system to encode significant amounts of information (McIntosh, 2004; Sporns et al., 2000; Tononi, Sporns, & Edelman, 1994; Tononi, Sporns, & Edelman, 1996). Response plasticity is the process that allows neurons to respond variably, depending on the nature and source of afferent stimulation that causes these neurons to fire. Thus, neurons can show rapid and transient shifts in response to differential stimulation (McIntosh, 2004).

Taken together, response plasticity and anatomical connectivity create an environment in which there is considerable flexibility in the neural responses across connected structures. Thus, there is an interaction between the external world and the internal neural responses. This structure means that examining regional activation may yield ambiguous results regarding the role that a particular region has in a specific cognitive operation (D'Esposito, Ballard, Aguirre, & Zarahn, 1998) and that the overall pattern of activity across regions may be a more appropriate and meaningful measure (Haxby et al., 2001). For example, a study by McIntosh Rajah, and Lobaugh (2003) demonstrated that the medial temporal lobe was involved in task facilitation regardless of whether or not participants were aware of a tone predicting a visual event. However, the medial temporal lobe demonstrated differential connectivity with other regions depending on whether the participants were aware or unaware of this predictive pattern (McIntosh et al., 2003). These findings lead McIntosh (2003, 2004) to propose that the medial temporal lobe is a "behavioural catalyst", which means that it is responsible for transmitting information between regions or enabling the shift from one network to another.

4.1.1. Connectivity of the pSTS

The pSTS has been implicated in many social-cognitive processes (Allison, Puce, & McCarthy, 2000; Calder et al., 2007; Hein & Knight, 2008; Schultz, Friston, O'Doherty, Wolpert, & Frith, 2005). It is possible that the pSTS is also a behavioural catalyst, as described by McIntosh (2004), and thus plays an important role in different functional networks depending on task-demands. For example, social processing is thought to involve the coordinated functioning of the anterior temporal lobes, ventral and medial prefrontal cortex, the pSTS/temporoparietal junction (associated with theory of mind), the fusiform gyrus (associated with face processing), anterior and posterior cingulate, parts of the insula (emotional responses), and somatosensory cortices (action understanding) (Adolphs, 2009; Amodio and Frith, 2006; Carmichael and Price, 1995; Frith, 2007; Gotts et al., 2012; Hamilton & Grafton, 2006; Maximo et al., 2014; Schipul, Keller, & Just, 2011; Singer, Kiebel, Winston, Dolan, & Frith., 2004; von dem Hagen et al., 2011). In addition to its role in social processing generally, there is compelling evidence that the pSTS is involved in understanding biological motion and it's intent, and is critically involved in multisensory integration. The hypothesis that the role of the pSTS in multisensory integration might depend on its transient interactions with a broader network of regions has been largely under-explored. Thus, it is crucial to examine areas that may be co-activated with the pSTS in order to identify its role in multisensory integration.

4.2. Development of Networks

To understand these complex networks and how they may be abnormally structurally and/or functionally connected in atypical populations, it is important to consider the complex process of network development and the factors underlying aberrant development. Early brain development is characterized by the “delicate balance between the functional specialization of specific regions as well as the formation of connections across these regions through integration” (Maximo et al., 2014 p.17). Over development, there is a shift from a focus on local connectivity to more distributed architecture (Fair et al., 2009; Kelly et al., 2009; Maximo et al., 2014). While short connections may be responsible for communication between nearby neuronal populations, long distance connections are likely responsible for playing a larger role in perception and information processing (Maximo et al., 2014). Therefore, these long-distance connections are important for linking association areas across the cortex that are involved in higher-order cognitive processes, such as social-cognitive processing, attention, memory, and language, which are all functions that are impaired in ASD (Maximo et al., 2014).

4.2.1. Atypical Neural Development in ASD

Researchers have found multiple genes that are related to ASD and ASD-like symptoms. There is some evidence of a relation between ASD and loci on chromosomes 2 and 7, but these results have only been partially replicated (Rutter, 2005). The 15q11-q13 loci, which are responsible for Angelman and Prader-Willi syndrome, have also been implicated in ASD. In ASD research, 15q11-q13 abnormalities have been linked to seizures, low levels of hippocampal $GABA_A$ (gamma-Aminobutyric acid) receptor binding, and low levels of platelet GABA (Blatt et al., 2001). 15q11-q13 has also been shown to play a role in building cortical columns and column arrangement in the brain (Belmonte, et al., 2004; Jiang, Tsai, Bressler, & Beaudet, 1998). Cortical columns are the basic functional units that form the fundamental organization of the cortex. If the cortical columns of individuals with ASD are structured and arranged differently, the development of structurally and functionally connected networks would be impacted. The presentation of gene abnormalities is heterogeneous in ASD, and a set of ASD-specific genetic anomalies has not yet been confirmed. However, the genes that have been identified as possible contributors to the development of ASD are also associated with neural development.

The specific abnormalities in structural morphology and connectivity associated with ASD have been inconsistent in previous research. It has been proposed that atypical changes in brain volume occur across development for individuals with ASD, and research findings are dependent upon the age of the participant groups used in the research (Akshoomoff, Pierce & Courchesne, 2002). Despite potentially having a smaller head at birth, infants and young children with ASD have larger brain volumes (Courchesne, Carper,& Akshoomoff, 2003; Redcay & Courchesne, 2005; Hazlett et al., 2005; Hultman, Sparen & Cnattingius, 2002), and post-mortem studies of children with ASD reveal an excess of neurons in the prefrontal cortex (Courchesne et al., 2011). In fact, larger brain volumes are found in 90% of 2- to 4-year-old children with ASD. Specifically, the enlarged brain volumes are caused by increases in both grey and white matter (Bartholomeusz, Courchesne, & Karns, 2002). The overgrowth of grey and white matter appears to progress in an anterior to posterior direction (Hughes, 2007). Taken together, these findings suggest that following a period of normal or slightly reduced growth, individuals with ASD display rapid overgrowth (Lewis & Elman, 2008). Many researchers have reported that the increased cortical volume in early childhood seems to be followed by a lack of growth throughout the lifespan (Hughes, 2007). Therefore, ASD is posited to be a disorder of growth regulation (Akshoomoff, Pierce & Courchesne, 2002). Importantly, the degree of abnormality in the growth rate has been connected with symptom severity (Akshoomoff et al., 2004; Courchesne et al., 2003; Deutsch & Joseph, 2003; Tager-Flusberg & Joseph, 2003)

DiCicco-Bloom and colleagues (2006) postulated that white and grey matter volume differences could result from a few abnormalities including changes in: “(1) the numbers and sizes of neurons and glia; (2) the elaboration of axons and dendrites and synapses; (3) axodentrical pruning; (4) programmed cell death; (5) production of cortical columns; and (6) myelination” (p. 6899). The increase in grey matter is most likely related to a failure of synaptic pruning, which is the necessary loss or destruction of neurons not needed for processing in the individual's current environment (Hughes, 2007). Interestingly, the increase in white matter is related to abnormal myelination, where not only is there a delay in compaction of white matter, but also the biochemical composition of white matter may actually differ in children with ASD (Akshoomoff et al., 2002). These aberrant growth patterns are thought to create changes in the organization of neurons, including abnormalities in cortical column structure and increases in axonal length, which results in reduced efficiency of axonal conduction. Specifically, children with ASD show an increase in the total number of cortical columns, but the columns are smaller and less densely packed than typically developing individuals (Casanova, Buxhoeveden, Switala, & Roy, 2002). This structure would lead to reduced network differentiation, “noisy” communication between regions, and shorter connectivity fibres (Belmonte et al., 2004; Casanova et al., 2006; Maximo et al., 2014; Rubenstein & Merzenich, 2003). This increase in column number has been postulated to be responsible for the low SNR found in research utilizing ERPs, where the signal of an electrode measuring neural activation thought to be related to a particular stimulus is less strong relative to the noise of the other electrodes measuring brain activation in other, unrelated regions (Cohen, 1994). Belmonte (2004) proposed that the low SNR in neural processing will cause a range of behavioural changes resulting in the wide range of symptoms associated with ASD. In particular, this structure could cause over-aroused and under-selective primary processing, which would overload later higher order cognition, and thus as an adaptation, the brain would place an emphasis on low-level features over high-level integrative processing (Belmonte, 2004).

Abnormalities in brain growth may also cause abnormalities in axon structure and the efficiency of connections in the brain. Lewis and Elman (2008) propose that because the conduction speed of myelinated axons is dependent on the diameter and the length of the axon (Waxman, 1977), and because larger brains have been found to have longer axons (Olivares, Montiel & Aboitiz, 2001), the connections in the brain of someone with ASD may be less efficient. Lewis and Elman further propose that in abnormal brain development, when the brain is abnormally small, the long distance connections would be relied on, and when the brain is abnormally large, it is the short distance connections that would be relied on. This means that brains that undergo developmentally inconsistent growth patterns will show a fluctuation in connectivity patterns across the course of development. This would lead to rapid changes in connectivity, which may mean that one connection might be abandoned in favour of another connection, and thus, cognitive processes that relied on the previous connection will be lost, causing behavioural abnormalities (Lewis & Elman, 2008).

4.2.1.1.Task-based functional connectivity in ASD. Abnormal brain growth patterns in ASD would have significant impacts on functional connectivity. Some researchers have postulated that the large brain volume, abnormal cortical column structure, and aberrant white matter growth in the first years of life may cause structural under connectivity of several cortical areas (Courchesne et al., 2011; Hazlett et al., 2005; Just, Cherkassky, Keller, Kana, & Minshew, 2007; Rippon, Brock, Brown, & Boucher, 2007; Sparks et al., 2002). Other researchers have posited that a lack of synaptic pruning could result in the preservation of unneeded structural connections (Hughes, 2007; Lewis and Elman, 2008). Further, some have proposed a combination of both structural over- and under- connectivity, as the overgrowth and microstructural abnormalities result in ineffective short-distance over-connectivity, and long-distance under-connectivity (Belmonte et al., 2004; Courchesne & Pierce, 2005).

Relevant to the current study, disturbances in functional connectivity have been noted between the visual cortex and frontal cortex (Villalobos et al. 2005); between striatal sub-regions and heteromodal association and limbic cortices, including insula and superior temporal gyrus (Di Martino et al., 2011); and among areas involved in language processing (Dinstein et al. 2011; Shih et al., 2010). Further, if the multisensory perception deficits observed in ASD are limited to socially-based stimuli, then the abnormal connectivity patterns for social cognition (see Just et al., 2012; Kleinhans et al., 2008; Kana et al., 2009) may also be relevant.

Recent reviews have attempted to understand the reasons for these equivocal results of under or over connectivity. Muller and colleagues (2011) highlighted methodological issues, while Uddin and colleagues (2013) proposed that inconsistencies were due to the developmental stage at which task-based connectivity was measured. Muller and colleagues (2011) reported that methodological differences could differentiate between studies that supported a general under-connectivity model versus those that did not. Specifically, studies that support under-connectivity were more likely to examine task-related functional connectivity among particular regions of interest, relative to studies that examined whole-brain intrinsic functional connectivity using low-pass filtering. Uddin and colleagues (2013), on the other hand, reviewed age-related differences in functional connectivity in ASD and proposed a developmental model. They argued that there is a shift in brain connectivity patterns following puberty. Hormones that are expressed during puberty have an impact on brain development (Cahill, 2006). Uddin and colleagues reported that studies that examined children with ASD younger than 12 years of age found over-connectivity, while studies of late adolescence and adulthood reported under-connectivity. Further complicating these developmental changes, near-infrared spectroscopy revealed that infants who are at high risk of having ASD (siblings of children diagnosed with ASD), relative to controls, show marginally increased functional connectivity at 3-months of age, which shifts to significantly decreased functional connectivity by 12-months of age (Keehn, et al., 2013).

Considering the dramatic changes in structure, periods of overgrowth, arrested growth, over connectivity, and under connectivity, ASD may be better characterized as a disorder with disrupted growth and connectivity, which would encompass the wide range of findings (Maximo et al., 2014).

4.3. Current Study

The focus of the current study was to compare task-based functional connectivity (whole-brain patterns of covariance) between young adults with and without ASD, in response to synchronous and asynchronous multisensory social-linguistic, nonsocial-nonlinguistic, and social-nonlinguistic stimuli. Patterns of whole-brain task-related covariance were explored using PLS. Given that whole-brain patterns of covariance for multisensory perception have not been previously examined in individuals without ASD, the current study was largely exploratory. It was hypothesized that the group without ASD would display a multisensory specific network that would differentiate synchronous from asynchronous stimuli that would include the pSTS, regardless of content condition. As individuals with ASD were found to display a differential BOLD response to temporally synchronous and asynchronous stimuli compared to young adults without ASD in study two, the current study hypothesized that individuals with ASD would display an altered network for synchronous multisensory integration. Specifically, it was proposed that individuals with ASD would not display a multisensory specific network that would differ as a function of synchronous from asynchronous stimuli, or would display a pattern of activity specific to synchronous multisensory integration that differed from the group without ASD (i.e., a different network), which would not include the pSTS.

4.4. Data Analysis

The preprocessed data from study 2 were analyzed using PLS (McIntosh, 1999; Krishnan et al., 2011). PLS is a multivariate technique that examines patterns of whole-brain covariance and how they differ across groups and/or conditions (McIntosh, Chau, & Protzner, 2004). PLS is similar to principle components analysis, but the critical difference is that the resulting components, or “latent variables” (LVs: also akin to factors in a factor analysis) are specifically attributable to the different conditions, groups, or behavioural metrics being explored. PLS offers many advantages over other statistical techniques; in particular, it is robust to many potential confounds. PLS is ideal for data sets where the dependent measures within a block are highly correlated, as in neuroimaging data (McIntosh et al., 2004). Because of its ability to identify groups of brain regions with covarying activity, this technique is methodologically suited to the investigation of large-scale brain networks. PLS was used to examine differential patterns of whole-brain covariance, across different conditions and stimulus-types, between groups.

Specifically, PLS was used to identify patterns of covariance that differed between young adults with and without ASD. Significance of the LVs was determined with permutation tests (500 permutations), using resampling without replacement. Robustness of each voxel's contribution to a LV across subjects was provided by bootstrap resampling, which resampled the data 100 times, with replacement, to estimate the standard error of the weight of each voxel on the LV. A bootstrap ratio, calculated as the ratio of each voxel's weight to its standard error, was thresholded at ±2.58, equivalent to $p < 0.01$. Unlike standard mass-univariate voxel-wise analysis techniques, PLS calculates values on the whole brain in a single mathematical step, thus, no corrections for multiple comparisons are required. For each participant, a composite brain score was calculated, which provides an index of how strongly each participant expresses the pattern of activity identified by that LV. To examine differences across groups, the mean brain scores for the different conditions within each group from each significant LV were compared. Confidence intervals (95%) were calculated from the bootstrap resampling, and differences between conditions and groups were determined by a lack of overlap in these confidence intervals.

PLS analyses were conducted on the same sample of participants described in study 2 (see study 2 for additional information regarding methods). Further, since the particular focus of the current study was to examine multisensory integration in conditions that varied in the level of social and linguistic content, the task-based PLS analyses examined only the social-linguistic, social-nonlinguistic, and mousetrap conditions (excluding the music condition).

In order to interpret and understand the significance of each PLS analysis, the statistical significance, percent of the covariance accounted for by the model, design scores, and mean brain scores will be considered. Design scores represent the contrast weightings that explain the most task-related variance in BOLD signal. Brain scores represent a composite measure of activity on a given LV for each participant, akin to factor scores, which indicate how strongly individual participants express the patterns on the LV. They are calculated by multiplying the salience by the BOLD signal change (relative to a reference scan) in every voxel and summing these values across the whole brain. Brain scores are then used to examine differences between experimental conditions and groups.

4.5. Results

4.5.1. Group without ASD

As the research on the networks underlying multisensory integration in neurotypical individuals is relatively scarce, it was worthwhile to first examine the whole-brain covariance for synchronous and asynchronous activation across the conditions within the group without ASD alone. A significant pattern of activity dissociating the mousetrap condition from the social-linguistic and the social-nonlinguistic conditions was found, which accounted for 62.93% of the covariance in the model, $p < 0.001$ (figure 19; table 12). The mousetrap condition was associated with regions of the dorsal attention network (Fox, Corbetta, Snyder, Vincent,& Raichle 2006), including the IPS, FEF, middle temporal motion complex (MT+), and lateral occipital regions. Also involved in processing the mousetrap stimuli was the bilateral parahippocampal gyrus, which extended into the entorhinal and perirhinal cortices, left anterior insula, nucleus accumbens, bilateral middle and posterior cingulate cortex, and right thalamus. The social and language conditions were associated with aspects of the “social brain” (for a review see Frith and Frith, 2007; Gotts et al., 2012) and areas that are often associated in language processing, including bilateral STS/STG (including the bilateral auditory cortex, the temporal poles, and right pSTS), and bilateral posterior cingulate cortex. Also included in the social and linguistic conditions was the bilateral medial occipital cortex (cuneus and lingual gyri), mammillary bodies, right substantia nigra, and left thalamus. A second LV that differentiated synchronous from asynchronous stimuli was not significant.

4.5.1.1. Differentiating Social and Linguistic Processes

As previous research has reported differential findings for multisensory integration of linguistic vs. non-linguistic stimuli it was particularly relevant to examine how the social-linguistic stimuli differed from the mousetrap stimuli. Further, previous research has failed to differentiate the social and linguistic aspects of social-linguistic multisensory integration. Stimuli that contain language information, such as a person telling a story, are also social, as they involve attending to another person. Therefore, the current study sought to differentiate social and linguistic multisensory integration through the inclusion of the social-nonlinguistic condition. Therefore, using data from group without ASD only, two independent PLS analyses were performed that included 1) social-linguistic and mousetrap synchronous and asynchronous conditions only and 2) social-linguistic and social-nonlinguistic synchronous and asynchronous conditions only.

4.5.1.1.1. Social-Linguistic vs. Mousetrap. The only significant LV in the social-linguistic and mousetrap analysis differentiated the social-linguistic conditions from the mousetrap conditions, regardless of synchrony (accounting for 74% of the covariance, $p < 0.001$) (figure 20, table 13). Similar to the overall analysis that was run, the mousetrap condition was associated with the dorsal attention network and bilateral parahippocampal regions. The social-linguistic conditions were associated with the regions previously demonstrated to be involved in social and language processing, including the STS/STG and frontal operculum (Broca's area) bilaterally, medial occipital cortex (cuneus and lingual gyri), right substantia nigra, and right amygdala. It is important to note that the activation of the left Broca's area is more prominent in this analysis (figure 20) relative to the overall analysis that differentiated social-linguistic and social-nonlinguistic processing together from the mousetrap stimuli (figure 19). While the second LV differentiated synchronous from asynchronous activation across both content conditions, it was not significant.

4.5.1.1.2. Social-Linguistic vs. Social-Nonlinguistic. The only significant LV in the social-linguistic and social-nonlinguistic analysis differentiated the content conditions regardless of synchrony (accounting for 43.32 % of the covariance, $p = .018$) (figure 21, table 14). The social-nonlinguistic condition included many areas that are implicated in social processing and direction of attention, including the left posterior middle temporal gyrus, the SC (the SC was not included in cluster report as it was under the cluster threshold of 20 voxels, and is not visible in the cortical surface projections, see figure 22 for the SC cluster), bilateral middle cingulate cortex, and the mammillary bodies. The social-linguistic condition involved regions frequently associated with language processing, including the bilateral STS, Broca's area, the posterior-most aspect of the superior temporal gyrus (planum temporale: i.e., Wernicke's area), and primary somatosensory and motor areas. Consistent with hemispheric asymmetry of language processing, activation for the linguistic stimuli was greater in the left hemisphere. Also included in the social-linguistic network was the hippocampus. Similar to the analyses with the social-linguistic and mousetrap conditions, the second LV differentiated synchronous from asynchronous activation across both content conditions, but it was not significant.

4.5.2. Comparison between ASD and Control

A single PLS analysis including both groups was conducted, which included the synchronous and asynchronous conditions for the social-linguistic, social-nonlinguistic, and mousetrap stimuli. Similar to the group without ASD analyses above, the first LV differentiated the mousetrap conditions from the social-nonlinguistic and social-linguistic conditions, regardless of synchrony, and this pattern was not different between the two groups of participants (accounting for 49.27% of the covariance in the model, $p < 0.001$) (figure 23, table 15). The social-linguistic and social-nonlinguistic content conditions involved the social network, including the bilateral STS/STG (stretching from the temporal poles to the pSTS), right amygdala, right substantia nigra, posterior and middle cingulate gyrus, left medial occipital cortex (cuneus and lingual gyri), and the SC (the SC was not included in cluster report as it was under the cluster threshold of 20 voxels, and is not visible in the cortical surface projections, see figure 24 for the SC cluster). The mousetrap condition was associated with the dorsal attention network (IPS, FEF, MT+, and lateral occipital regions), as well as frontal control regions (rostrolateral prefrontal cortex, middle frontal gyrus, frontal operculum, cingulate cortex), parahippocampal gyrus, entorhinal and perirhinal cortices. Also coactivated were regions involved in stimulus learning and anticipatory approach behaviours, including bilateral thalamus and the left insula (Kirsch et al., 2003).

The second LV differentiated different conditions for each group, and synchrony was found to play an important role in these different patterns (accounting for 14.35% of the covariance, $p < 0.001$) (figure 25, table 16). For the group without ASD, the LV differentiated the synchronous social-linguistic condition from the asynchronous mousetrap condition, the 2 most disparate conditions. For the group with ASD, the LV differentiated synchronous mousetrap and social-linguistic stimuli from all other conditions. Therefore, as can be seen in figure 25, the group with ASD did not differentiate between processing of synchronous social-linguistic and mousetrap stimuli. Further, the pattern of activation associated with the mousetrap condition in the group with ASD was the same as for social-linguistic processing in both groups, and notably, included left lateralized pSTS activation. A distinct left pSTS activation, similar to the location of the study 2 group findings, was also found. Further, important social and language regions were also involved, including right amygdala, bilateral Broca's area, Wernicke's area, bilateral thalamus, areas along the posterior, middle, and anterior cingulate gyrus, medial prefrontal cortex, and the motor and somatosensory cortices. In contrast, the pattern associated with asynchronous mousetrap processing in the group without ASD and both asynchronous social-linguistic and mousetrap processing in the group with ASD included two main areas: the left posterior middle temporal sulcus and a small cluster in the right STG (this cluster was not visible in the cortical surface projections, see figure 26).

The third LV accounted for 9.21% of the covariance, *p* = 0.01. Similar to the second LV, it differentiated different conditions for each group (figure 27, table17). In the group without ASD, it differentiated both synchronous and asynchronous social-nonlinguistic stimuli from the asynchronous mousetrap condition. In the group with ASD, the LV differentiated the synchronous and asynchronous mousetrap conditions from the synchronous and asynchronous social-linguistic conditions. These conditions were differentiated such that the group with ASD was using the same regions to process nonsocial-nonlinguistic information as the group without ASD used to processes social-nonlinguistic information. This included the mammillary bodies and posterior cingulate cortex (these clusters were not visible in the cortical surface projections). Further the group with ASD was also using the same regions to process social-linguistic stimuli, regardless of synchrony, as the group without ASD used to process asynchronous mousetrap stimuli. This included bilateral middle STG (auditory cortex) and bilateral pSTS (with greater involvement from the left), the central sulcus (with both motor and somatosensory cortex involvement), anterior, middle, and posterior cingulate cortex, left fusiform gyrus, bilateral insula, bilateral precuneus, right parietal-occipital sulcus, and right superior parietal cortex. This pattern of activation seems to be consistent with the ventral attention network (Fox et al., 2006).

4.6. Discussion

The goal of the current study was to examine the whole-brain patterns of covariance in young adults with ASD compared to young adults without ASD, in order to examine if the atypical multisensory perception of individuals with ASD identified in previous research was related to aberrant neural connectivity. The networks involved in processing social and linguistic multisensory information relative to nonsocial-nonlinguistic (mousetrap) multisensory information were of particular relevance.

No one network for processing all forms of synchronous multisensory information was identified. However, the results of this study revealed many important findings to be discussed. Prior to discussing group differences, an examination of the networks that were involved in processing the linguistic, social, and nonsocial-nonlinguistic components of each stimulus set is needed.

4.6.1. Nonsocial-Nonlinguistic Processing

Analyses involving both the groups with and without ASD consistently implicated the dorsal attention network in the processing of the nonsocial-nonlinguistic stimuli. The dorsal attention network is often activated during visual search tasks, and is thought to reflect top-down attentional control (i.e. voluntary attentional control; Corbetta & Shulman, 2002; Ozaki, 2011). While the dorsal attention network has been primarily implicated in spatial attention, it has also been reported to be involved in other types of top-down visual attentional processes (Corbetta & Shulman, 2002). Consistent with this, the mousetrap condition most likely engaged top-down cognitive processing related to the task of visually searching for and tracking the marble as it travelled through the run.

Given that the dorsal attention network was strongly involved in the processing of the mousetrap stimuli, it is also not surprising that some regions of the frontoparietal control network (Vincent et al., 2008) also appeared to be involved. Although the parietal regions of the frontoparietal control network did not emerge in the analyses, it is important to note that this lack of parietal association does not indicate that this entire network was not involved in the processing of the stimuli. Rather, the parietal region may have been involved in both the social-linguistic and nonsocial-nonlinguistic conditions, and therefore did not emerge as dissociating between the stimuli. The frontoparietal control network has been proposed to regulate switching between the dorsal-attention network, which is responsible for externally focused cognition, and the default network, which is responsible for internally focused cognition (Spreng, Stevens, Chamberlain, Gilmore, & Schacter, 2010). It has been suggested that, from an anatomical perspective, the frontoparietal control network is positioned in a way that would facilitate the integration of information from both the dorsal attention and default networks (Vincent et al., 2008). Therefore, the frontoparietal control network is engaged by tasks that demand top-down cognitive control. In particular, the areas of activation in the frontoparietal control network (e.g. insula and the anterior cingulate cortex) are proposed to be involved in focal attention and initiating top-down regulation (Petersen and Posner, 2012).

Additionally, the nucleus accumbens, left insula, middle and posterior cingulate, and thalamus were also involved in processing the mousetrap stimuli. The interaction between the anterior insula with the thalamus and nucleus accumbens is often associated with reward driven learning (Cho et al., 2013; Kirsch et al., 2003; Thut et al., 1997; Elliott, Friston, & Dolan, 2000; Knutson, Adams, Fong, & Hommer, 2001). It appears that there is something unique about the nonsocial-nonlinguistic stimuli that was driving this reward learning circuitry. At the end of the mousetrap clip, the marble completed the run by setting off a trap. It is possible that the repetitive nature of the marble stimuli, coupled with this end result of a trap being set off, was unique enough to this stimuli set (relative to the social-linguistic and social-nonlinguistic stimuli) that it resulted in these reward circuits being involved.

The parahippocampal gyrus, entorhinal cortex, and perirhinal cortex were also consistently implicated in the processing of the nonsocial-nonlinguistic stimuli. It is not surprising that these regions were involved in the processing of multisensory stimuli (regardless of synchrony) given that they receive inputs from both the visual cortex and the auditory cortex (Aminoff, Kveraga, & Bar, 2013; Murray & Bussey, 1999). In particular, with regards to the ventral-visual stream, the perirhinal cortex is the first cortical area where there is integration of information from different sensory modalities (Murray & Bussey, 1999). Further, given the novel nature of the mousetrap stimuli, coupled with the unique task of visually tracking the marble while it descended through the run and learning the sounds that each part of the run made, it is intuitive that these areas would be involved in the cognitive processing of this stimulus. Together, the perirhinal, entorhinal, and parahippocampal areas are involved in visual stimulus recognition, learning and retention of visual-auditory associations, visual-spatial processing, visual object identification, and contextual associations (Higuchi & Miyashita,1996; Insausti, Amaral, & Cowan, 1987; Murray & Bussey, 1999).

Finally, consistent with previous research with other nonsocial-nonlinguistic stimuli, a small area of activation in the left middle temporal gyrus was also unique to the mousetrap stimuli. This area has been found to respond preferentially to tool-based stimuli (relative to scenes, animals, etc.; Chao, Haxby, Martin, 1999; Beauchamp et al.,2002). Additionally, this area has been found to display even greater activation for moving objects relative to stable objects (Beauchamp, et al., 2002).

Taken together, it appears that networks critical for identifying and remembering nonsocial visual objects and their spatial location worked in concert with top-down cognitive control processes in order to process the nonsocial-nonlinguistic stimuli in both groups.

4.6.2. Linguistic and Social Processing

The social-linguistic and social-nonlinguistic stimuli engaged many of the same regions, which suggested common underlying social processing, regardless of linguistic content. Consistent with this interpretation, areas of the “social brain” (Frith, 2007; Frith & Frith, 2007) were strongly involved in the processing of these social stimuli. The social brain is critical for making sense of social interactions, understanding others’ intentions, predicting what is going to happen next, and planning responses in social exchanges (Frith, 2007; Frith & Frith, 2007). The social brain is composed of connections between cortical and sub-cortical circuits (Gotts et al., 2012). The amygdala, temporal poles, pSTS (which will be discussed in further detail below), lingual gyrus, and posterior cingulate, were consistently involved in processing of the social stimuli in the current study and have all been well-established as important areas for social-cognitive processing. The amygdala has long been associated with social processing. Interestingly, amygdala volume in humans has been found to correspond to variations in social network complexity and size, such that individuals with larger social groups have greater amygdala volumes (Bickart, Wright, Dautoff, Dickerson, & Barrett 2011). The amygdala has been proposed to have a role in associating value (positive and negative judgments) and recognizing the emotional states of others (Frith, 2007; Morris et al, 1996; Pelphry, et al., 2011). The temporal poles seem to be more involved in interpreting changes in others' behaviour in reaction to different situational demands (Funnell, 2001). Thus, the temporal poles play a vital role in mentalizing and applying one's general knowledge of social norms to the particular situation at hand on a moment-to-moment basis (Ganis & Kutas 2003; Frith, 2007). The lingual gyrus was also consistently implicated in the processing of the social stimuli. Previous research has also included the lingual gyrus in the social brain as it has been frequently implicated in face processing (Gobbini, Leibenluft, Santiago, & Haxby, 2004; Puce, Allison, Gore, & McCarthy, 1995).

The thalamus, SC, substantia nigra, posterior cingulate, mammillary bodies, and the cuneus were also involved in processing social stimuli in the current study. These areas have been found to have important connections amongst themselves and with the previously mentioned social brain cortical areas, and are critically involved in integrating sensory stimuli, directing attention, self-focused cognition, and memory retrieval (Cabeza & St. Jacques, 2007; Cavanna & Trimble, 2006; Frischen, Bayliss, & Tipper, 2007; Hoffman, Gothard, Schmid, & Logothetis, 2007; Johnson et al., 2005; Maddock, Garrett, & Buonocore, 2001, 2003; Tekin, & Cummings, 2002; Van Der Werf, Jolles, Witter, & Uylings, 2003; Vuilleumier et al., 2002). In particular, the connections between the STS, amygdala, motor cortex, and basal ganglia have been proposed to be a network that is involved in moving from perception to action (Alisson, et al, 2000; Pelphry, Shultz, Hudac, & Vander Wyk, 2011).

4.6.3. Linguistic Specific Processing

Naturalistic linguistic stimuli are inherently social as they involve attending to someone else talking. Therefore, it was not surprising that the social-linguistic and social-nonlinguistic stimuli in the current study both involved many of the same social brain regions. However, through the direct comparison of the regions involved in processing both the social-linguistic and social-nonlinguistic stimuli, some regions that preferentially responded to the linguistic information were identified. In particular, the social-linguistic condition involved greater activation bilaterally of the STS, hippocampus, Broca’s area, Wernicke’s area, and the somatosensory and motor strips. Consistent with hemispheric asymmetry of language processing, activation for the linguistic stimuli was greater in the left hemisphere. Activation of the STS, hippocampus, Broca’s area, and Wernicke’s area is also consistent with previous research that has examined language comprehension and verbal memory (Dronkers, 2011; Dronkers, Wilkins, Van Valin, Redfern,& Jaeger, 2004; Friederici, 2002; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Jung-Beeman, 2005; Smith & Milner, 1989).

Interestingly, the motor and sensory cortices were also involved in processing of the social-linguistic stimuli. In line with theories regarding mirror neurons, these areas have been found to be involved in social brain circuitry that is activated when humans observe others moving or being touched (Blakemore et al. 2005; Chartrand & Bargh 1999; Keysers et al. 2004; Frith, 2007). Further, the motor and sensory cortices have been implicated in processing and understanding the actions of others (Hamilton & Grafton, 2008). However, in the current study, these areas were uniquely involved in processing stimuli of a woman telling a story, but not a woman making non-verbal oral sounds/movements. The story did include some action and sensory words (e.g. pat, hot), and researchers have found that the motor and premotor cortices are activated during language comprehension tasks that include action words (Rapso, Moss, Stamatakis, & Tyler, 2009). Therefore it is possible that the action words were driving activation of these areas.

4.6.4. Group Differences

4.6.4.1. Synchronous Multisensory Linguistic Network.

When examining the PLS results in order to identify networks that may be responsible for uniquely processing synchronous audiovisual information there were important differences found between the groups. One pattern of brain activation emerged that differentiated the synchronous social-linguistic stimuli from the asynchronous nonsocial-nonlinguistic stimuli, only in the group without ASD. These two conditions (social-linguistic and nonsocial-nonlinguistic stimuli) were the two most disparate conditions in terms of structure and content. There was also a trend towards a dissociation between the synchronous and asynchronous social-linguistic stimuli in this group. It is possible that there were not enough trials to fully differentiate between synchronous and asynchronous social-linguistic stimuli in the group without ASD, and future research should explore this relationship further.

In contrast, the group with ASD processed synchronous social-linguistic and nonsocial-nonlinguistic stimuli with the same network, and differentiated these two conditions from all other stimulus types. Thus, unlike the group without ASD, the group with ASD did not display specificity for synchronous multisensory linguistic processing, but rather, engaged the same network for processing both linguistic and nonsocial-nonlinguistic synchronous multisensory stimuli. Notably, this pattern of activation included a region in the left pSTS overlapping/ proximal to the area that displayed underlying group differences in study two. Importantly, the brain regions underlying the synchronous social-linguistic and nonsocial-nonlinguistic processing also included regions involved in social and language processing discussed earlier (bilateral amygdala, Broca's area, bilateral thalamus, posterior cingulate), as well as additional areas also implicated in social and language processing (anterior cingulate cortex, medial prefrontal cortex, and ventrolateral frontal cortex).

In previous behavioural research, a linguistic-specific deficit for multisensory perception in individuals with ASD has been proposed (Bryson, 1972). Much of the support for this linguistic-specific deficit has come from studies employing preferential looking paradigms. When viewing synchronous and asynchronous information, individuals without ASD displayed a "linguistic boost" (i.e., displayed a preference for the synchronous quadrant in a preferential looking display) relative to individuals with ASD who did not display this boost (Lavoie et al., in preparation). In other studies, individuals without ASD did not display this linguistic boost, however, participants with ASD showed a "linguistic drop" for multisensory information (i.e., children without ASD displayed preferential looking to the temporally synchronous audiovisual screen in linguistic and nonlinguistic conditions, while children with ASD only displayed preferential looking to the temporally synchronous audiovisual screen in the nonlinguistic condition; Bebko et al., 2006). Notwithstanding this discrepancy, both studies provided evidence that individuals with ASD do not display the same linguistic-specific multisensory processing as individuals without ASD. However, in the current study, it appeared that individuals with ASD were using the same network to process the synchronous audiovisual linguistic information and synchronous audiovisual nonsocial-nonlinguistic information. Therefore, the group with ASD did not show a linguistic-specific network, which is consistent with the behavioural findings. The absence of linguistic-specific processing of multisensory information may interact with other attention regulation and control networks in order to explain the aberrant behavioural responses of individuals with ASD in preferential looking paradigms. Further research using preferential looking paradigms should explore how attentional control networks interact with the regions that process synchronous multisensory linguistic stimuli in neurotypical individuals.

As indicated previously in the current study, the areas involved in processing the synchronous social-linguistic stimuli for the group without ASD and synchronous social-linguistic and nonsocial-nonlinguistic stimuli in the group with ASD are recognized as being components of important social and language networks. Prominently involved in this network was Broca’s area (bilaterally, with a left side predominance), Wernicke’s area, bilateral STS, right amygdala, and many segments of the cingulate cortex (anterior, middle, and posterior). The medial prefrontal cortex and the anterior cingulate cortex were also included in this network, although they were not part of the pattern that dissociated social from nonsocial stimuli per se in a previous analysis. With respect to social processing, the medial prefrontal cortex is involved in emotional processing, perceptions and judgments of others, and theory of mind (see Amodio & Frith, 2006 for a review). The anterior cingulate cortex has been associated with a diverse range of functions, including social feedback, emotional learning, and assigning emotional valence to stimuli (Craig, 2008; Devinsky, Morrell,& Vogt, 1995; Somerville, Heatherton, Kelley, 2006). Together, the anterior cingulate cortex and the medial prefrontal cortex have also been implicated in joint attention (i.e. the shared attentional focus of two individuals), which is known to be impaired in individuals with ASD (Mundy, 1995).

4.6.4.2. Altered Networks for Social and Linguistic Processing in ASD.

While the second LV suggested a lack of language-specific processing in individuals with ASD, the third LV indicated even more aberrant networks involved in processing language and social information in the group with ASD. Specifically, it appeared that the underlying network that the group without ASD was using to process the asynchronous nonsocial-nonlinguistic information was the same that the group with ASD engaged to process language information, regardless of synchrony. In addition to the areas discussed previously (bilateral middle STS and bilateral pSTS, with greater involvement of the left, the left central sulcus, posterior and middle cingulate cortex), the ventral attention network was involved in processing the stimuli.

One of the key roles of the ventral attention network is to direct cognitive attention to unexpected or low-frequency events (for a review see Corbetta & Shulman, 2002). Therefore it is not surprising that this network was activated for the group without ASD when they viewed a highly novel class of stimuli where the audio and visual content was asynchronous. What is concerning about the group with ASD is that ventral attention system was activated for social-linguistic information, regardless of synchrony. This implies that the group with ASD was engaging a network involved in processing unexpected stimuli when presented with social-linguistic information, which should be relatively common in their everyday environment.

Additionally, the pattern of brain activity identified by the third LV indicated that the group with ASD processed the nonsocial-nonlinguistic stimuli, regardless of synchrony, similar to the way the group without ASD processed social-nonlinguistic information. In particular, the right amygdala, posterior cingulate, and the mammillary bodies were involved. As discussed previously, these regions have all been connected with the social brain. Thus, this finding suggests that the group with ASD engaged the social-brain to process the nonsocial stimuli, but not the social stimuli. Further discussion about the social brain in ASD is provided below.

4.6.5. Social Brain in ASD and Multisensory Integration

Regions implicated in the social brain were of particular relevance in the current study, given that multisensory processing deficits have been proposed to be more evident for social-linguistic stimuli. The social brain contains many of the same underlying regions involved in the processing of language information. Given that language information is inherently social and that language development is tied to social interactions (for a review see Kuhl, 2007), this overlap makes sense. The results of this study indicate that individuals with ASD were using social and language based networks to process nonsocial-nonlinguistic stimuli, suggesting that while these networks are social and/or social-linguistic specific for the group without ASD, these networks appear not to be specialized in individuals with ASD.

In individuals with ASD, many areas of the social brain discussed above display atypical functioning, including the amygdala (Di Martino et al., 2009; Hadjikhani et al., 2007; Kleinhans, et al., 2010; Pinkham et al., 2008), pSTS (Castelli, et al., 2002; Hadjikhani, et al., 2007; Pinkham, et al., 2008), somatosensory and premotor cortex (Hadjikhani, et al., 2007), anterior cingulate cortex (Thakkar, et al., 2008), posterior cingulate cortex (Di Martino, et al., 2009), and superior colliculi (Kleinhans et al., 2010). Given that many areas of the social brain seem to be functioning atypically in ASD, a more parsimonious explanation is that it is the interactions among these regions within the social brain network that are atypical, rather than each region being impaired in isolation. More recently, researchers have been examining the underlying network of connections across the social brain. By examining the brain at the network-level, abnormal connectivity between these regions has been found in ASD. In particular, a lack of functional connectivity between the limbic system and cortical areas involved in social processing has been found (Gotts et al., 2012). Broadly, the limbic circuitry is involved in the processing of emotional aspects of social interactions and the cortical areas are involved in higher-order social processing (e.g., making complex social interpretations of others' actions, and planning responses to social interactions) (Gotts et al., 2012). Further, using a large multi-site database of resting-state functional connectivity, Di Martino et al. (2014) found hyper-connectivity between the subcortical regions and hypo-connectivity between cortical regions in ASD.

It is likely that these networks develop atypically in individuals with ASD due to early brain developmental abnormalities. Perlman and colleagues (2010) proposed that one or more of the specialized functions of neuroanatomical structures in the social brain develop atypically. Further, because this atypical development happens prenatally or early in infancy, the downstream consequence is that the development of other brain regions in the social network, and the connections between them, are impacted. The current study expands upon previous findings of aberrant connectivity in the social brain to further demonstrate that it is not functionally specialized for processing social information in individuals with ASD. From a developmental perspective, it would make sense that there would be impairment in both social and language processing networks for individuals with ASD, given that these networks overlap and their development is so interdependent.

Particularly interesting in the current study, is that this lack of social and language specific networks for the group with ASD emerged in the context of brain patterns that also differentiated some synchronous and asynchronous stimuli. In the everyday environment, there is a multitude of information and the brain must automatically select what to attend to and process from one moment to the next. Both social information (Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998) and synchronous multisensory stimuli (Bahrick & Lickliter, 2002; Bahrick & Todd, 2012) direct attention and are preferentially attended to in the environment. Therefore, the combination of both synchronous and asynchronous multisensory linguistic and non-linguistic information in the current study helped to illuminate potential differences between individuals with and without ASD in social and language processing.

4.6.6. Multisensory Integration and the pSTS

The pSTS in particular is an area that is thought to be critically involved in multisensory integration. However, this region has also been implicated in many social-cognitive processes. Specifically, parts of the pSTS have been implicated in processing biological movement and understanding the intentions of these movements (Pelphrey et al, 2004; Pelphrey et al, 2005; Saxe et al., 2004). The finding of the pSTS differentiating the social stimuli from the nonsocial stimuli would appear to fit with this hypothesis. However, in the second LV, the left pSTS also seemed to be associated with synchronous information specifically (social-linguistic in the group without ASD, and both social-linguistic and nonsocial-nonlinguistic in the group with ASD). This pSTS involvement in multisensory integration suggests that this region may not only be specialized for viewing and interpreting biological motion, but may have a function in multiple cognitive processes. Although it is possible that different subsections of the pSTS are involved in distinct processes (e.g. distinct areas for biological motion vs. multisensory integration), the pSTS might instead have different functions depending on the network of brain regions with which it interacts when processing specific types of stimuli, consistent with the ideas of neural context and response plasticity (McIntosh, 2004).

It is possible that multisensory neurons are found throughout important sensory hubs in the brain, including the pSTS, SC, and the thalamus. These hubs might participate in many important functional networks, including the social brain. The role of the multisensory neurons in these hubs may be to assist with more rapidly processing and preferentially attending to relevant stimuli. Further, these sensory relay or integration sites (thalamus, SC, STS), which are known to be involved in processing multiple modalities of information and directing attention, appear to overlap with the social brain network. Given that humans are a socially dependent species, one of the most important sources of information is socially relevant stimuli. As such, this overlap between sensory, attention, and social networks may result in a greater concentration of multisensory neurons in hubs that are within the social networks. This overlap would allow for the quick recognition of social stimuli and speeded facilitation of social interactions. Further, given that it has been proposed that individuals with ASD have impairments in selectively orienting and shifting their attention to social stimuli (Dawson et al., 1998), this hypothesis of overlap in multisensory and social networks makes sense given the current findings of a lack of social-specificity in the ASD social brain network.

4.6.7. Conclusions

While the pSTS has been heavily implicated as an area responsible for multisensory integration, it appears that this region alone does not account for differences in processing temporally synchronous and asynchronous stimuli. Importantly, this study demonstrated that in addition to the aberrant connectivity of the social brain network, individuals with ASD display a lack of functional specificity for processing of social and language information in their putative social and language networks. This lack of specificity for social and language information becomes clear when comparing synchronous to asynchronous stimuli in particular. Future research should further explore how multisensory information may enhance social and language information processing and how multisensory neurons are distributed within the human brain.

Final Discussion

The overall goal of the current study was to explore the neurological mechanisms that underlie multisensory perception deficits in individuals with ASD. These deficits were proposed to be a core underlying symptom of individuals with ASD that would explain downstream consequences in social, language, and behavioural functioning. Further, they have been suggested to be most evident for linguistic information.

The collection of studies sought to elucidate the nature of these multisensory integration deficits by delineating social and linguistic processing and examining differences in neural activity across populations with and without ASD. Because temporal synchrony is thought to be the basis for more complex forms of multisensory integration, the current study focused on multisensory processing abnormalities with temporal synchrony variations. The overall goal was accomplished in three ways: 1) by demonstrating the efficacy of the temporal synchrony method in examining neurological mechanisms of multisensory integration; 2) by examining the role of the pSTS in multisensory integration of individuals with and without ASD; and 3) by examining how the networks responsible for processing social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic multisensory stimuli differed in young adults with and without ASD.

5.1. Review of Study Findings

5.1.1.Study One: Temporal Synchrony

In order to explore multisensory integration in individuals with ASD, it was first necessary to ensure that the method being used to identify multisensory areas of the brain was both theoretically valid and empirically practical. Thus, the temporal synchrony method was compared to the three most commonly used methods for the identification of multisensory regions: the mean, max, and super-additive methods. Overall, the results of study one demonstrated the efficacy of the temporal synchrony method by demonstrating that it is theoretically valid and is the most empirically tenable of the methods examined.

In comparison to the only other method in the current study that is theoretically valid - the super-additive method - the temporal synchrony method was demonstrated to be more empirically practical. Due to the heterogeneity of cells at the voxel-level, super-additivity (which requires the synchronous audiovisual response to be greater than the sum of the unisensory responses) is unlikely to be observed, as the co-localization of non-multisensory neurons within the voxel would wash out the signal associated with the unique response properties of the minority multisensory neurons. Given the low multisensory signal to noise ratio in a heterogeneous population of cells in a voxel, many trials would be required to yield sufficient power to detect a super-additive response. The temporal synchrony method is not similarly impacted by the heterogeneity of cells at the voxel level, as the non-multisensory neurons in the voxels would respond equivalently to all conditions. Thus, by subtracting the asynchronous condition from the synchronous condition, the contribution of the non-multisensory neurons to voxel activation (i.e., the "noise") is removed. Therefore, the temporal synchrony method should require fewer trials in order to be able to identify regions demonstrating multisensory integration at the individual level. Consistent with this, we found that the temporal synchrony method identified a multisensory region in the pSTS in the largest number of participants bilaterally and in at least one hemisphere. Further, the temporal synchrony method was supported in three out of the four conditions, as indicated by the synchronous audiovisual activation being significantly greater than the asynchronous audiovisual activation in the social-linguistic, social-nonlinguistic, and nonsocial-nonlinguistic mousetrap conditions.

5.1.2. Study Two: Comparison Between Groups using ROI Analyses

Study two used the temporal synchrony method to identify ROIs in the pSTS that display temporal synchrony enhancement. Targeted contrasts were used to compare the BOLD response of individuals with and without ASD in ROIs identified at the individual- and group-level.

Individuals with and without ASD demonstrated a differential response at the group-level using whole-brain analyses. There were many regions where the group without ASD showed greater temporal synchrony enhancement (synchronous multisensory stimuli > asynchronous multisensory stimuli) relative to individuals with ASD. Through the examination of an ROI in the pSTS, a region thought to be critical for multisensory integration, individuals with ASD displayed either a lack of multisensory enhancement to synchronous multisensory stimuli or a reversed response (asynchronous > synchronous). This atypical multisensory response in the pSTS for individuals with ASD was apparent for both social-linguistic and nonsocial-nonlinguistic stimuli.

5.1.3. Study Three: Comparison of Whole-Brain Activity Patterns Between Groups

Study three compared whole-brain patterns of activity associated with observing synchronous and asynchronous multisensory stimuli between young adults with and without ASD, in response to social-linguistic, nonsocial-nonlinguistic, and social-nonlinguistic stimuli. The results of study three suggested that individuals with ASD display a distinct lack of functional specificity for processing of social and language information in their putative social and language networks. This lack of specificity for social and language information became clear when comparing synchronous to asynchronous stimuli in particular.

5.2. Integration of Findings

5.2.1. Temporal Synchrony Method

While study one demonstrated that the temporal synchrony method was empirically practical with a neurotypical population, study two demonstrated that not all populations display temporal synchrony enhancement (synchronous audiovisual > asynchronous audiovisual) in regions thought to be responsible for multisensory integration. In fact, individuals with ASD showed the opposite relative response in some cases (asynchronous audiovisual > synchronous audiovisual). Therefore, in addition to identifying brain regions that show a greater relative response to synchronous stimuli, future studies should also investigate regions that demonstrate the opposite response (i.e., a greater relative response to asynchronous stimuli).

5.2.2. The role of the pSTS in Multisensory Integration

Through the three studies presented, it is clear that the pSTS is not exclusively engaged in multisensory integration. While some propose that the pSTS is an important part of the social brain, specific for understanding and interpreting biological motion (Materna et al., 2008; Pelphrey et al., 2004; Pelphrey et al., 2005; Redcay, 2008; Saxe et al., 2004), the results of the current study were inconsistent with this hypothesis; rather, the pSTS was also involved in multisensory integration (as indicated by the finding in study two that the pSTS demonstrated temporal synchrony enhancement, and in study three that the pSTS was involved in synchronous linguistic multisensory integration in the group without ASD). While further research is needed in order to fully clarify the implications of the current results, the pSTS appears to be variably engaged with dissociable networks of regions involved in integrating multisensory stimuli (both social-linguistic and nonsocial) and processing social stimuli per se. Applying the ideas of neural context and response plasticity (McIntosh 2000), it is proposed that the pSTS may function differently in discrete and/or overlapping brain networks for the processing of synchronous audiovisual social and nonsocial information. McIntosh (2000) proposed that critical nodes in the brain could serve as behavioural catalysts. Due to the anatomical connections that catalysts have with other regions, catalysts facilitate shifting from one network to another. An intriguing possibility, consistent with the results of the current studies, is that the pSTS could serve as one of these important catalyst nodes. This idea is proposed here as the "multisensory catalyst nodes" hypothesis.

5.2.3. Multisensory Catalyst Nodes Hypothesis

In animal research, multisensory neurons have been identified in homologues of both the human pSTS and SC (Calvert et al., 2000; Stein & Meredith, 1993). However, in these structures, only 25-60% of the cells are proposed to be multisensory neurons (Beauchamp et al., 2004; Laurienti et al., 2005). If these structures are not entirely composed of multisensory neurons, it suggests that they also serve purposes beyond multisensory integration. Further, the results of the PLS analyses in study three indicate that there is not one single specialized network devoted exclusively to multisensory integration. A synthesis of previous research and the results of the current study suggest that synchronous multisensory information may be processed by multiple areas of the brain that contain these specific multisensory neurons, and also contain other populations of neurons. These areas can be thought of as important hubs (or catalyst nodes) in the brain that have both: 1) neurons designated for responding to multisensory information, and 2) neurons that are essential components of other networks, such as the social brain.

As multisensory neurons direct attention (Stein et al., 1989), the overlap between these multisensory catalyst nodes and hubs for other brain networks may facilitate the rapid response of these other networks to important stimuli in the environment. For example, in animals, the SC is one of the most well understood structures for multisensory integration. The SC receives input from visual, auditory, vestibular, and proprioceptive areas, and has projections to motor and premotor areas, as well as control over the orienting of the eyes to stimuli in the environment (Wallace & Stein, 1994; Meredith & Stein, 1986). Multisensory neurons also respond more rapidly to multisensory stimuli than they do to unisensory stimuli (Stein & Meredith, 1993). Thus, the overlap of the dense population of multisensory neurons in the SC, combined with its role in networks responsible for orienting humans' attention to the environment, ensures that the important multisensory information in the environment is preferentially attended to over unisensory information. This hypothesis of multisensory catalyst nodes is consistent with research demonstrating that humans provide quicker behavioural responses to multisensory stimuli than to unisensory stimuli (Calvert et al., 2000; Frens & Van Opstal, 1995; Henderson, 1962; Morrell, 1968), and that multisensory stimuli direct selective attention (Bahirck & Lickliter, 2002; Bahrick & Todd, 2012). Following this hypothesis, it is possible that multisensory neurons are found throughout important sensory hubs in the brain, including limbic structures, such as the thalamus and the cingulate cortex, as well as cortical areas, such as the pSTS and occipital cortex. Supporting this idea, the current study found that the pSTS, thalamus, and cingulate cortex were involved in processing the synchronous social-linguistic stimuli.

Given the role of limbic and sub-cortical structures in orienting attention and learning, future research should further explore these regions for the existence of potential multisensory catalyst nodes and the function of these nodes in directing attention towards multisensory information.

The social brain depends on communication between cortical and limbic structures (Gotts et al., 20102). If more of these putative multisensory catalyst nodes were found in limbic structures, it would make sense that they would have substantial overlap with the social brain. Given that humans are a socially dependent species, one of the most important sources of information for humans is socially relevant stimuli. This overlap between sensory, attention, and social networks would essentially allow multisensory catalyst nodes to facilitate the rapid engagement of the social brain in response to relevant sensory stimuli. Therefore, these multisensory catalyst nodes would result in rapid recognition of social stimuli and speeded processing of social interactions. Accordingly, it has been proposed that individuals with ASD have impairments in selectively orienting and shifting their attention to social stimuli (Dawson et al., 1998). The finding of a lack of social-specific network dissociations in individuals with ASD, in conjunction with the hypothesized preponderance of multisensory catalyst nodes within the social brain, is consistent with the fact that individuals with ASD have difficulty shifting and orienting their attention towards social stimuli.

As indicated above, it is proposed that multisensory catalyst nodes are found in regions of the brain that also contain neurons that are essential components of other networks. The proposed role of these multisensory catalyst nodes is to facilitate shifting from one network to another. Thus, these nodes do not form their own global multisensory-specific network, but rather, they overlap and are functionally connected with other networks. This hypothesis of regions that have multisensory catalyst nodes co-localized with other neurons would explain why it might be difficult for studies to identify multisensory specific brain networks. The substantial overlap between multisensory catalyst nodes and other neurons in a specific region would result in a decreased multisensory signal relative to the noise of other neurons. For example, the social brain network would respond to social stimuli regardless of audio-visual synchrony, which is consistent with the first latent variable in the PLS analysis of both groups in study three. As multisensory catalyst nodes are proposed to overlap and be functionally connected with the social brain network, a slightly increased BOLD response to synchronous audiovisual social stimuli relative to asynchronous audiovisual social stimuli would occur. However, due to the heterogeneity of the cells in these regions, the power required to detect this increased signal in response to synchronous audiovisual stimuli may be substantial. Thus, there would need to be a significant increase in the power required to detect specific synchronous multisensory social-nonlinguistic, social-linguistic, and nonsocial-nonlinguistic networks. This hypothesis might explain why the synchronous and asynchronous information was only significantly differentiated between the most opposite stimuli in the neurotypical population (i.e., synchronous social-linguistic stimuli and asynchronous mousetrap stimuli) in study three. Further research with neurotypical populations is needed in order to better understand how multisensory neurons are organized in the human brain and how they may be concentrated within areas that are hubs of other important networks. The current study was able to identify these possible multisensory catalyst nodes that overlap with other networks by varying the temporal synchrony of the stimuli. A potential next step in this research field would be to vary the SNR of the stimuli relative to background noise in order to use the inverse effectiveness principal (Calvert & Thesen, 2004;

Meredith & Stein, 1983; Stevenson & James, 2009), which would make the processing of the stimuli more reliant on multisensory integration.

5.3. Multisensory Perception in ASD

The results from studies two and three indicate that at a neurological level, individuals with ASD are not processing multisensory stimuli in the same manner as neurotypical individuals. While study two indicated that individuals with ASD display a reverse response to temporal synchrony variations (asynchronous > synchronous), this pattern of response was not identified at the network level in study three. Given the proposal that very high statistical power is required in order to detect multisensory-specific networks, it is possible that there was too little power in the current study to detect region-specific differences in the functioning of multisensory hubs in individuals with ASD at the large-scale network-level. Therefore the targeted contrast in study two may have reduced some of the noise from social and language processing per se within respective networks that were present in the overall whole-brain covariance analyses. An important next step will be to use seed-PLS - a method of directly investigating the task-related functional connectivity of particular brain regions - to examine regions that display a differential response to synchronous and asynchronous stimuli in targeted contrasts, such as the pSTS.

Individuals with ASD have been found to integrate even low-level (simple) stimuli over a wider temporal window (Foss-Feig, et al., 2010). If individuals with ASD integrate stimuli atypically at a lower level, it would follow that the integration of high-order, more complex social, linguistic, and nonsocial-nonlinguistic stimuli would also be effected. However, using higher-order stimuli, the results of studies investigating global versus social-linguistic specific deficits are equivocal (Bebko et al., 2006; Bryson, 1972; Mongillo et al., 2008; Willams et al., 2004). Based on the multisensory catalyst nodes hypothesis proposed here, it is possible that regions containing these nodes have developed atypically in individuals with ASD. Aberrant structure and/or function of these multisensory catalyst nodes would cause down-stream consequences in brain development for individuals with ASD. Given the hypothesized significant overlap between multisensory catalyst nodes and the social brain, and the proposed role of the multisensory catalyst nodes in orienting attention, it may be that individuals with ASD do not preferentially attend to social stimuli, and therefore do not develop typical social-specific networks. Further, as language development is so reliant on multisensory perception (Bahrick, 2010; Barhrik & Lickliter, 2002; Bahrick & Todd, 2012; Edelman, 1992; Gibson, 1969; Gogate & Bahrick, 1998; Lewkowicz, 2000; Thelen & Smith, 1994) and closely intertwined with social development (Kuhl, 2007), the language networks would also be impacted. Therefore, the addition of a linguistic component to social stimuli, making them more complex, might further exacerbate the difficulties that individuals with ASD have in preferentially attending to synchronous social-linguistic stimuli (as demonstrated in Bebko et al., 2006 and Lavoie et al., in preparation). It would be useful to explore the function of multisensory catalyst nodes in orienting attention to synchronous audio-visual displays in preferential looking paradigms.

5.4. Conclusions and Future Directions

Iarocci and McDonald (2006) proposed that the perceptual atypicalities observed in individuals with ASD “may arise from the integration of specific processes rather than solely from impairments in the different components” (p.85). Consistent with this proposal, the results of the current studies provide clear evidence that at a neurological level, individuals with ASD do not process all multisensory stimuli in the same way as neurotypical individuals. These findings are particularly relevant in light of the fact that this study used temporal synchrony.

The use of amodal properties for the integration of sensory experiences develops in a hierarchical fashion, with infants developing the ability to integrate based on increasingly complex forms of amodal information. Temporal synchrony is thought to be the foundation for the development of multisensory perception (Lekowicz, 1999; 2000; Radeau, 1994; Stein & Meredith, 1993; Welch & Warren, 1980). In support of this, infants have been found to lack the ability to integrate audiovisual information based on duration and rate if the development of integration based on temporal synchrony was not formed (Lewkowicz, 2000). The current study demonstrated that individuals with ASD are not using the foundational amodal property for multisensory integration in the same fashion as typically developing individuals. If individuals with ASD are not using amodal properties such as temporal synchrony to differentiate and integrate sensory experiences (especially those that are social-linguistic in nature), there would be profound consequences for many aspects of cognitive, language, and social development and functioning. Social interactions and communication in particular are highly dependent on multisensory processing, and thus, it is not surprising that difficulties with multisensory processing have been found to correlate with ASD symptomatology (Donohue et al., 2012). It is possible that atypicalities in multisensory integration at the neurological level are at the core of many symptoms that are characteristic of ASD. This idea also implies that focusing on the integration of sensory experiences through conscious direction of attention or through techniques such as biofeedback could be important avenues for intervention research. Future research should also examine how individuals with ASD integrate sensory experiences with more complex forms of amodal information to determine if the integration of sensory experiences is altered at a neurological level for all forms of amodal information.

The pSTS in particular is an area that is thought to be critically involved in multisensory integration, and this area was implicated as a region that differentiated temporal synchrony integration in individuals with ASD from those without ASD. In study two, individuals with ASD failed to display temporal synchrony enhancement (synchronous > asynchronous) in the pSTS for both social-linguistic and nonsocial-nonlinguistic stimuli. In study three, through the examination of whole-brain patterns of covariance, pSTS involvement was not limited to multisensory integration. However, the pSTS was found to co-activate with other regions to process synchronous social-linguistic information in the group without ASD. However, the group with ASD did not display specificity for synchronous multisensory linguistic processing. This lack of specificity of synchronous multisensory linguistic processing only became clear when examining the PLS analysis that differentiated temporal synchrony variations. Therefore the examination of multisensory information illuminated the underlying atypical neurological functioning of individuals with ASD. A hypothesis of multisensory catalyst nodes was proposed in order to explain the findings from the current studies and connect them to previous behavioural literature.

Overall, given the impact of multisensory perception on social, language, and cognitive development and functioning, it would be important to further understand the nature of these multisensory integration deficits, both from a developmental and neurological level. In particular, research should examine how multisensory catalyst nodes may overlap with other networks in typically developing populations, and then examine how these may develop and function atypically in individuals with ASD.

Finally, ASD, as a diagnosis, encapsulates a wide spectrum of differing abilities and impairments across multiple domains, including social-cognitive processing abilities. It has been proposed that there may be sub-groups of differing social, cognitive, sensory, adaptive, and language profiles in ASD (Joseph, Tager-Flusberg, & Lord, 2002; Kjelgaard & Tager-Flusberg, 2001; Lane, Young, Baker, & Angley, 2009; Tager-Flusber & Joseph, 2003). Exploring how sub-groups of individuals with ASD, particularly those with and without language and intellectual impairments, might differ with respect to the development and functioning of multisensory catalyst nodes will be important.

References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual review of psychology*, *60*, 693. doi: 10.1146/annurev.psych.60.110707.163514

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of cognitive neuroscience*, *13*(2), 232-240. doi: 10.1162/089892901564289

Akshoomoff, N., Lord, C., Lincoln, A. J., Courchesne, R. Y., Carper, R. A., Townsend, J., & Courchesne, E. (2004). Outcome classification of preschool children with autism spectrum disorders using MRI brain measures. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(3), 349-357. doi: 10.1097/00004583-200403000-00018

Akshoomoff, N., Pierce, K., & Courchesne, E. (2002). The neurobiological basis of autism from a developmental perspective. *Development and psychopathology*, *14*(3), 613-634. doi: 10.1017/S0954579402003115

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in cognitive sciences*, *4*(7), 267-278. doi:10.1016/S1364-6613(00)01501-1

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders,* Fifth Edition. Washing, DC: American Psychiatric Association.

Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S., Naumer, M.J. (2005). Functional imaging of human crossmodal indentification and object recognition. Exp Brain Res 166: 559–571. doi: 10.1007/s00221-005-2396-5

Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in cognitive sciences*, *17*(8), 379-390. doi:10.1016/j.tics.2013.06.009

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268-277. doi:10.1038/nrn1884

Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., & Bullmore, E. T. (2007). Differential activation of the amygdala and the 'social brain'during fearful face-processing in Asperger Syndrome. *Neuropsychologia*, *45*(1), 2-14. doi:10.1016/ j.neuropsychologia.2006.04.014

Bahrick, L. E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infants Behavior and Development, 10*(4), 287-416. doi: 10.1016/0163-6383(87)90039-7

Bahrick, L.E. (1988). Intermodal Learning in infancy: Leraning on the basis of two kinds or invariance relations in audible and visible events. *Child Development, 59*(1), 197-209. DOI: 10.2307/1130402

Bahrick, L. E. (1994). The development of infants' sensitivity to arbitrary intermodal relations. *Ecological Psychology, 6*(2), 111-123. doi:10.1207/s15326969eco0602_2

Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. *Blackwell handbook of infant development*, *1*, 120-165. doi: 10.1002/9781444327564

Bahrick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimuliation. Journal of Experimental Child Psychology, 79(3), 253-270. doi:10.1006/jecp.2000.2588

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*(2), 190-201. doi: 10.1037/0012-1649.36.2.190

Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In R. V. Kail (Ed.), *Advances in child development and behavior, vol. 30.* (pp. 153-187). San Diego, CA, US: Academic Press.

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science, 13*(3), 99-102. doi:10.1111/j. 09637214.2004.00283.x

Bahrick, L. E., Netto, D., & Hernandez-Keif, M. (1998). Intermodal perception of adult and child faces and voices by infants. *Child Development*, *69*(5), 1263-1275. doi: 10.1111/j. 1467-8624.1998.tb06210.x

Bahrick, L. E., & Todd, J. T. (2012). Multisensory processing in autism spectrum disorders: Intersensory processing disturbance as a basis for atypical development. *The new handbook of multisensory processes* .In B. E. Stein (Ed.), The new handbook of multisensory processes (p. 657-674), Cambridge, MA: MIT Press

Banati, R. B., Goerres, G. W., Tjoa, C., Aggleton, J. P., & Grasby, P. (2000). The functional anatomy of visual-tactile integration in man: a study using positron emission tomography. *Neuropsychologia*, *38*(2), 115-124. doi:10.1016/S0028-3932(99)00074-3

Bartholomeusz, H. H., Courchesne, E., & Karns, C. M. (2002). Relationship between head circumference and brain volume in healthy normal toddlers, children, and adults. *Neuropediatrics*, *33*(05), 239-241.

Beauchamp, M. S. (2005). Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics*, *3*(2), 93-113. doi: 10.1385/NI:3:2:093

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*(5), 809-823. doi:10.1016/S0896-6273(04)00070-4

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, *34*(1), 149-159. doi:10.1016/S0896-6273(02)00642-6

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *The Journal of Neuroscience*, *30*(7), 2414-2417. doi: 10.1523/JNEUROSCI. 4865-09.2010

Bebko, J. M., Weiss, J. A., Demark, J. L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *Journal of Child Psychology and Psychiatry, 47*(1), 88-98. doi:10.1111/j.1469- 7610.2005.01443.x

Belmonte, M. K., Allen, G., Beckel-Mitchener, A., Boulanger, L. M., Carper, R. A., & Webb, S. J. (2004). Autism and abnormal development of brain connectivity. *The Journal of Neuroscience*, *24*(42), 9228-9231. doi: 10.1523/JNEUROSCI.3340-04.2004

Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., & Barrett, L. F. (2011). Corrigendum: Amygdala volume and social network size in humans. *Nature Neuroscience*, *14*(9), 1217-1217. doi:10.1038/nn0911-1217e

Blakemore, S. J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. *Brain*, *128*(7), 1571-1583. DOI: 10.1093/brain/awh500

Blatt, G. J., Fitzgerald, C. M., Guptill, J. T., Booker, A. B., Kemper, T. L., & Bauman, M. L. (2001). Density and distribution of hippocampal neurotransmitter receptors in autism: an autoradiographic study. *Journal of Autism and Developmental Disorders*, *31*(6), 537-543. doi: 10.1023/A:1013238809666

Boddaert, N., Chabane, N., Gervais, H., Good, C. D., Bourgeois, M., Plumet, M. H., Barthelemy, X. & Zilbovicius, M. (2004). Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. *Neuroimage*, *23*(1), 364-369. doi:10.1016/j.neuroimage.2004.06.016

Bryson, C.Q. (1972). Short-term memory and crossmodal information processing in autistic children. *Journal of Learning Disabilities, 5,* 81–91. 1141. doi: 10.1177/002221947200500204

Bushara, K. O., Grafman, J., & Hallett, M. (2001). Neural correlates of auditory–visual stimulus onset asynchrony detection. *The Journal of Neuroscience*, *21*(1), 300-304.

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in cognitive sciences*, *11*(5), 219-227. doi:10.1016/j.tics.2007.02.005

Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience*, *7*(6), 477-484. doi:10.1038/nrn1909

Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology*, *17*(1), 20-25. doi:10.1016/j.cub.2006.10.052

Callan, D. E., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. *Cognitive Brain Research*, *10*(3), 349-353. doi:10.1016/S0926-6410(00)00054-9

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral cortex*, *11*(12), 1110-1123. doi: 10.1093/cercor/11.12.1110

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, W. R. P., Iversen, S. D., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*(5312), 593-596. doi: 10.1126/science.276.5312.593

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*(11), 649-657. doi:10.1016/S0960-9822(00)00513-3

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1), 191-205. doi:10.1016/j.jphysparis.2004.03.018

Campbell, R. & Dodd, B. (1980). Hearing by eye. Quarterly Journal of Experimental Psychology, 32, 85-99. doi:10.1080/00335558008248235

Carmichael, S. T., & Price, J. L. (1995). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology*, *363*(4), 642-664. doi: 10.1002/cne.903630409

Casanova, M. F., Buxhoeveden, D. P., Switala, A. E., & Roy, E. (2002). Minicolumnar pathology in autism. *Neurology*, *58*(3), 428-432. doi: 10.1212/ WNL.58.3.428

Casanova, M. F., van Kooten, I. A., Switala, A. E., van Engeland, H., Heinsen, H., Steinbusch, H. W., ... & Schmitz, C. (2006). Minicolumnar abnormalities in autism. *Acta neuropathologica*, *112*(3), 287-303. doi: 10.1007/s00401-006-0085-5

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839-1849. doi:10.1093/brain/awf189

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, *129*(3), 564-583. doi:10.1093/brain/awl004

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Jama*, *285*(24), 3093-3099. doi:10.1001/jama.285.24.3093.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature neuroscience*, *2*(10), 913-919. doi:10.1038/13217

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893. doi: 10.1037/0022-3514.76.6.893

Cho, Y. T., Fromm, S., Guyer, A. E., Detloff, A., Pine, D. S., Fudge, J. L., & Ernst, M. (2013). Nucleus accumbens, thalamus and insula connectivity during incentive anticipation in typical adults and adolescents. *Neuroimage*, *66*, 508-521. doi:10.1016/j.neuroimage. 2012.10.013

Cohen, I. (1994). An artificial neural network analogue of learning in autism. *Biological Psychiatry, 36*, 5-20. doi:10.1016/0006-3223(94)90057-4

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, *3*(3), 201-215. doi:10.1038/nrn755

Courchesne, E., Campbell, K., & Solso, S. (2011). Brain growth across the life span in autism: age-specific changes in anatomical pathology. *Brain research*, *1380*, 138-145. doi: 10.1016/j.brainres.2010.09.101

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *Journal of the American Medical Association*, *290*(3), 337-344. doi:10.1001/jama.290.3.337.

Courchesne, E., & Pierce, K. (2005)**.** Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Current Opinion in Neurobiology*, *15*(2), 225-230. doi:10.1016/j.conb.2005.03.001

Cox, R. W. (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, An International Journal 29*,162-173. doi:10.1006/cbmr.1996.0014

Craig, A. D. (2008). Interoception and emotion: a neuroanatomical perspective. *Handbook of emotions*, *3*, 272-88.

Dahlgren, S. O., & Gillberg, C. (1989). Symptoms in the first two years of life. *European Archives of Psychiatry and Neurological sciences*, *238*(3), 169-174. doi: 10.1007/BF00451006

Dawson, G., & Watling, R. (2000). Interventions to facilitate auditory, visual, and motor integration in autism: A review of the evidence. *Journal of Autism and Developmental Disorders. Special Issue: Treatments for People with Autism and Other Pervasive Developmental Disorders: Research Perspectives, 30*(5), 415-421. doi:10.1023/A:1005547422749

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of autism and developmental disorders*, *28*(6), 479-485. doi: 10.1023/A:1026043926488

D'Esposito, M., Ballard, D., Aguirre, G. K., & Zarahn, E. (1998) Human prefrontal cortex is not specific for working memory: A functional MRI study. *Neuroimage 8*, 274–282. doi: 10.1016/S0926-6410(98)00004-4

de Boer-Schellekens, L., Eussen, M., & Vroomen, J. (2013). Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Frontiers in Integrative Neuroscience*, *7*. doi: 10.3389/fnint.2013.00008 de Gelder, B., Vroomen, J., & Van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology. Special Issue: Face Recognition, 3*(1), 69-86. doi:10.1080/09541449108406220

Deutsch, C. K., & Joseph, R. M. (2003). Brief report: cognitive correlates of enlarged head circumference in children with autism. *Journal of Autism and Developmental Disorders*, *33*(2), 209-215. doi: 10.1023/A:1022903913547

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, *118*(1), 279-306. doi: 10.1093/brain/118.1.279

DiCicco-Bloom, E., Lord, C., Zwaigenbaum, L., Courchesne, E., Dager, S. R., Schmitz, C., ... & Young, L. J. (2006). The developmental neurobiology of autism spectrum disorder. *The Journal of Neuroscience*, *26*(26), 6897-6906. doi: 10.1523/JNEUROSCI.1712-06.2006

Di Martino, A., Ross, K., Uddin, L. Q., Sklar, A. B., Castellanos, F. X., & Milham, M. P. (2009). Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. *Biological psychiatry*, *65*(1), 63-74. doi:10.1016/j.biopsych.2008.09.022

Di Martino, A., Kelly, C., Grzadzinski, R., Zuo, X. N., Mennes, M., Mairena, M. A., ... & Milham, M. P. (2011). Aberrant striatal functional connectivity in children with autism. *Biological psychiatry*, *69*(9), 847-856. doi: 10.1016/j.biopsych.2010.10.029

Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., ... & Deen, B. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, *19*(6), 659-667. doi:10.1038/mp.2013.78

Dinstein, I., Pierce, K., Eyler, L., Solso, S., Malach, R., Behrmann, M., & Courchesne, E. (2011). Disrupted neural synchronization in toddlers with autism. *Neuron*, *70*(6), 1218-1225. doi: 10.1016/j.neuron.2011.04.018

Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology, 11*(4), 478-484. doi:10.1016/0010-0285(79)90021-5

Doehrmann, O., Naumer, M. J., Volz, S., Kaiser, J., & Altmann, C. F. (2008). Probing category selectivity for environmental sounds in the human auditory brain. *Neuropsychologia*, *46*(11), 2776-2786. doi:10.1016/j.neuropsychologia.2008.05.011

Donohue, S. E., Darling, E. F., & Mitroff, S. R. (2012). Links between multisensory processing and autism. *Experimental Brain Research*, *222*(4), 377-387.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1), 145-177. doi:10.1016/j.cognition.2003.11.002

Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in systems neuroscience*, *5*, 1. doi: 10.3389/fnsys.2011.00001

Edelman, G. M (1992). *Bright air, brilliant fire: On the matter of the mind.* New York : Basic Books.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *The Journal of Neuroscience*, *20*(16), 6159-6165.

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U., Church, J. A., Miezin, F. M., ... & Petersen, S. E. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Computational Biology*, *5*(5), 1-14. doi: 10.1371/journal.pcbi. 1000381

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1-47. doi: 10.1093/cercor/1.1.1

Fenwick, K. D., & Morrongiello, M. A. (1998). Spatial co-location and infants' learning of auditory-visual associations. *Infant Behavior and Development*, *21*(4), 745-759. doi: 10.1016/S0163-6383(98)90042

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron 33*, 341-355. doi:10.1016/S0896-6273(02)00569-X

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*(2), 381-389. doi: 10.1007/ s00221-010-2240-4

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, *103*(26), 10046-10051. doi: 10.1073/pnas. 0604187103

Frens, M. A., & Van Opstal, A. J. (1995). A quantitative study of auditory-evoked saccadic eye movements in two dimensions. *Experimental Brain Research*, *107*(1), 103-117.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78-84. doi:10.1016/S1364-6613(00)01839-8

Fries, W. (1984). Cortical projections to the superior colliculus in the macaque monkey: a retrograde study using horseradish peroxidase. *Journal of Comparative Neurology*, *230*(1), 55-76. doi: 10.1002/cne.902300106

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, *133*(4), 694. doi: 10.1037/0033-2909.133.4.694

Frith, C. D. (2007). The social brain?. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362*(1480), 671-678. doi: 10.1098/rstb.2006.2003

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*(16), R724-R732. doi:10.1016/j.cub.2007.05.068

Funnell, E. (2001). Evidence for scripts in semantic dementia: Implications for theories of semantic memory. *Cognitive Neuropsychology*, *18*(4), 323-341 doi: 10.1080/02643290042

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of ‘theory of mind’. *Trends in cognitive sciences*, *7*(2), 77-83. doi:10.1016/S1364-6613(02)00025-6

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, *16*(2), 123-144. doi:10.1016/S0926-6410(02)00244-6

Ghazanfar, A. A., Chandrasekaran, C., & Logothetis, N. K. (2008). Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *The Journal of Neuroscience*, *28*(17), 4457-4469. doi: 10.1523/ JNEUROSCI.0541-08.2008

Gibson, E. J. (1969) *Principles of perceptual learning and development*. New York : Appleton Century –Crofts.

Gibson, J. J. (1966). *The sense considered as perceptual systems*. Boston, MA : Houghton Mifflin.

Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *Neuroimage*, *22*(4), 1628-1635. doi: 10.1016/j.neuroimage.2004.03.049

Goebel, R., & van Atteveldt, N. (2009). Multisensory functional magnetic resonance imaging: a future perspective. *Experimental brain research*, *198*(2-3), 153-164. doi:10.1007/ s00221-009-1881-7

Gogate, L. J., & Bahrick, L. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology, 69*(2), 133-149. doi:10.1006/jecp.1998.2438

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development, 71*(4), 878-894. doi:10.1111/1467-8624.00197

Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word-comprehension: an ecological–dynamic systems view. *Developmental Science*, *4*(1), 1-18. doi:10.1111/1467-7687.00143

Glezer, L. S., & Riesenhuber, M. (2013). Individual variability in location impacts orthographic selectivity in the “visual word form area”. *The Journal of Neuroscience*, *33*(27), 11221-11226. doi:10.1523/JNEUROSCI.5002-12.2013

Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic resonance in medicine*, *44*(1), 162-167. doi: 10.1002/1522-2594(200007)

Gotts, S. J., Simmons, W. K., Milbury, L. A., Wallace, G. L., Cox, R. W., & Martin, A. (2012). Fractionation of social brain circuits in autism spectrum disorders. *Brain*, *160,* 1-15. doi: 10.1093/brain/aws160

Grandin, T. (1995). How people with autism think. In *Learning and cognition in autism* (pp. 137-156). Springer US.

Hadjikhani, N., & Roland, P. E. (1998). Cross-modal transfer of information between the tactile and the visual representations in the human brain: a positron emission tomographic study. *The Journal of Neuroscience*, *18*(3), 1072-1084.

Hadjikhani, N., Joseph, R. M., Snyder, J., & Tager-Flusberg, H. (2006). Anatomical Difference in the Mirror Neuron System and Social Cognition Network in Autism. *Cerebral Cortex, 16,* 1276-1282. do:10.1093/cercor/hhj069.

Hadjikhani, N., Joseph, R. M., Snyder, J., & Tager-Flusberg, H. (2007). Abnormal activation of the social brain during face perception in autism. *Human Brain Mapping*, *28*(5), 441-449. doi: 10.1002/hbm.20283

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. S*cience*, *304*(5669), 438-441. doi: 10.1126/science.1095455

Hamilton, A. F. D. C., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *The Journal of Neuroscience*, *26*(4), 1133-1137. doi: 10.1523/ JNEUROSCI.4551-05.2006

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science 293*, 2425–2430. doi: 10.1126/science.1063736

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223-233. doi:10.1016/ S1364-6613(00)01482-0

Hazlett, H. C., Poe, M., Gerig, G., Smith, R. G., Provenzale, J., Ross, A., ... & Piven, J. (2005). Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. *Archives of General Psychiatry*, *62*(12), 1366-1376. doi:10.1001/ archpsyc.62.12.1366.

Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *The Journal of Neuroscience*, *27*(30), 7881-7887. doi: 10.1523/ JNEUROSCI.1740-07.2007

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—it's my area: or is it?. Journal of *Cognitive Neuroscience, 20*(12), 2125-2136. doi: 10.1162/jocn.2008.20148

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, *63*(3), 289. do: 0.1037/h0039516

Higuchi, S. I., & Miyashita, Y. (1996). Formation of mnemonic neuronal responses to visual paired associates in inferotemporal cortex is impaired by perirhinal and entorhinal lesions. *Proceedings of the National Academy of Sciences*, *93*(2), 739-743.

Hill, A., Bölte, S., Petrova, G., Beltcheva, D., Tacheva, S., & Poustka, F. (2001). Stability and interpersonal agreement of the interview-based diagnosis of autism. *Psychopathology*, *34*(4), 187-191. doi:10.1159/000049305

Hoffman, K. L., Gothard, K. M., Schmid, M. C., & Logothetis, N. K. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Current Biology*, *17*(9), 766-772. doi:10.1016/j.cub.2007.03.040

Hughes, J. R. (2007). Autism: the first firm finding= underconnectivity?. *Epilepsy & Behavior*, *11*(1), 20-24. doi:10.1016/j.yebeh.2007.03.010

Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G., & Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(1), 131. doi: 10.1037/0096-1523.20.1.131

Hultman, C. M., Sparén, P., & Cnattingius, S. (2002). Perinatal risk factors for infantile autism. *Epidemiology*, *13*(4), 417-423.

Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders, 36*, 77–90. doi: 10.1007/s10803-005-0044-3

Insausti, R., Amaral, D. G., & Cowan, W. M. (1987). The entorhinal cortex of the monkey: III. Subcortical afferents. *Journal of Comparative Neurology*, *264*(3), 396-408. doi: 10.1002/cne.902640307

James, T. W., Stevenson, R. A., & Kim, S. (2012). Inverse effectiveness in multisensory processing. *The new handbook of multisensory processes. MIT Press, Cambridge.*

Jiang, W., Jiang, H., & Stein, B. E. (2002). Two corticotectal areas facilitate multisensory orientation behavior. *Journal of Cognitive Neuroscience*, *14*(8), 1240-1255. doi: 10.1162/089892902760807230

Jiang, W., Jiang, H., & Stein, B. E. (2006). Neonatal cortical ablation disrupts multisensory development in superior colliculus. *Journal of Neurophysiology*, *95*(3), 1380-1396. doi: 10.1152/jn.00880.2005

Jiang, Y. H., Tsai, T. F., Bressler, J., & Beaudet, A. L. (1998). Imprinting in angelman and Prader-Willi syndromes. *Current Opinion in Genetics & Development*, *8*(3), 334-342. doi: 10.1016/S0959-437X(98)80091-9

Joassin, F., Pesenti, M., Maurage, P., Verreckt, E., Bruyer, R., & Campanella, S. (2011). Cross-modal interactions between human faces and voices involved in person recognition. *Cortex*, *47*(3), 367-376. doi: 10.1016/j.cortex.2010.03.003

Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., De Haan, M., ... & Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology*, *17*(3), 599-619. doi: 10.1017/S0954579405050297

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, *9*(11), 512-518. doi:10.1016/j.tics.2005.09.009

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: evidence from an FMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex*, *17*(4), 951-961. doi: 10.1093/cercor/bhl006

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Social Neuroscience*, *4*(2), 135-152. doi:10.1080/17470910802198510

Keehn, B., Shih, P., Brenner, L. A., Townsend, J., & Müller, R. A. (2013). Functional connectivity for an “Island of sparing” in autism spectrum disorder: an fMRI study of visual search. *Human Brain Mapping*, *34*(10), 2524-2537. doi: 10.1002/hbm.22084

Kelly, A. C., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., Margulies, D. S., Castellanos, F. X. & Milham, M. P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cerebral Cortex*, *19*(3), 640-657. doi: 10.1093/cercor/bhn117

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L. & Gallese, V. 2004 A touching sight: SII/PV activation during the observation and experience of touch. *Neuron 42,* 335–346. doi:10.1016/S0896-6273(04)00156-4

King, A. J., & Carlile, S. (1993). Changes induced in the representation of auditory space in the superior colliculus by rearing ferrets with binocular eyelid suture. *Experimental Brain Research*, *94*(3), 444-455. doi: 10.1007/BF00230202

Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., ... & Vaitl, D. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system:: an event-related fMRI study. *Neuroimage*, *20*(2), 1086-1095. doi: 10.1016/S1053-8119(03)00381-1

Kleinhans, N. M., Richards, T., Sterling, L., Stegbauer, K. C., Mahurin, R., Johnson, L. C., Greenson, J., Dawson, G. & Aylward, E. (2008). Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain*, *131*(4), doi: 1000-1012. 10.1093/brain/awm334

Kleinhans, N. M., Richards, T., Weaver, K., Johnson, L. C., Greenson, J., Dawson, G., & Aylward, E. (2010). Association between amygdala response to emotional faces and social anxiety in autism spectrum disorders. *Neuropsychologia*, *48*(12), 3665-3670. doi: 10.1016/j.neuropsychologia.2010.07.022

Klinger, L., Dawson, G., Autistic disorder. In E. J. Mash, & R.A. Barkley. (1996). *Child psychopathology.* New York: Guilford, 409-454.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal Neuroscience*, *21*(16), RC159, 1-5.

Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage*, *37*(4), 1445-1456. doi:10.1016/j.neuroimage.2007.06.020

Kuhl, P. K. (2007). Is speech learning 'gated'by the social brain?. *Developmental Science*, *10*(1), 110-120. DOI: 10.1111/j.1467-7687.2007.00572

Kuhl, P. K., & Meltzoff, A. N. (1982). The biomodal perception of speech in infancy. *Science, 218*, 1138-1141.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior & Development, 7*(3), 361-381. doi:10.1016/S0163-6383(84)80050-8

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., & Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in Integrative Neuroscience*, *4*, 129. doi: 10.3389/fnint.2010.00129

Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., & Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Experimental Brain Research*, *166*(3-4), 289-297. doi: 10.1007/ s00221-005-2370-2

Lavoie, S. L., Hancock, L. N., Bebko, J. M. (in preparation). Intermodal perception in children with ASDs: A linguistic specific deficit?

Lecavalier, L., Aman, M. G., Scahill, L., McDougle, C. J., McCracken, J. T., Vitiello, B., ... & Cronin, P. (2006). Validity of the autism diagnostic interview-revised. *American Journal on Mental Retardation*, 111, 199–215, 228.

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain*, *137*(1), 12-32. doi:10.1093/brain/awt162

Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behavior and Development*, *13*(3), 343-354. doi:10.1016/0163-6383(90)90039-B

Levitt, J. G., Blanton, R. E. Smalley, S., Thompson, P.M., Guthrie, D., McCracken, J. T., Sadoun, T., Heinichen, L., & Toga, A. W. (2003). Cortical sulcal maps in autism. *Cebral Cortex, 13* (7), 728-735. doi: 10.1093/cercor/13.7.728.

Lewis, J. D., & Elman, J. L. (2008). Growth-related neural reorganization and the autism phenotype: a test of the hypothesis that altered brain growth leads to altered connectivity. *Developmental science*, *11*(1), 135-155. DOI: 10.1111/j.1467-7687.2007.00634

Lewkowicz, D. J. (1999). The development of temporal and spatial intermodal perception. In G. Aschersleben, T. Bachmann, & J. Musseler, (Eds.), *Cognitive contributions to the perception of spatial and temporal events. Advances in Psychology, 129* (pp. 395-420). New York: Elsevier Science. doi:10.1016/S0166-4115(99)80038-7

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: an epigenetic systems/limitations view. *Psychological Bulletin*, *126*(2), 281. doi: 10.1037/0033-2909.126.2.281

Lord, C., Rutter, M., & Le Couteru, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659-685.

Lord, C., Rutter, M., & LeCouteur, A. (2003). *Autism Diagnostic Interview-Revised*. Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). *Autism Diagnostic Observation Schedule.* Western Psychological Services, Los Angeles

Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Jr, Leventhal, B.L., DiLavre, P.C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205–223.

Love, S. A., Pollick, F. E., & Latinus, M. (2011). Cerebral correlates and statistical criteria of cross-modal face and voice integration. *Seeing and Perceiving*, *24*(4), 351-367. doi: 10.1163/187847511X584452

Loveland, K.A., Tunali-Kotoski, B., Chen, R., Brelsford, K.A., Ortegon, J., & Pearson, D.H. (1995). Intermodal perception of affect in persons with autism or Down syndrome. *Development and Psychopathology*, *7*, 409–418.

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage*, *21*(2), 725-732.

Maddock, R. J., Garrett, A. S., Buonocore, M. H. 2001. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience 104,* 667–676. doi:10.1016/S0306-4522(01)00108-7

Maddock, R. J., Garrett, A. S., & Buonocore, M. H. (2003). Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Human Brain Mapping*, *18*(1), 30-41. doi: 10.1002/hbm.10075

Marchant, J. L., Ruff, C. C., & Driver, J. (2012). Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Human brain mapping*, *33*(5), 1212-1224. doi: 10.1002/hbm.21278

Materna, S., Dicke, P. W., & Thier, P. (2008). The posterior superior temporal sulcus is involved in social communication not specific for the eyes. *Neuropsychologia*, *46*(11), 2759-2765.

Maximo, J. O., Cadena, E. J., & Kana, R. K. (2014). The implications of brain connectivity in the neuropsychology of autism. *Neuropsychology Review*, *24*(1), 16-31. doi:10.1016/ j.neuropsychologia.2008.05.016

McGurk, H., & MacDonald, J. (1978). Auditory–visual coordination in the first year of life. *International Journal of Behavioral Development, 1*(3), 229-239. doi:10.1016 S0163-6383(78)80035-6

McIntosh, A. R. (1999). Mapping cognition to the brain through neural interactions. *Memory*, *7*(5-6), 523-548. doi:10.1080/096582199387733

McIntosh, A. R. (2004). Contexts and catalysts. *Neuroinformatics*, *2*(2), 175-181. 10.1385/NI: 2:2:175

McIntosh, A. R., Chau, W. K., & Protzner, A. B. (2004). Spatiotemporal analysis of event-related fMRI data using partial least squares. *Neuroimage*, *23*(2), 764-775.

McIntosh, A. R., Rajah, M. N., Lobaugh, N. J (2003). Functional connectivity of the medial temporal lobe related to learning and awareness. *Neuroscience, 23*, 6250-6528. doi: 10.1016/j.neuroimage.2004.05.018

Megnin, O., Flitton, A., RG Jones, C., de Haan, M., Baldeweg, T., & Charman, T. (2012). Audiovisual speech integration in autism spectrum disorders: ERP evidence for atypicalities in lexical-semantic processing. *Autism Research*, *5*(1), 39-48. doi: 10.1002/ aur.231

Mendelson, M. J. (1979). Acoustic-optical correspondence and auditory-visual coordination in infancy. *Canadian Journal of Psychology, 33*(4), 334-346. doi: 10.1037/h0081730

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, *221*(4608), 389-391. doi: 10.1126/science.6867718

Meredith, A. M., & Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, *365*(2), 350-354. doi: 10.1016/0006-8993(86)91648-3

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology, 75*(5), 1843-1857. doi: 10.1016/0006-8993(86)91648-3

Mesulam, M., & Mufson, E. J. (1982). Insula of the old world monkey. III: Efferent cortical output and comments on function. *Journal of Comparative Neurology*, *212*(1), 38-52. doi: 10.1002/cne.902120104

Mesulam, M. M., Mufson, E. J., Levey, A. I., & Wainer, B. H. (1984). Atlas of cholinergic neurons in the forebrain and upper brainstem of the macaque based on monoclonal choline acetyltransferase immunohistochemistry and acetylcholinesterase histochemistry. *Neuroscience*, *12*(3), 669-686. doi:10.1016/0306-4522(84)90163-5

Miller, L. M., & D'esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *The Journal of Neuroscience*, *25*(25), 5884-5893. doi: 10. 1523/ JNEUROSCI. 0896-05. 2005

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*(7), 1349-1358. doi: 10.1007/ s10803-007-0521-y

Morrell, L. K. (1968). Temporal characteristics of sensory interaction in choice reaction times. *Journal of Experimental Psychology*, *77*(1), 14. doi: 10.1037/h0021289

Morris, J. P., Pelphrey, K. A., & McCarthy, G. (2008). Perceived causality influences brain activity evoked by biological motion. *Social neuroscience*, *3*(1), 16-25. doi: 10.1080/17470910701476686

Muller, R., Shih, P., Keehn, B., Deyoe, J. R., Leyden, K. M., Shukla, D. K. (2011). Underconnected, but how? A survey of functional connectivity MRI studies in Autism Spectrum Disorders. *Cerebral Cortex, 21,* 2233-2243. doi: 10.1093/cercor/bhq296

Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Development and Psychopathology*, *7*(01), 63-82. doi: 10.1017/ S0954579400006349

Murray, E. A., & Bussey, T. J. (1999). Perceptual–mnemonic functions of the perirhinal cortex. *Trends in cognitive sciences*, *3*(4), 142-151. doi:10.1016/S1364-6613(99)01303-0

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage*, *59*(1), 781-787. doi:10.1016/ j.neuroimage.2011.07.024

Nichols, T., Brett, M., Andersson, J., Wagner, T., & Poline, J. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage, 25*(3), 653-660. doi: 10.1016/j.neuroimage. 2004.12.005

Noesselt, T., Bergmann, D., Heinze, H. J., Münte, T., & Spence, C. (2012). Coding of multisensory temporal patterns in human superior temporal sulcus. *Multisensory Perception and Action: Psychophysics, Neural Mechanisms, and Applications*, 64(6), 44-54. doi: 10.3389/fnint.2012.00064

O'Hare, J. (1991). Perceptual integration. *Journal of the Washington Academy of Sciences, 81,* pp. 44–59.

Olivares, R., Montiel, J., & Aboitiz, F. (2001). Species differences and similarities in the fine structure of the mammalian corpus callosum. *Brain, Behavior and Evolution*, *57*(2), 98-105. doi:10.1159/000047229

Ornitz, E. M. Autism at the interface between sensory and information processing. In Autism: Nature, diagnosis, and treatment. Dawson, Geraldine (Ed), (1989). (pp. 174-207). New York, NY, US: Guilford Press, xxv, 417 pp.

Ozaki, T. J. (2011). Frontal-to-parietal top-down causal streams along the dorsal attention network exclusively mediate voluntary orienting of attention. *PLoS One*, *6*(5), doi: e20079. 10.1371/journal.pone.0020079

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523-540. doi:10.1207/s15374424jccp3403_8

Pearson, R. C. A., Brodal, P., Gatter, K. C., & Powell, T. P. S. (1982). The organization of the connections between the cortex and the claustrum in the monkey. *Brain research*, *234*(2), 435-441. doi:10.1016/0006-8993(82)90883-6

Pelphrey, K. A., Shultz, S., Hudac, C. M., & Vander Wyk, B. C. (2011). Research review: constraining heterogeneity: the social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *52*(6), 631-644. doi: 10.1111/j. 1469-7610.2010.02349.x

Pelphrey, K. A., Morris, J., & Mccarthy, G. (2004). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. Journal of *Cognitive Neuroscience*, *16*(10), 1706-1716. doi: 10.1162/0898929042947900

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review Psychology.*, *56*, 89-114. doi: 10.1146/annurev.psych.56.091103.070225

Perlman, S. B., Hudac, C. M., Pegors, T., Minshew, N. J., & Pelphrey, K. A. (2011). Experimental manipulation of face-evoked activity in the fusiform gyrus of individuals with autism. *Social Neuroscience*, *6*(1), 22-30. doi:10.1080/17470911003683185

Perrott, D. R., Saberi, K., Brown, K., & Strybel, T. Z. (1990). Auditory psychomotor coordination and visual search performance. *Perception & Psychophysics*, *48*(3), 214-226. doi: 10.3758/BF03211521

Pinkham, A. E., Hopfinger, J. B., Pelphrey, K. A., Piven, J., & Penn, D. L. (2008). Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophrenia research*, *99*(1), 164-175. doi:10.1016/j.schres.2007.10.024

Polley, D. B., Hillock, A. R., Spankovich, C., Popescu, M. V., Royal, D. W., & Wallace, M. T. (2008). Development and plasticity of intra-and intersensory information processing. *Journal of the American Academy of Audiology*, *19*(10), 780-798. doi: 10.3766/jaaa. 19.10.6

Poustka, F., Lisch, S., Rühl, D., Sacher, A., Schmötzer, G., & Werner, K. (1996). The standardized diagnosis of autism, Autism Diagnostic Interview-Revised: interrater reliability of the German form of the interview. *Psychopathology*, *29*(3), 145-153. doi: 10.1159/000284984

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of neurophysiology*, *74*(3), 1192-1199.

Radeau, M. (1994). Auditory-visual spatial interaction and modularity. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 13*(1), 3-51.

Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, *28*(2), 617-625. doi:10.1016/S0896-6273(00)00138-0

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, *47*(2), 388-396. doi:10.1016/j.neuropsychologia.2008.09.017

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *32*(1), 123-142. doi:10.1016/j.neubiorev.2007.06.004

Redcay, E., & Courchesne, E. (2005). When is the brain enlarged in autism? A meta-analysis of all brain size reports. *Biological psychiatry*, *58*(1), 1-9. doi:10.1016/j.biopsych. 2005.03.026

Reisberg, D., McLean, J., & Goldfield, A. (1987). A lipreading advantage with intact auditory stimuli. *Hearing by Eye: The Psychology of Lipreading, London: Lawrence Earlbaum Associates*.

Rippon, G., Brock, J., Brown, C., & Boucher, J. (2007). Disordered connectivity in the autistic brain: challenges for the 'new psychophysiology'. *International Journal of Psychophysiology*, *63*(2), 164-172. doi:10.1016/j.ijpsycho.2006.03.012

Rosenblum, L. D., Johnson, J. A., & Saldana, H. M. (1996). Point-light facial displays enhance comprehension of speech in noise. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1159-1170. doi:10.1044/jshr.3906.1159

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics*, *59*(3), 347-357. doi: 10.3758/BF03211902

Royal, D. W., Carriere, B. N., & Wallace, M. T. (2009). Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Experimental brain research*, *198*(2-3), 127-136. doi: 10.1007/s00221-009-1772-y

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, *2*(5), 255-267. doi: 10.1034/j.1601-183X.2003.00037.x

Rutter, M. (2005). Aetiology of autism: findings and questions. *Journal of Intellectual Disability Research*, *49*(4), 231-238. DOI: 10.1111/j.1365-2788.2005.00676.x

Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., & Cox, R. W. (2004, April). SUMA: an interface for surface-based intra-and inter-subject analysis with AFNI. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on* (pp. 1510-1513). IEEE. doi: 10.1109/ISBI.2004.1398837

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*(2), 235-239. doi:10.1016/j.conb.2006.03.001

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*(11), 1435-1446. doi:10.1016/j.neuropsychologia.2004.04.015

Scannell, J. W., Burns, G. A. P. C., Hilgetag, C. C., O'Neil, M. A., & Young, M. P. (1999). The connectional organization of the cortico-thalamic system of the cat. *Cerebral Cortex*, *9*(3), 277-299. doi: 10.1093/cercor/9.3.277

Schipul, S. E., Keller, T. A., & Just, M. A. (2011). Inter-regional brain communication and its disturbance in autism. *Frontiers in Systems Neuroscience*, *5*(10), 1-11. doi: 10.3389/ fnsys.2011.00010

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., & Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron*, *45*(4), 625-635. doi:10.1016/j.neuron.2004.12.052

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400-2406. doi:10.1093/ brain/123.12.2400

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308-308. doi: 10.1038/385308a0

Seltzer, B., Cola, M. G., Gutierrez, C., Massee, M., Weldon, C., & Cusick, C. G. (1996). Overlapping and nonoverlapping cortical projections to cortex of the superior temporal sulcus in the rhesus monkey: double anterograde tracer studies. *Journal of Comparative Neurology*, *370*(2), 173-190. doi: 10.1002/(SICI)1096-9861(19960624)370:2

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: the role of the right ventromedial prefrontal cortex. *Journal of Cognitive Neuroscience*, *15*(3), 324-337. doi: 10.1162/089892903321593063

Shih, P., Shen, M., Öttl, B., Keehn, B., Gaffrey, M. S., & Müller, R. A. (2010). Atypical network connectivity for imitation in autism spectrum disorder. *Neuropsychologia*, *48*(10), 2931-2939. doi:10.1016/j.neuropsychologia.2010.05.035

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron*, *41*(4), 653-662. doi:10.1016/S0896-6273(04)00014-5

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry, 48*(8), 813-821. doi:10.1111/j.1469-7610.2007.01766.x

Smith, M. L., & Milner, B. (1989). Right hippocampal impairment in the recall of spatial location: encoding deficit or rapid forgetting?. *Neuropsychologia*, *27*(1), 71-81. doi:10.1016/0028-3932(89)90091-2

Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, *9*(8), 1007-1008. doi:10.1038/nn1728

Sparks, B. F., Friedman, S. D., Shaw, D. W., Aylward, E. H., Echelard, D., Artru, A. A., ... & Dager, S. R. (2002). Brain structural abnormalities in young children with autism spectrum disorder. *Neurology*, *59*(2), 184-192. doi: 10. 1212/ WNL. 59. 2. 184

Spelke, E. S., & Owsley, C. J. (1979). Intermodal exploration and knowledge in infancy. *Infant Behavior & Development, 2*(1), 13-27. doi:10.1016/S0163-6383(79)80004-1

Sporns, O., Tononi, G., & Edelman, G. M. (2000). Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Networks*, *13*(8), 909-922. doi:10.1016/S0893-6080(00)00053-8

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage*, *53*(1), 303-317. doi: 10.1016/j.neuroimage.2010.06.016

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT press.

Stein, B. E., Meredith, M., Huneycutt, W., & McDade, L. (1989). Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, *1*(1), 12-24. doi: 10.1162/jocn.1989.1.1.12

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255-266. doi: 10.1038/nrn2331

Stephan, K. E., Hilgetag, C. C., Burns, G. A., O'Neill, M. A., Young, M. P., & Kotter, R. (2000). Computational analysis of functional connectivity between areas of primate cerebral cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1393), 111-126. doi: 10.1098/rstb.2000.0552

Stevens, W. D., Kahn, I., Wig, G. S., & Schacter, D. L. (2012): Hemispheric asymmetry of visual scene processing in the human brain: Evidence from repetition priming and intrinsic activity. *Cerebral Cortex, 22,*1935–1949. doi: 10.1093/cercor/bhr273

Stevens, W. D., Tessler, M. H., Peng, C. S., & Martin, A. (2015). Functional connectivity constrains the category-related organization of human ventral occipitotemporal cortex. *Human Brain Mapping*, *36*(6), 2187-2206. doi: 10.1002/hbm.22764

Stevenson, R. A., Geoghegan, M. L., & James, T. W. (2007). Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Experimental Brain Research*, *179*(1), 85-95. doi: 10.1007/s00221-006-0770-6

Stevenson, R. A., & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage*, *44*(3), 1210-1223. doi:10.1016/j.neuroimage.2008.09.034

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, *34*(3), 691-697. doi: 10.1523/JNEUROSCI. 3615-13.2014

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*, 212-215. doi:10.1121/1.1907309

Szycik, G. R., Tausche, P., & Münte, T. F. (2008). A novel approach to study audiovisual integration in speech perception: localizer fMRI and sparse sampling. *Brain Research*, *1220*, 142-149. doi:10.1016/j.brainres.2007.08.027

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1430), 303-314. doi: 10.1098/rstb.2002.1198

Tekin, S., & Cummings, J. L. (2002). Frontal–subcortical neuronal circuits and clinical neuropsychiatry: an update. *Journal of Psychosomatic Research*, *53*(2), 647-654. doi: 10.1016/S0022-3999(02)00428-2

Thakkar, K. N., Polli, F. E., Joseph, R. M., Tuch, D. S., Hadjikhani, N., Barton, J. J., & Manoach, D. S. (2008). Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD). *Brain*, *131*(9), 2464-2478. doi: 10.1093/brain/awn099

Thelen, E. & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

Thut, G., Schultz, W., Roelcke, U., Nienhusmeier, M., Missimer, J., Maguire, R. P., & Leenders, K. L. (1997). Activation of the human brain by monetary reward. *Neuroreport*, *8*(5), 1225-1228.

Tomchek, S. D., & Dunn, W. (2007). Sensory processing in children with and without autism: a comparative study using the short sensory profile. *American Journal of Occupational Therapy*, *61*(2), 190-200. doi:10.5014/ajot.61.2.190

Tononi, G., Sporns, O., & Edelman, G. M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, *91*(11), 5033-5037.

Tononi, G., Sporns, O., & Edelman, G. M. (1996). A complexity measure for selective matching of signals by the brain. *Proceedings of the National Academy of Sciences*, *93*(8), 3422-3427.

Tranel, D., Bechara, A., & Denburg, N. L. (2002). Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. *Cortex*, *38*(4), 589-612. doi:10.1016/S0010-9452(08)70024-8

Uddin, L. Q., Supekar, K., & Menon, V. (2013). Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Frontiers in human neuroscience*, *7*. doi: 10.3389/fnhum.2013.00458

van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex*, *17*(4), 962-974. doi: 10.1093/cercor/bhl007

van der Smagt, M. J., van Engeland, H., & Kemner, C. (2007). Brief report: Can you see what is not there? low-level auditory-visual integration in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(10), 2014-2019. doi:10.1007/s10803-006-0346-0

Van Der Werf, Y. D., Jolles, J., Witter, M. P., & Uylings, H. B. (2003). Contributions of thalamic nuclei to declarative memory functioning. *Cortex*, *39*(4), 1047-1062. doi:10.1016/S0010-9452(08)70877-3

Villalobos, M. E., Mizuno, A., Dahl, B. C., Kemmotsu, N., & Müller, R. A. (2005). Reduced functional connectivity between V1 and inferior frontal cortex associated with visuomotor performance in autism. *Neuroimage*, *25*(3), 916-925. doi:10.1016/j.neuroimage.2004.12.022

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*(6), 3328-3342. doi: 10.1152/jn.90355.2008 von dem Hagen, E. A., Nummenmaa, L., Yu, R., Engell, A. D., Ewbank, M. P., & Calder, A. J. (2011). Autism spectrum traits in the typical population predict structure and function in the posterior superior temporal sulcus. *Cerebral Cortex*, *21*(3), 493-500. doi: 10.1093/cercor/bhq062

Vuilleumier, P., Armony, J. L., Clarke, K., Husain, M., Driver, J., & Dolan, R. J. (2002). Neural response to emotional faces with and without awareness: event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia*, *40*(12), 2156-2166. doi:10.1016/S0028-3932(02)00045-3

Walker-Andrews, A. S., Bahrick, L. E., Raglione, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology, 3*(2), 55-75. doi:10.1207/s15326969eco0302_1

Wallace, M. T., & Stein, B. E. (1994). Cross-modal synthesis in the midbrain depends on input from cortex. *Journal of Neurophysiology*, *71*(1), 429-432.

Wallace, M. T., Meredith, M. A., & Stein, B. E. (1992). Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research*, *91*(3), 484-488. doi:10.1007/BF00227844

Wallace, M. T., & Stein, B. E. (2000). Onset of cross-modal synthesis in the neonatal superior colliculus is gated by the development of cortical influences. *Journal of Neurophysiology*, *83*(6), 3578-3582.

Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *Journal of Neurophysiology*, *97*(1), 921-926. doi: 10.1152/jn.00497.2006

Watson, R., Latinus, M., Charest, I., Crabbe, F., & Belin, P. (2014). People-selctivity, audiovisual integration and hetermodality in the superior temporal sulcus. *Cortex, 50*(100), 1250136. Doi: 10.1016/j.cortex.2013.07.011

Waxman, S. G. (1977). Conduction in myelinated, unmyelinated, and demyelinated fibers. *Archives of Neurology*, *34*(10), 585-589. doi:10.1001/archneur.1977.00500220019003.

Wechsler, D. (1999). Wachsler Abbreviated Scale of Inelligence (WASI). *San Antonio, Psychological Corporation*

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638. doi: 10.1037/0033-2909.88.3.638

Wilkinson, L. K., Meredith, M. A., & Stein, B. E. (1996). The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Experimental Brain Research*, *112*(1), 1-10. doi: 10.1007/BF00227172

Williams, J. H., Massaro, D. W., Peel, N. J., Bosseler, A., & Suddendorf, T. (2004). Visual–auditory integration during speech imitation in autism. *Research in Developmental disabilities*, *25*(6), 559-575. doi:10.1016/j.ridd.2004.01.008

Wolf, N. S., Gales, M. E., Shane, E., & Shane, M. (2001). The developmental trajectory from amodal perception to empathy and communication: The role of mirror neurons in this process. *Psychoanalytic Inquiry*, *21*(1), 94-112. doi:10.1080/07351692109348925

Williams, J.H.G., Massaro, D.W., Peel, N.J., Bosseler, A., & Suddendorf, T. (2004). Visual-auditory integration during speech imitation in autism. *Research in Developmental Disabilities, 25,* 559– 575. doi:10.1016/j.ridd.2004.01.008

Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex*, *13*(10), 1034-1043. doi: 10.1093/cercor/13.10.1034

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in Neurosciences*, *29*(7), 359-366. doi:10.1016/j.tins.2006.06.004

Appendix A: Tables

Table 1

Number of participants with an ROI identified at the individual-level in the left and right pSTS for each method. The t-statistic is provided for the peak activation value.

Method	**Hemisphere**	**N (%)**	**t-statistic Mean (range)**
Mean			
	Left Only	0 (0%)	NA
	Right Only	2 (12%)	5.0 (3.80 - 6.20)
	Both Left and Right	8 (47%)	5.31 (3.72 - 7.20)
	Total	10 (59%)	5.27 (3.72 - 7.20)
Max			
	Left Only	3 (18%)	5.46 (4.10 - 6.45)
	Right Only	0 (0%)	NA
	Both Left and Right	0 (0%)	NA
	Total	3 (18%)	5.46 (4.10 - 6.45)
Superadditive			
	Left Only	2 (12%)	5.17 (5.03 - 5.30)
	Right Only	2 (12%)	3.45 (3.10 - 3.89)
	Both Left and Right	7 (41%)	4.71 (3.50 - 6.30)
	Total	11 (65%)	4.46 (3.10 - 6.30)
Temporal Synchrony			
	Left Only	2 (12%)	2.91 (2.85 - 2.97)
	Right Only	1 (6%)	3.62 (NA)
	Both Left and Right	11 (65%)	4.41 (2.75 - 6.81)
	Total	14 (82%)	4.25 (2.75 - 6.81)

Table 2

Beta values for the ROIs created with the Mean method per content condition in left and right hemispheres. The table also demonstrates the t-test used to test the Mean method within each content condition in each hemisphere.

Condition by Hemisphere	**Synchronous audio-visual**	**Mean of visual only and audio only**	***t-test***	**Cohen's d**
	M (SD)	**M (SD)**		
Social-Linguistic				
Left	0.56 (0.41)	0.26 (0.17)	$t(7) = 2.20, p = .06$	0.77
Right	0.32 (0.31)	0.16 (0.15)	$t(9) = 1.60\ p = .14$	0.50
Social-Nonlinguistic				
Left	0.44 (0.32)	0.12 (0.17)	$t(7) = 2.62, p = .03$	0.92
Right	0.35 (0.12)	0.17 (0..13)	$t(9) = 1.77, p = .11$	0.56
Mousetrap				
Left	0.39 (0.38)	0.19 (0.26)	$t(7) = 1.77, p = .12$	0.62
Right	0.43 (0.33)	0.21 (0.18)	$t(9) = 254, p = .03$	0.80
Music				
Left	0.42 (0.46)	0.11 (0.13)	$t(7) = 2.11, p = .007$	0.74
Right	0.23 (0.15)	0.07 (0.08)	$t(9) = 3.35, p = .009$	1.06

Table 3

Beta values for the ROIs created with the Max method per content condition in left hemisphere only (no ROIs were identified for any participants in the right hemisphere using the Max method). The table also demonstrates the t-test used to test the Max method within each content condition.

Condition by Hemisphere	**Synchronous audio-visual** **M (SD)**	**Max value of visual only or audio only** **M (SD)**	***t-test***	**Cohen's d**
Social-Linguistic				
Left	0.81 (0.81)	1.02 (0.63)	*t*(2) = -0.26, *p* = .82	0.15
Social-Nonlinguistic				
Left	1.05 (0.41)	0.30 (0.37)	*t*(2) = 2.09, *p* = .17	1.21
Mousetrap				
Left	0.65 (0.19)	0.39 (0.40)	*t*(2) = 1.12, *p* = .38	0.64
Music				
Left	0.87 (0.82)	0.11 (0.07)	t(2) = 1.55, p = .26	0.89

Table 4

Beta values for the ROIs created with the Super-Additive method per content condition in left and right hemispheres. The table also demonstrates the t-test used to test the Super-Additive method within each content condition in each hemisphere

Condition by Hemisphere	**Synchronous audio-visual**	**Sum of visual only and audio only**	***t-test***	**Cohen's d**
	M (SD)	**M (SD)**		
Social-Linguistic				
Left	0.49 (0.46)	0.43 (0.33)	$t(8) = 0.41, p = .69$	0.13
Right	0.43 (0.38)	0.33 (0.25)	$t(8) = 0.75, p = .47$	0.24
Social-Nonlinguistic				
Left	0.42 (0.48)	0.32 (0.37)	$t(8) = 0.55, p = .56$	0.18
Right	0.29 (0.38)	0.38 (0.27)	$t(8) = -0.94, p = .38$	0.31
Mousetrap				
Left	0.27 (0.21)	0.18 (0.38)	$t(8) = 0.55, p = .60$	0.18
Right	0.45 (0.39)	0.40 (0.35)	$t(8) = 0.51, p = .62$	0.17
Music				
Left	0.29 (0.46)	0.14 (0.11)	$t(8) = 0.92\ p = .39$	0.31
Right	0.38 (0.50)	0.25 (0.26)	$t(8) = 0.62, p = .55$	0.20

Table 5

Beta values for the ROIs created with the Temporal Synchrony method per content condition in left and right hemispheres. The table also demonstrates the t-test used to test the Temporal Synchrony method within each content condition in each hemisphere

Condition by Hemisphere	**Synchronous audio-visual**	**Asynchronous audio-visual**	***t-test***	**Cohen's d**
	M (SD)	**M (SD)**		
Social-Linguistic				
Left	0.19 (0.26)	-0.17 (0.34)	*t*(12) = 2.58, *p* = .02	0.71
Right	0.18 (0.20)	-0.28 (0.47)	*t*(11) = 2.70, *p* = .02	0.78
Social-Nonlinguistic				
Left	0.18 (0.35)	-0.17 (0.28)	*t*(12) = 2.94, *p* = .01	0.81
Right	0.24 (0.50)	-0.12 (0.32)	*t*(11) = 2.54, *p* = .03	0.73
Mousetrap				
Left	0.13 (0.33)	-0.11 (0.34)	*t*(12) = 3.58, *p* = .004	0.99
Right	0.37 (0.49)	-0.08 (0.32)	*t*(11) = 3.25, *p* = .008	0.93
Music				
Left	0.04 (0.28)	-0.02 (0.30)	*t*(12) = 0.53, *p* = .61	0.14
Right	0.09 (0.25)	0.13 (0.40)	*t*(11) = -0.33, *p* = .75	0.09

Table 6

Comparison of Age and IQ scores between young adults with (N = 15) and without (N = 17) ASD.

	ASD **M (SD)**	**Without ASD** **M (SD)**	**T-test**
Age	21.87 (3.55)	21.67 (3.13)	$t(29) = 1.73, p = .86$
WASI Vocabulary T-Score	56.50 (8.29)	59.40 (13.27)	$t(29) = -.73, p = .47$
WASI Matrix Reasoning T-Score	52.94 (7.36)	56.49 (6.17)	$t(29) = -1.41, p = .17$
Fill Scale IQ Standard Score	108.37 (11.74)	114.20 (13.80)	$t(29) = -1.27, p = .22$

Table 7

Number and percent of participants with (N = 15) and without (N = 17) ASD with left and right pSTS ROIs.

Hemisphere by Group	**N (%)**	**t-statistic Mean (range)**
Without ASD		
Left Only	2 (12%)	2.91 (2.85 - 2.97)
Right Only	1 (6%)	3.62 (NA)
Both Left and Right	11 (65%)	4.41 (2.75 - 6.81)
Total	14 (82%)	4.25 (2.75 - 6.81)
ASD		
Left Only	1 (7%)	6.43 (NA)
Right Only	1 (7%)	4.28 (NA)
Both Left and Right	11 (73%)	4.25 (2.61 - 5.91)
Total	13 (87%)	4.36 (2.61 - 6.43)

Table 8

Temporal Synchrony ROI beta values per content condition in left and right hemispheres

Group	**Condition by Hemisphere**	**Synchronous audio-visual**	**Asynchronous audio-visual**	***t-test***
		M (SD)	**M (SD)**	
Without ASD	***Social-Linguistic***			
	Left	0.19 (0.26)	-0.17 (0.34)	*t*(12) = 2.58, *p* = .02
	Right	0.18 (0.20)	-0.28 (0.47)	*t*(11) = 2.70, *p* = .02
	Social-Nonlinguistic			
	Left	0.18 (0.35)	-0.17 (0.28)	*t*(12) = 2.94, *p* = .01
	Right	0.24 (0.50)	-0.12 (0.32)	*t*(11) = 2.54, *p* = .03
	Mousetrap			
	Left	0.13 (0.33)	-0.11 (0.34)	*t*(12) = 3.58, *p* = .004
	Right	0.37 (0.49)	-0.08 (0.32)	*t*(11) = 3.25, *p* = .008
	Music			
	Left	0.04 (0.28)	-0.02 (0.30)	*t*(12) = 0.53, *p* = .61
	Right	0.09 (0.25)	0.13 (0.40)	*t*(11) = -0.33, *p* = .75
ASD	***Social-Linguistic***			
	Left	0.21 (0.50)	-0.09 (0.36)	*t*(11) = 1.42, *p* = .18
	Right	0.21 (0.22)	-0.05 (0.44)	*t*(11) = 1.69, *p* = .12
	Social-Nonlinguistic			
	Left	0.09 (0.28)	-0.25 (0.40)	*t*(11) = 3.71, *p* = .003
	Right	-.03 (0.21)	-0.28 (0.41)	*t*(11) = 2.83, *p* = .01
	Mousetrap			
	Left	0.19 (0.52)	-0.08 (0.44)	*t*(11) = 2.31, *p* = .04
	Right	0.10 (0.24)	-0.09 (0.29)	*t*(11) = 2.35, *p* = .03
	Music			
	Left	0.10 (0.63)	-0.16 (0.37)	*t*(11) = 1.08, *p* = .30
	Right	0.18 (0.25)	-0.16 (0.50)	*t*(11) = 2.03, *p* = .07

Table 9

Number and percent of participants with (N = 15) and without (N = 17) ASD in which the left or right pSTS ROI was used when collapsing across hemispheres by using the peak voxel value from the collapsed condition.

Hemisphere by Group	**N (%)**
Without ASD	
Left	4 (23%)
Right	10 (59%)
Total	14 (82%)
ASD	
Left	7 (47%)
Right	6 (40%)
Total	13 (87%)

Table 10

Number and percent of participants with (N = 15) and without (N = 17) ASD in which the left or right pSTS ROI was used when collapsing across hemispheres by using the max BOLD beta value.

Group	**Condition**	**Hemisphere**	**N (%)**
Without ASD	**Collapsed**	Left	6 (35%)
		Right	8 (47%)
	Social-Linguistic	Left	6 (35%)
		Right	8 (47%)
	Social-Nonlinguistic	Left	6 (35%)
		Right	8 (47%)
	Mousetrap	Left	4 (23%)
		Right	10 (59%)
	Music	Left	7 (41%)
		Right	7 (41%)
	Total		14 (82%)
ASD	**Collapsed**	Left	8 (53%)
		Right	5 (33%)
	Social-Linguistic	Left	5 (33%)
		Right	8 (53%)
	Social-Nonlinguistic	Left	8 (53%)
		Right	5 (33%)
	Mousetrap	Left	7 (47%)
		Right	6 (40%)
	Music	Left	5 (33%)
		Right	8 (53%)
	Total		13 (87%)

Table 11

Number and percent of participants with and without ASD that were classified as displaying temporal synchrony enhancement (synchronous beta value > asynchronous beta value) for each condition.

Condition	**Left Hemisphere**		**Right Hemisphere**	
	Without ASD N (%)	**ASD N (%)**	**Without ASD N (%)**	**ASD N (%)**
Social-Linguistic	10 (71%)	7 (54%)	11 (79%)	9 (69%)
Social-Nonlinguistic	10 (71%)	10 (77%)	9 (64%)	9 (69%)
Mousetrap	12 (86%)	7 (54%)	10 (71%)	7 (54%)
Music	4 (29%)	9 (69%)	6 (43%)	9 (69%)

Table 12 *Activation peaks for clusters that survived the threshold of + 2.58 for the PLS analyses of task-related brain activity for the synchronous and asynchronous mousetrap nonsocial-nonlinguistic, social-linguistic, and social-nonlinguistic content conditions in the group without ASD. Including the Laterality (L- Left; R- Right), Location, coordinates, Brodmann Area (BA), Brain Score (BSR), p-value, and cluster-size.*

Laterality	**Location**	**X**	**Y**	**Z**	**BA**	**BSR**	**P-Value**	**Cluster Size**
R	Middle Temporal Gyrus/ STS	58.0	-2.0	-4.0	21	15.3133	<.0001	4675
L	Middle Temporal Gyrus/ STS	-64.0	-30.0	4.0	22	10.8231	<.0001	3897
L	Cuneus	-2.0	-94.0	16.0	18	8.3170	<.0001	839
L	Culmen	-10.0	-44.0	-10.0	*	4.9206	<.0001	52
R	Lingual Gyrus	4.0	-82.0	-4.0	18	4.5372	<.0001	139
L	Superior Frontal Gyrus	-4.0	6.0	66.0	6	3.7888	0.0002	35
R	Inferior Frontal Gyrus (Broca)	56.0	20.0	16.0	45	3.7372	0.0002	97
R	Insula	36.0	4.0	12.0	13	3.6690	0.0002	35
R	Amygdala	24.0	-14.0	-20.0	28	3.5930	0.0003	52
R	Cerebellar Tonsil	6.0	-44.0	-38.0	*	3.5394	0.0004	56
R	Superior Frontal Gyrus	6.0	18.0	62.0	6	3.3060	0.0009	24
L	Inferior Frontal Gyrus (Broca)	-52.0	22.0	20.0	45	3.2149	0.0013	51
L	Middle Occipital Gyrus	-40.0	-80.0	12.0	19	-11.7758	<.0001	9456
L	Sub-Gyral	-24.0	0.0	54.0	6	-11.4502	<.0001	1652
R	Precuneus	32.0	-72.0	36.0	19	-11.4415	<.0001	7916
R	Middle Frontal Gyrus	28.0	4.0	50.0	6	-8.7424	<.0001	1285
L	Precentral Gyrus	-56.0	-2.0	36.0	6	-6.2214	<.0001	396
R	Subcallosal Gyrus	2.0	12.0	-12.0	25	-5.4708	<.0001	40
L	Rectal Gyrus	-10.0	12.0	-24.0	11	-4.3886	<.0001	133
R	Middle Frontal Gyrus	38.0	30.0	24.0	9	-4.2279	<.0001	270
R	Cingulate Gyrus	16.0	-30.0	42.0	31	-4.1494	<.0001	72
L	Middle Frontal Gyrus	-40.0	30.0	34.0	9	-4.0333	0.0001	144
L	Cingulate Gyrus	-12.0	-26.0	38.0	31	-3.7293	0.0002	38
R	Precuneus	16.0	-58.0	24.0	31	-3.5549	0.0004	84
R	Paracentral Lobule	6.0	-40.0	62.0	5	-3.1741	0.0015	27

Table 13

Activation peaks for clusters that survived the threshold of + 2.58 for the PLS task-related brain activity for the synchronous and asynchronous mousetrap nonsocial-nonlinguisti and social-linguistic, and conditions in the group without ASD. Including the Laterality (L- Left; R- Right), Location, coordinates, Brodmann Area (BA), Brain Score (BSR), p-value, and cluster-size.

Laterality	**Location**	**X**	**Y**	**Z**	**BA**	**BSR**	**P-Value**	**Cluster Size**
R	Middle Temporal Gyrus/ STS	58.0	-2.0	-4.0	21	15.3133	<.0001	4675
L	Middle Temporal Gyrus/ STS	-64.0	-30.0	4.0	22	10.8231	<.0001	3897
L	Cuneus	-2.0	-94.0	16.0	18	8.3170	<.0001	839
L	Culmen	-10.0	-44.0	-10.0	*	4.9206	<.0001	52
R	Lingual Gyrus	4.0	-82.0	-4.0	18	4.5372	<.0001	139
L	Superior Frontal Gyrus	-4.0	6.0	66.0	6	3.7888	0.0002	35
R	Inferior Frontal Gyrus (Broca)	56.0	20.0	16.0	45	3.7372	0.0002	97
R	Insula	36.0	4.0	12.0	13	3.6690	0.0002	35
R	Amygdala	24.0	-14.0	-20.0	28	3.5930	0.0003	52
R	Cerebellar Tonsil	6.0	-44.0	-38.0	*	3.5394	0.0004	56
R	Superior Frontal Gyrus	6.0	18.0	62.0	6	3.3060	0.0009	24
L	Inferior Frontal Gyrus (Broca)	-52.0	22.0	20.0	45	3.2149	0.0013	51
L	Middle Occipital Gyrus	-40.0	-80.0	12.0	19	-11.7758	<.0001	9456
L	Sub-Gyral	-24.0	0.0	54.0	6	-11.4502	<.0001	1652
R	Precuneus	32.0	-72.0	36.0	19	-11.4415	<.0001	7916
R	Middle Frontal Gyrus	28.0	4.0	50.0	6	-8.7424	<.0001	1285
L	Precentral Gyrus	-56.0	-2.0	36.0	6	-6.2214	<.0001	396
R	Subcallosal Gyrus	2.0	12.0	-12.0	25	-5.4708	<.0001	40
L	Rectal Gyrus	-10.0	12.0	-24.0	11	-4.3886	<.0001	133
R	Middle Frontal Gyrus	38.0	30.0	24.0	9	-4.2279	<.0001	270
R	Cingulate Gyrus	16.0	-30.0	42.0	31	-4.1494	<.0001	72
L	Middle Frontal Gyrus	-40.0	30.0	34.0	9	-4.0333	0.0001	144
L	Cingulate Gyrus	-12.0	-26.0	38.0	31	-3.7293	0.0002	38
R	Precuneus	16.0	-58.0	24.0	31	-3.5549	0.0004	84
R	Paracentral Lobule	6.0	-40.0	62.0	5	-3.1741	0.0015	27

Table 14 *Activation peaks for clusters that survived the threshold of + 2.58 for the PLS task-related brain activity for the PLS analyses of synchronous and asynchronous social-linguistic and social-nonlinguistic content conditions in the group without ASD. Including the Laterality (L-Left; R- Right), Location, coordinates, Brodmann Area (BA), Brain Score (BSR), p-value, and cluster-size.*

Laterality	Location	X	Y	Z	BA	BSR	P-Value	Cluster Size
R	Cingulate Gyrus	16.0	-26.0	34.0	31	5.1237	0.0000	28
L	Middle Occipital Gyrus	-50.0	-68.0	4.0	37	4.6088	0.0000	152
L	Superior Temporal Gyrus	-18.0	16.0	-30.0	38	4.5404	0.0000	24
L	Posterior Cingulate	-12.0	-38.0	18.0	29	4.4476	0.0000	24
R	Middle Frontal Gyrus	28.0	-2.0	38.0	6	4.1354	0.0000	65
L	Cingulate Gyrus	-26.0	-8.0	36.0	31	4.0874	0.0000	23
R	Culmen	38.0	-40.0	-24.0	*	4.0116	0.0001	33
R	Insula	50.0	-32.0	22.0	13	3.5442	0.0004	42
L	Insula	-26.0	-32.0	16.0	13	3.4797	0.0005	29
R	Culmen	6.0	-36.0	-6.0	*	3.3721	0.0007	24
L	Middle Temporal Gyrus/ STS	-60.0	-32.0	4.0	22	-11.5135	0.0000	3664
R	Superior Temporal Gyrus	58.0	-14.0	6.0	22	-9.9939	0.0000	3022
R	Parahippocampal Gyrus	20.0	-54.0	-2.0	19	-5.9506	0.0000	1947
L	Lingual Gyrus	-12.0	-76.0	-4.0	18	-5.1164	0.0000	335
L	Medial Frontal Gyrus	-6.0	-10.0	64.0	6	-4.8072	0.0000	238
R	Inferior Frontal Gyrus	30.0	26.0	-8.0	47	-4.7498	0.0000	94
L	Supramarginal Gyrus	-62.0	-50.0	22.0	40	-4.7449	0.0000	130
L	Postcentral Gyrus	-36.0	-18.0	44.0	4	-4.5085	0.0000	342
R	Inferior Frontal Gyrus	56.0	20.0	22.0	9	-4.3064	0.0000	44
L	Superior Temporal Gyrus	-28.0	18.0	-34.0	38	-3.9830	0.0001	60
R	Putamen	22.0	6.0	-4.0	*	-3.9607	0.0001	129
L	Fusiform Gyrus	-40.0	-10.0	-26.0	20	-3.9029	0.0001	49
L	Amygdala	-28.0	4.0	-30.0	*	-3.8190	0.0001	41
R	Posterior Cingulate	16.0	-62.0	12.0	30	-3.6925	0.0002	27
L	Superior Parietal Lobule	-32.0	-46.0	62.0	7	-3.6707	0.0002	30
R	Precentral Gyrus	44.0	-12.0	42.0	4	-3.6549	0.0003	40
L	Precentral Gyrus	-36.0	-2.0	40.0	6	-3.6506	0.0003	24
R	Middle Frontal Gyrus	30.0	40.0	24.0	10	-3.6380	0.0003	24
L	Insula	-36.0	0.0	14.0	13	-3.5167	0.0004	36
R	Cingulate Gyrus	4.0	0.0	44.0	24	-3.4230	0.0006	49
R	Sub-Gyral	24.0	-44.0	52.0	7	-3.3731	0.0007	64
L	Precentral Gyrus	-14.0	-26.0	66.0	4	-3.2839	0.0010	21
L	Precentral Gyrus	-24.0	-12.0	52.0	6	-3.2709	0.0011	55

Table 15

Activation peaks for clusters that survived the threshold of + 2.58 for the first latent variable for the PLS analyses of task-related brain activity for the synchronous and asynchronous mousetrap nonsocial-nonlinguistic, social-linguistic, and social-nonlinguistic content conditions in the both the participants with and without ASD. Including the Laterality (L- Left; R- Right), Location, coordinates, Brodmann Area (BA), Brain Score (BSR), p-value, and cluster-size.

Laterality	Location	X	Y	Z	BA	BSR	P-Value	Cluster Size
L	Superior Parietal Lobule	-24.0	-60.0	58.0	7	12.3830	<.0001	23103
L	Middle Frontal Gyrus	-26.0	-6.0	56.0	6	10.9846	<.0001	1953
R	Middle Frontal Gyrus	26.0	-4.0	62.0	6	8.6191	<.0001	1817
L	Inferior Frontal Gyrus	-48.0	6.0	32.0	9	7.0397	<.0001	622
L	Superior Frontal Gyrus	-8.0	64.0	-14.0	11	5.9856	<.0001	401
L	Paracentral Lobule	-12.0	-24.0	46.0	6	5.8765	<.0001	230
R	Pyramis (Cerebellum)	4.0	-72.0	-24.0	*	5.2002	<.0001	111
L	Inferior Frontal Gyrus	-18.0	10.0	-24.0	47	5.1398	<.0001	145
L	Middle Frontal Gyrus	-42.0	32.0	32.0	9	4.9249	<.0001	198
R	Middle Frontal Gyrus	38.0	32.0	26.0	9	4.8759	<.0001	401
L	Thalamus	-18.0	-30.0	10.0		4.5324	<.0001	164
R	Superior Frontal Gyrus	18.0	68.0	10.0	10	4.4808	<.0001	32
L	Middle Frontal Gyrus	-38.0	52.0	18.0	10	4.2598	<.0001	113
L	Superior Frontal Gyrus	-24.0	56.0	-12.0	11	4.2170	<.0001	342
R	Putamen	26.0	-16.0	16.0	*	3.8659	0.0001	138
R	Middle Frontal Gyrus	44.0	42.0	-14.0	11	3.8512	0.0001	88
L	Inferior Temporal Gyrus	-62.0	-16.0	-18.0	20	3.3677	0.0008	28
L	Cingulate Gyrus	0.0	2.0	32.0	24	3.3452	0.0008	30
R	Superior Frontal Gyrus	24.0	42.0	-16.0	11	3.3365	0.0008	25
R	Superior Frontal Gyrus	26.0	60.0	-2.0	10	3.2444	0.0012	45
R	Inferior Frontal Gyrus	46.0	4.0	26.0	9	3.2402	0.0012	71
R	Middle Frontal Gyrus	32.0	44.0	-6.0	11	3.0023	0.0027	23
R	Superior Temporal Gyrus	52.0	-16.0	4.0	22	-15.0848	<.0001	5449
L	Superior Temporal Gyrus	-58.0	-18.0	2.0	22	-13.9432	<.0001	3895
L	Cuneus	-2.0	-94.0	14.0	18	-4.6957	<.0001	272
R	Precuneus	20.0	-52.0	36.0	31	-4.2457	<.0001	23
R	Lingual Gyrus	6.0	-82.0	-4.0	18	-4.2307	<.0001	61
L	Posterior Cingulate	-10.0	-32.0	20.0	23	-3.7317	0.0002	97
R	Amygdala	30.0	-8.0	-18.0	*	-3.4184	0.0006	26
R	Cerebellum	2.0	-48.0	-26.0	*	-3.3411	0.0008	24
R	Cerebellum	20.0	-44.0	-28.0	*	-3.3347	0.0009	22

Table 16

Activation peaks for clusters that survived the threshold of + 2.58 for the second latent variable for the PLS analyses of task-related brain activity for the synchronous and asynchronous mousetrap nonsocial-nonlinguistic, social-linguistic, and social-nonlinguistic content conditions in participants with and without ASD. Including the Laterality (L- Left; R- Right), Location, coordinates, Brodmann Area (BA), Brain Score (BSR), p-value, and cluster-size.

Laterality	Location	X	Y	Z	BA	BSR	P-Value	Cluster Size
R	Middle Temporal Gyrus (STS)	58.0	0.0	-10.0	21	10.6633	<.0001	3115
L	Middle Temporal Gyrus (STS)	-56.0	-6.0	-6.0	21	9.6795	<.0001	5077
R	Middle Occipital Gyrus	10.0	-96.0	14.0	18	7.6648	<.0001	5861
L	Paracentral Lobule	-12.0	-32.0	48.0	5	6.3243	<.0001	3748
L	Middle Frontal Gyrus	-54.0	14.0	32.0	9	6.2707	<.0001	951
R	Putamen/Thalamus	28.0	-2.0	14.0	*	5.7712	<.0001	1644
L	Anterior Cingulate	-18.0	18.0	4.0	24	5.0248	<.0001	345
R	Postcentral Gyrus	26.0	-32.0	62.0	3	4.8953	<.0001	175
L	Medial Frontal Gyrus	-4.0	52.0	14.0	9	4.8924	<.0001	229
L	Superior Temporal Gyrus	-40.0	-54.0	24.0	39	4.7021	<.0001	93
R	Subcallosal Gyrus	18.0	6.0	-14.0	34	4.4903	<.0001	74
R	Medial Frontal Gyrus	6.0	68.0	8.0	10	4.3736	<.0001	58
L	Cingulate Gyrus	-10.0	28.0	26.0	32	4.1554	<.0001	327
R	Superior Frontal Gyrus	2.0	12.0	62.0	6	4.1356	<.0001	78
R	Amygdala	30	-10.0	-18	*	4.1281	<.0001	25
R	Inferior Frontal Gyrus	54.0	18.0	24.0	9	4.1083	<.0001	67
L	Superior Frontal Gyrus	-10.0	62.0	14.0	10	4.0737	<.0001	78
L	Hippocampus	-34.0	-32.0	-8.0	*	3.8485	0.0001	176
L	Culmen	-40.0	-44.0	-20.0	*	3.6597	0.0003	61
R	Parahippocampal Gyrus	18.0	-22.0	-14.0	35	3.5845	0.0003	24
L	Parahippocampal Gyrus	-20.0	2.0	-16.0	34	3.5132	0.0004	35
R	Inferior Parietal Lobule	44.0	-64.0	38.0	39	3.5080	0.0005	30
L	Middle Frontal Gyrus	-38.0	4.0	56.0	6	3.4512	0.0006	26
R	Postcentral Gyrus	36.0	-24.0	30.0	2	3.4495	0.0006	36
R	Cerebellar Tonsil	34.0	-38.0	-34.0	*	3.4296	0.0006	133
L	Posterior Cingulate	-16.0	-30.0	16.0	23	3.4134	0.0006	26
L	Superior Frontal Gyrus	-20.0	38.0	46.0	8	3.3215	0.0009	35
L	Superior Frontal Gyrus	-18.0	42.0	34.0	9	3.0936	0.0020	24
R	Inferior Parietal Lobule	42.0	-54.0	44.0	40	3.0911	0.0020	20
L	Inferior Temporal Gyrus	-56.0	-56.0	-14.0	20	3.0696	0.0021	21
R	Middle Temporal Gyrus	52.0	-62.0	24.0	39	2.8083	0.0050	25
L	Inferior Temporal Gyrus	-42.0	-66.0	-2.0	37	-4.0461	0.0001	125
R	Superior Temporal Gyrus	60.0	-34.0	20.0	42	-3.5819	0.0003	86
L	Middle Occipital Gyrus	-36.0	-82.0	2.0	18	-3.2553	0.0011	30
L	Precuneus	-12.0	-76.0	44.0	7	-3.1956	0.0014	33

Table 17

Activation peaks for clusters that survived the threshold of + 2.58 for the third latent variable for the PLS analyses of task-related brain activity for the synchronous and asynchronous mousetrap nonsocial-nonlinguistic, social-linguistic, and social-nonlinguistic content conditions in participants with and without ASD. Including the Laterality (L- Left; R- Right), Location, coordinates, Brodmann Area

Laterality	Location	X	Y	Z	BA	BSR	P-Value	Cluster Size
R	Posterior Cingulate	14.0	-22.0	28.0	29	4.4776	<.0001	34
R	Mammillary Body	2.0	-10.0	-6.0	*	3.3308	<.0001	47
L	Cingulate Gyrus	-12.0	2.0	46.0	24	-5.3694	<.0001	2669
L	Posterior Cingulate	-8.0	-42.0	20.0	29	-5.1904	<.0001	220
R	Precuneus	36.0	-66.0	30.0	39	-5.0524	<.0001	386
R	Superior Temporal Gyrus	48	-44.0	10.0	22	-4.9224	<.0001	180
L	Superior Parietal Lobule	-16.0	-52.0	60.0	7	-4.7831	<.0001	327
L	Superior Temporal Gyrus	-56.0	-12.0	0.0	21	-4.7504	<.0001	774
R	Insula	36.0	-18.0	16.0	13	-4.5836	<.0001	272
R	Postcentral Gyrus	42.0	-28.0	30.0	2	-4.5428	<.0001	60
L	Fusiform Gyrus	-38.0	-20.0	-26.0	20	-4.5070	<.0001	58
R	Superior Temporal Gyrus	64.0	-8.0	2.0	22	-4.1175	<.0001	470
R	Postcentral Gyrus	46.0	-18.0	54.0	3	-4.0688	<.0001	40
R	Middle Frontal Gyrus	42.0	14.0	26.0	9	-4.0542	0.0001	46
R	Lateral Globus Pallidus	16.0	2.0	-2.0	*	-4.0029	0.0001	84
R	Posterior Cingulate	8.0	-44.0	8.0	29	-3.9741	0.0001	49
R	Precuneus	2.0	-50.0	62.0	7	-3.9597	0.0001	130
R	Middle Frontal Gyrus	28.0	22.0	34.0	9	-3.9539	0.0001	84
L	Precuneus	-4.0	-74.0	44.0	7	-3.8497	0.0001	500
L	Declive	-32.0	-56.0	-18.0	*	-3.8285	0.0001	74
R	Posterior Cingulate	8.0	-38.0	26.0	23	-3.8029	0.0001	141
R	Inferior Frontal Gyrus	32.0	24.0	-4.0	47	-3.7464	0.0002	82
R	Inferior Frontal Gyrus	44.0	32.0	6.0	46	-3.6314	0.0003	29
R	Superior Temporal Gyrus	50.0	20.0	-16.0	38	-3.5950	0.0003	68
L	Anterior Cingulate	-2.0	24.0	12.0	24	-3.5597	0.0004	54
L	Superior Parietal Lobule	-30.0	-70.0	48.0	7	-3.5119	0.0004	67
L	Superior Temporal Gyrus	-46.0	14.0	-10.0	38	-3.3846	0.0007	72
R	Precentral Gyrus	36.0	-22.0	60.0	4	-3.3803	0.0007	28
L	Precentral Gyrus	-38.0	-16.0	34.0	4	-3.3661	0.0008	30
L	Superior Temporal Gyrus	-32.0	6.0	-22.0	38	-3.3401	0.0008	25
L	Insula	-38.0	18.0	2.0	13	-3.2662	0.0011	84
R	Paracentral Lobule	6.0	-40.0	48.0	5	-3.2563	0.0011	52
R	Middle Frontal Gyrus	30.0	10.0	40.0	6	-3.2484	0.0012	26
R	Precuneus	10.0	-68.0	50.0	7	-3.2230	0.0013	31
L	Middle Temporal Gyrus	-40.0	-70.0	24.0	39	-3.1777	0.0015	73
R	Paracentral Lobule	4.0	-28.0	70.0	6	-3.1432	0.0017	21
L	Putamen	-22.0	-2.0	6.0	*	-3.1160	0.0018	26

Appendix B: Figures

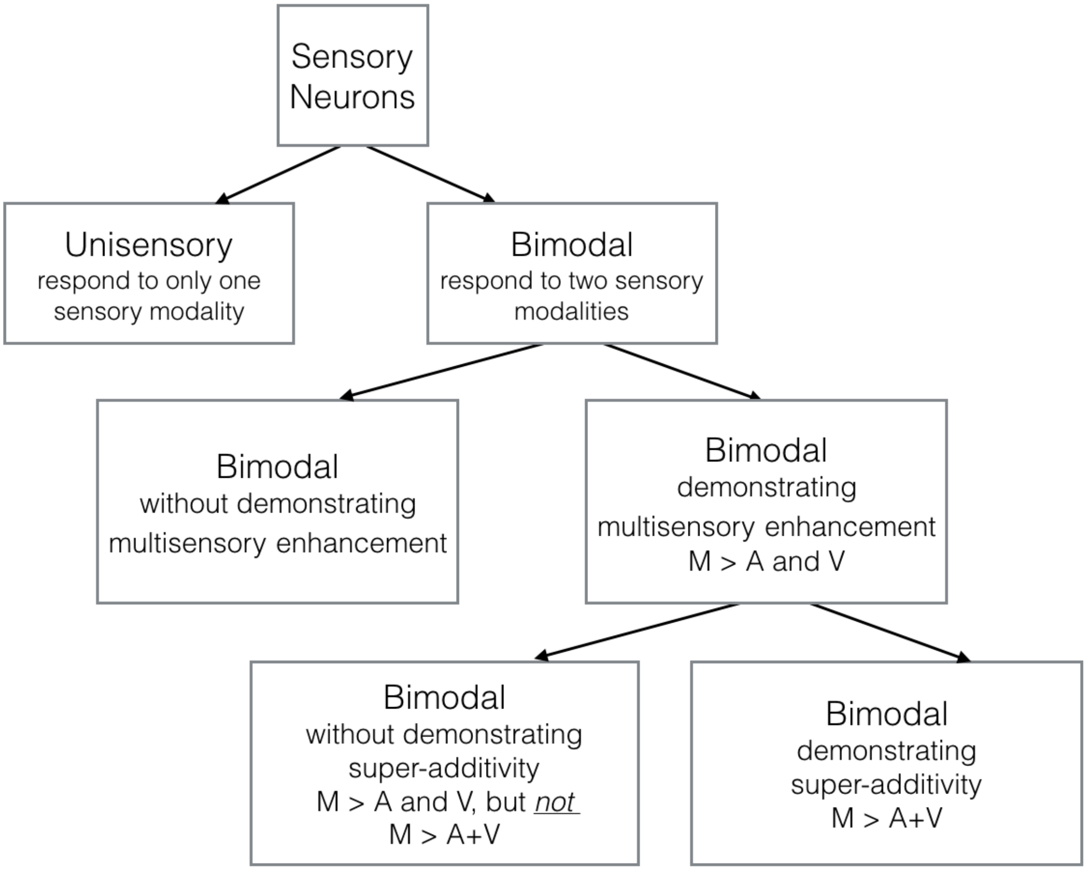
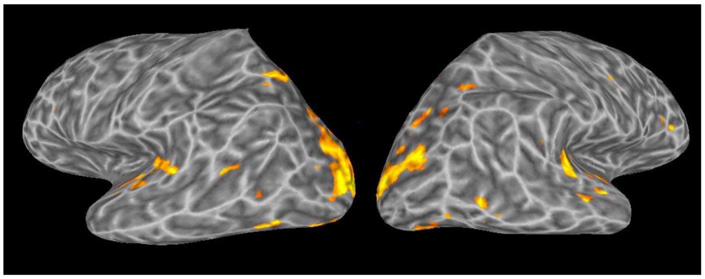
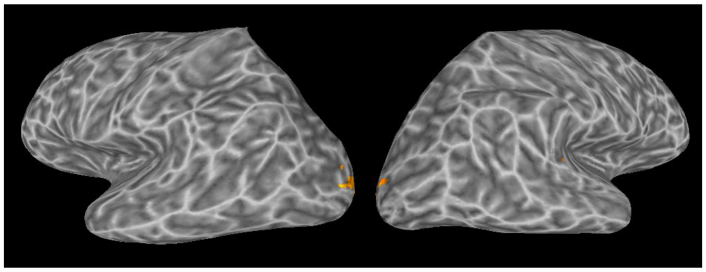


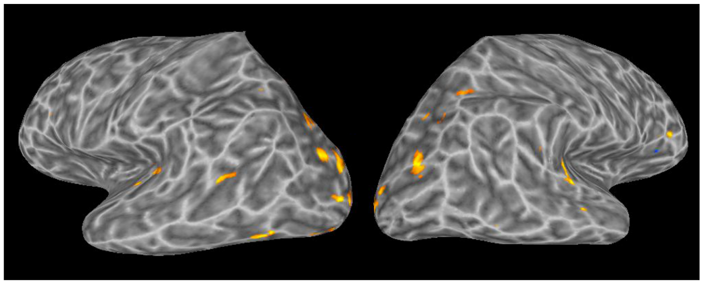
Figure 1. Types of multisensory neurons that are proposed to be present in the human and animal brain, with M representing multisensory stimuli, A representing auditory stimuli, and visual representing visual stimuli.



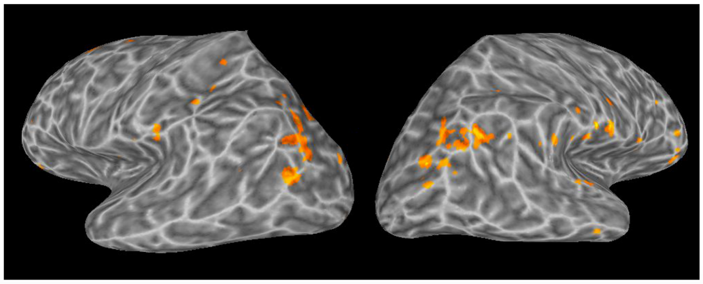
Mean



Max



Super-Additive



Temporal Synchrony

Figure 2. The brain images show the activation patterns for each of the methods in one of the participants. The activations are all threshold at a value of $p = .0001$ in order to demonstrate the small clusters of activation in each method.

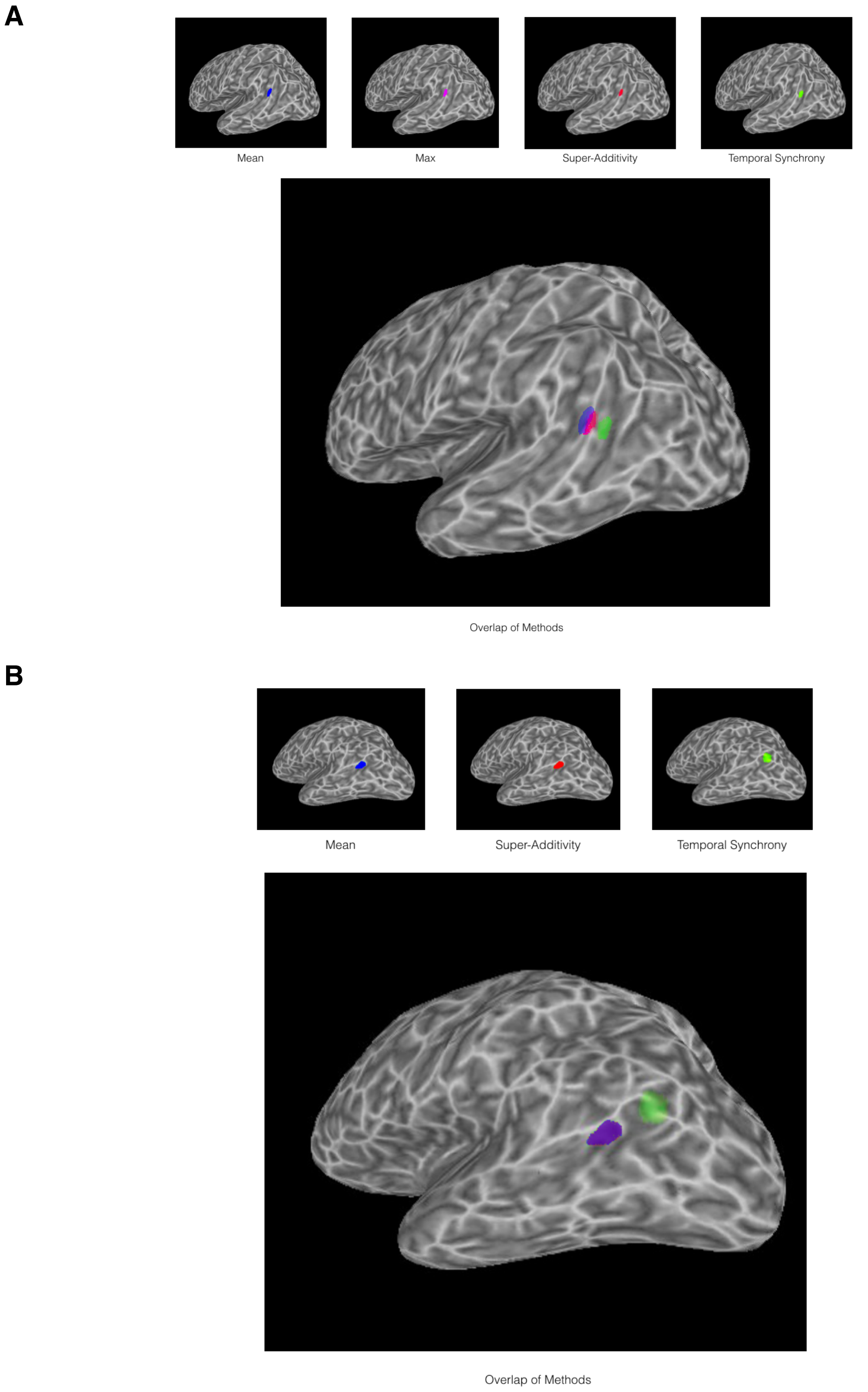


Figure 3. Display of ROIs for each method created for two different participants (A and B) in the left hemisphere. The second participant (B) did not have a Max ROI.

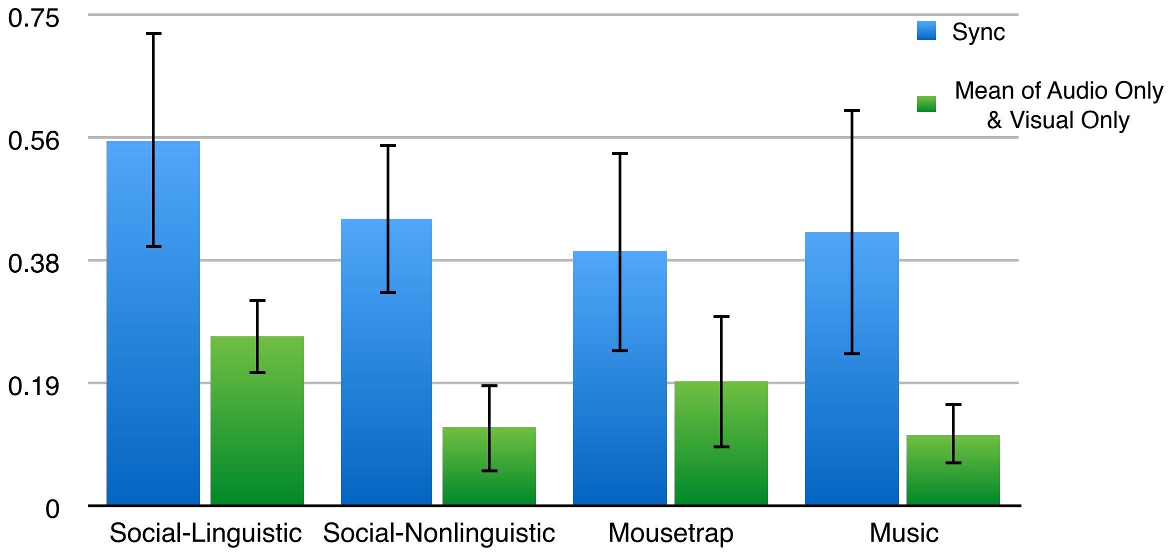


Figure 4. Mean beta values for the ROIs created with the Mean method at the collapsed level in the left hemisphere per content condition. Error bars represent the standard error of the mean.

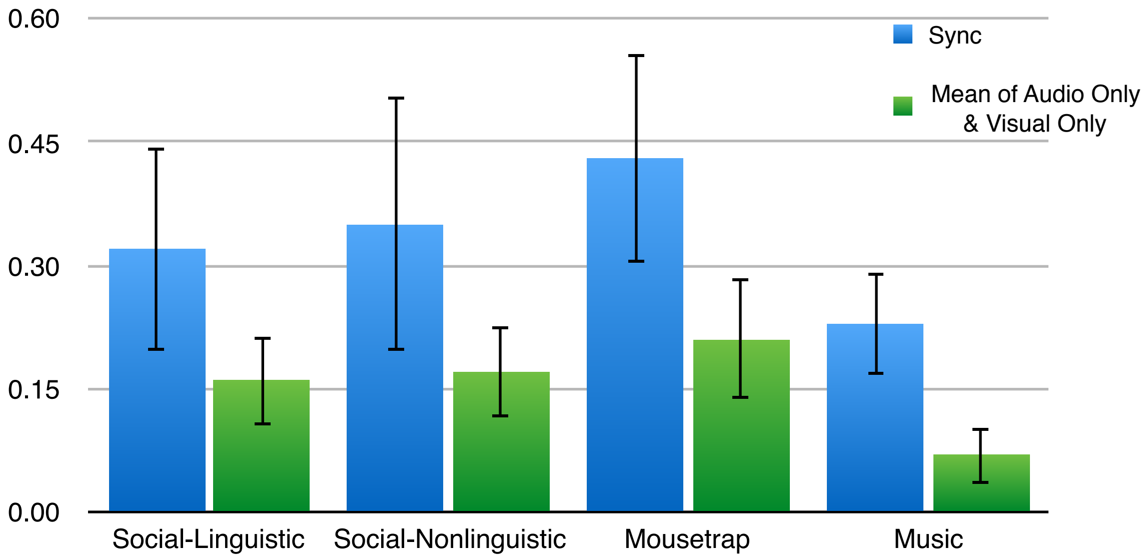


Figure 5. Mean beta values for the ROIs created with the Mean method at the collapsed level in the right hemisphere per content condition. Error bars represent the standard error of the mean.

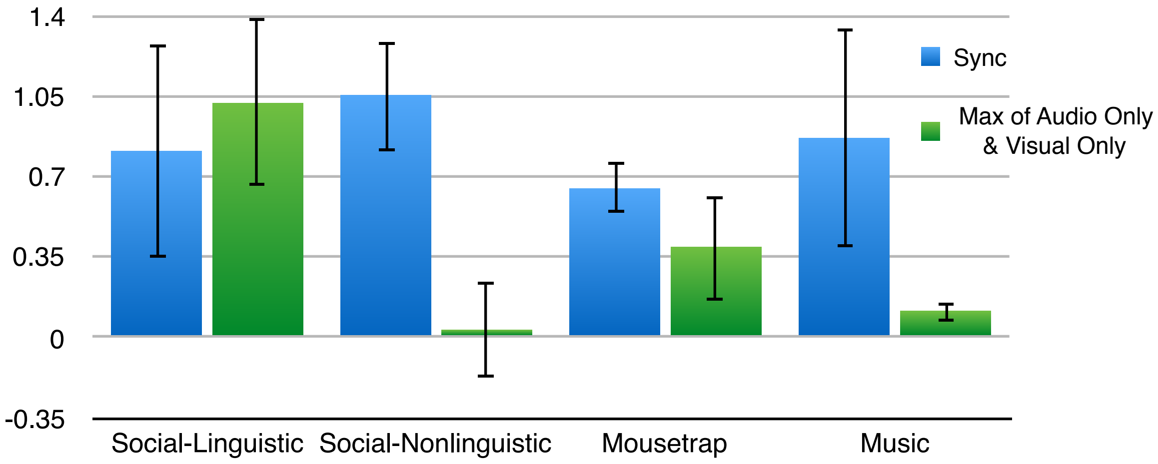


Figure 6. Mean beta values for the ROIs created with the Max method at the collapsed level in the left hemisphere per content condition. Error bars represent the standard error of the mean.

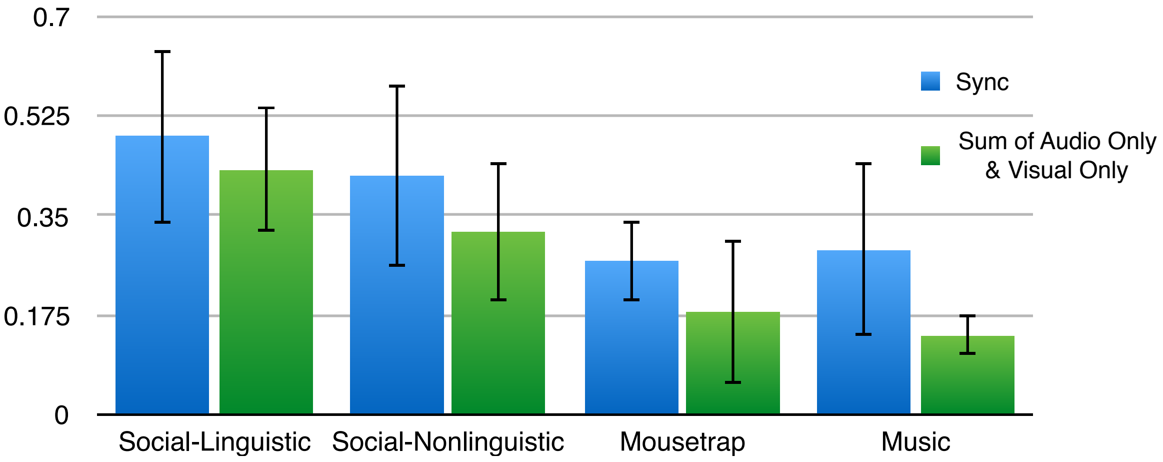


Figure 7. Mean beta values for the ROIs created with the Super-Additive method at the collapsed level in the left hemisphere per content condition. Error bars represent the standard error of the mean.

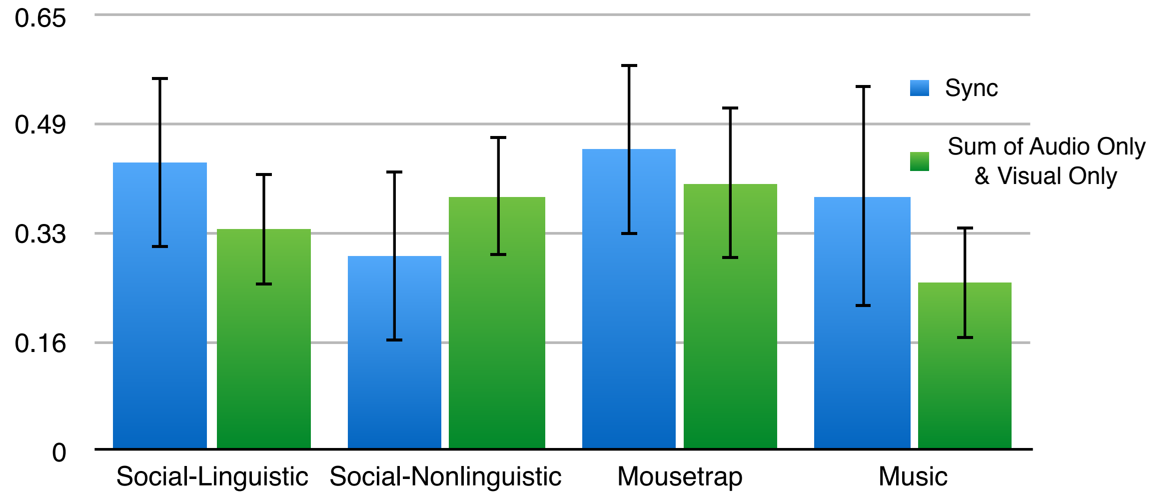


Figure 8. Mean beta values for the ROIs created with the Super-Additive method at the collapsed level in the right hemisphere per content condition. Error bars represent the standard error of the mean.

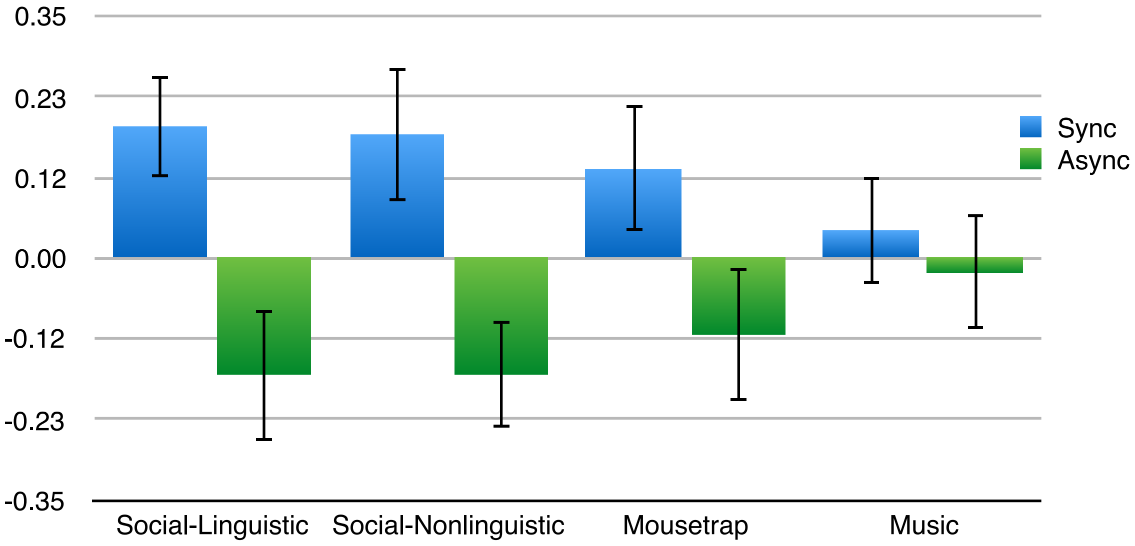


Figure 9. Mean beta values for the ROIs created with the Temporal Synchrony method at the collapsed level in the left hemisphere per content condition. Error bars represent the standard error of the mean.

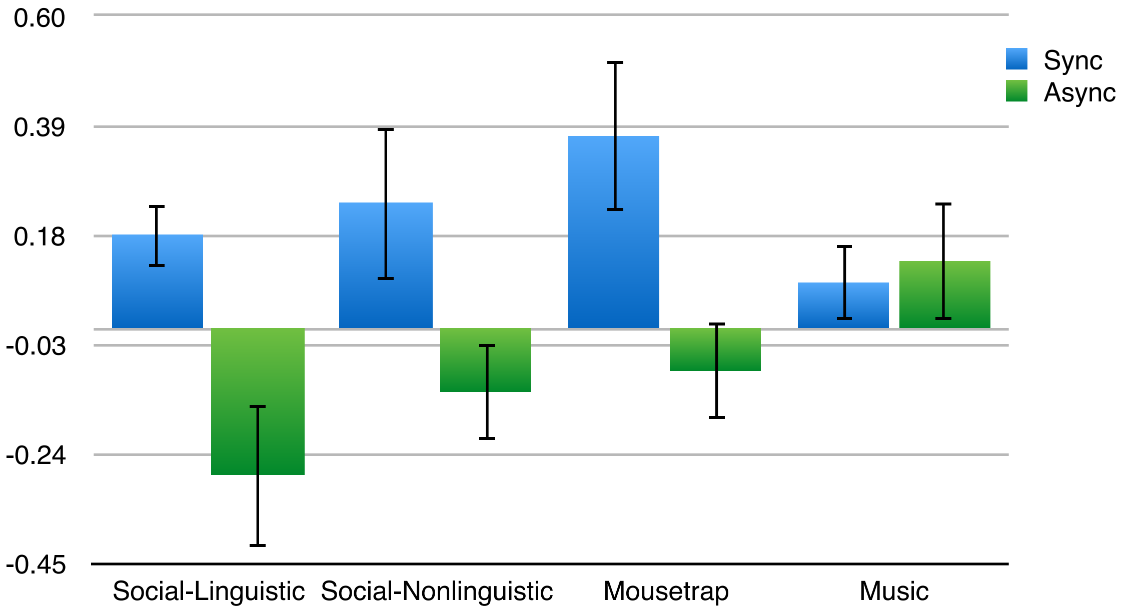


Figure 10. Mean beta values for the ROIs created with the Temporal Synchrony method at the collapsed level in the right hemisphere per content condition. Error bars represent the standard error of the mean.

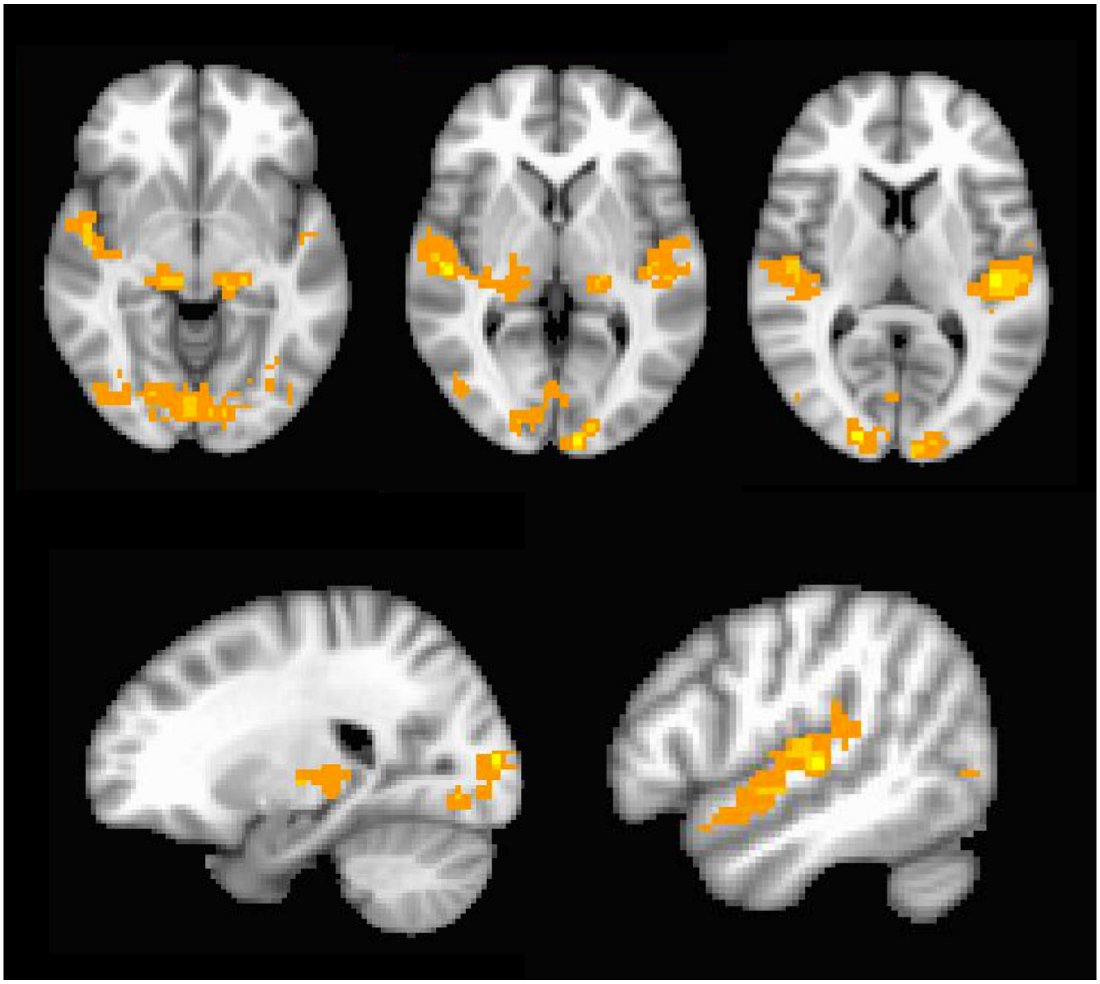


Figure 11. Group-level whole-brain analyses for the Mean method. The figure shows clusters where activation for synchronous audiovisual stimuli is greater than activation for the mean of the audio only and visual only stimuli (3cClustSim, AFNI, p <0.01, $\alpha < 0.05$, p<0.05, cluster size > 20). From left to right the axial slices (top row) were taken at z = -6, z = 2, z = 10 and sagittal slices (bottom row) were taken at x = -22 and x = -52.

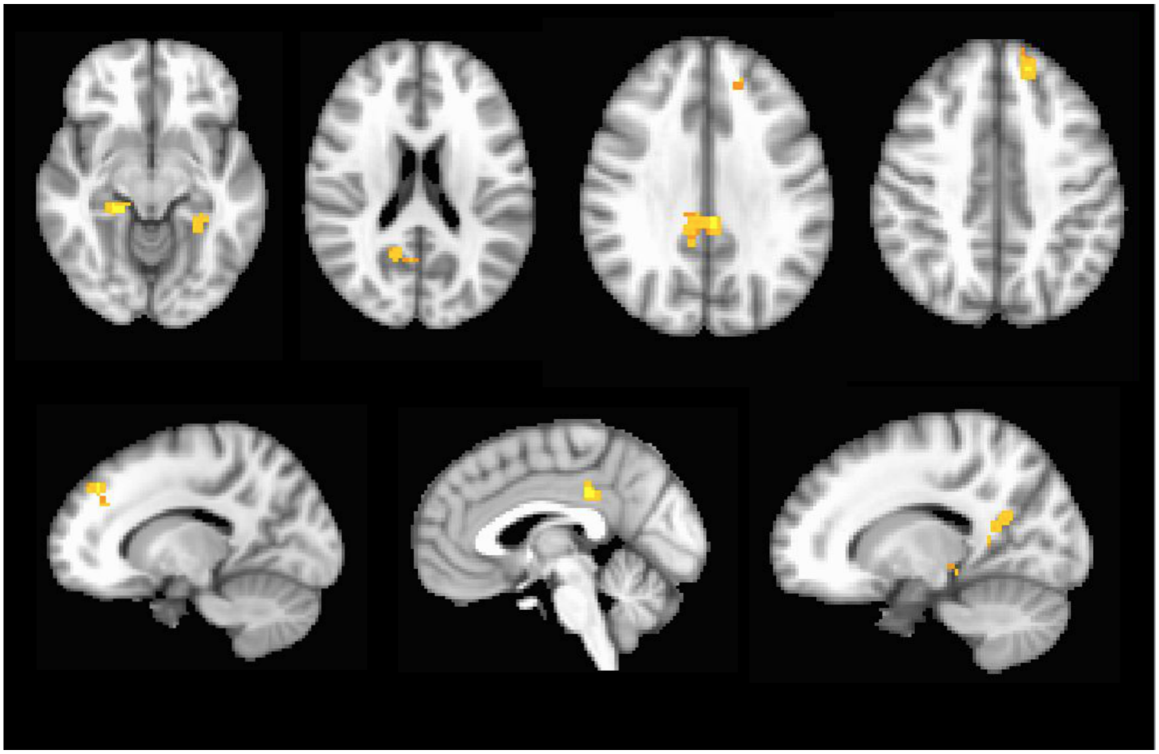


Figure 12. Group-level whole-brain analyses for the Temporal Synchrony method. The figure shows clusters where activation for the the synchronous audiovisual stimuli is greater than activation for the asynchronous audiovisual stimuli (3cClustSim, AFNI, $p < 0.01$, $\alpha < 0.05$, $p<0.05$, cluster size > 20). From left to right the axial slices (top row) were taken a z = -10, z = 20, z = 32, z = 42, and the sagittal slices (bottom row) were taken at x = 14, x = 2, x = - 16.

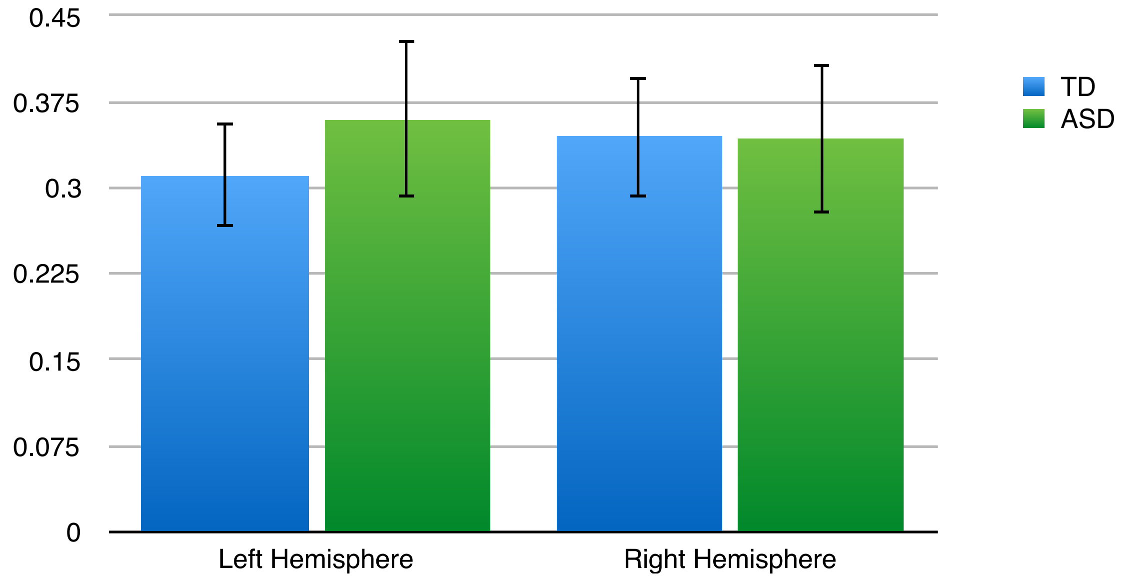


Figure 13. Mean beta values for the Temporal Synchrony (synchronous AV - asynchronous AV) ROIs in both hemispheres collapsed across content conditions for both groups. Error bars represent standard error of the mean.

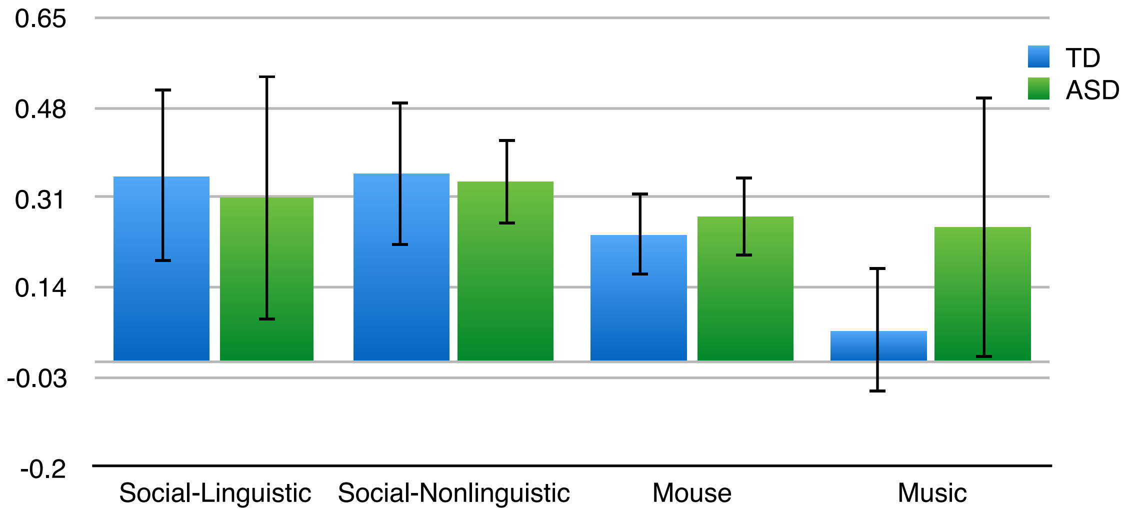


Figure 14. Mean beta values for the Temporal Synchrony (synchronous AV - asynchronous AV) ROIs in the left hemisphere for each content condition for both groups. Error bars represent standard error of the mean.

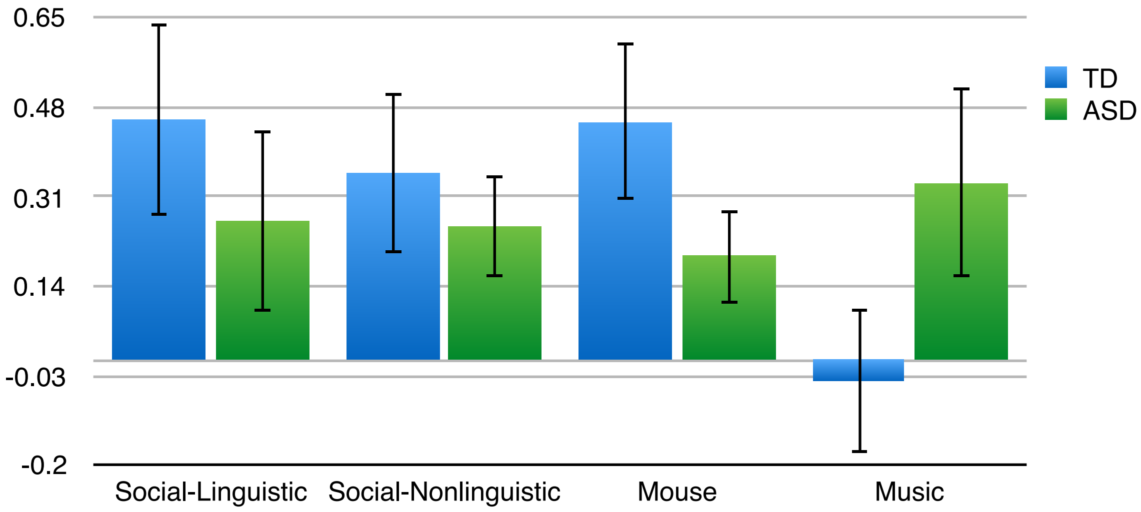


Figure 15. Mean beta values for the Temporal Synchrony (synchronous AV - asynchronous AV) ROIs in the right hemisphere for each content condition for both groups. Error bars represent standard error of the mean.

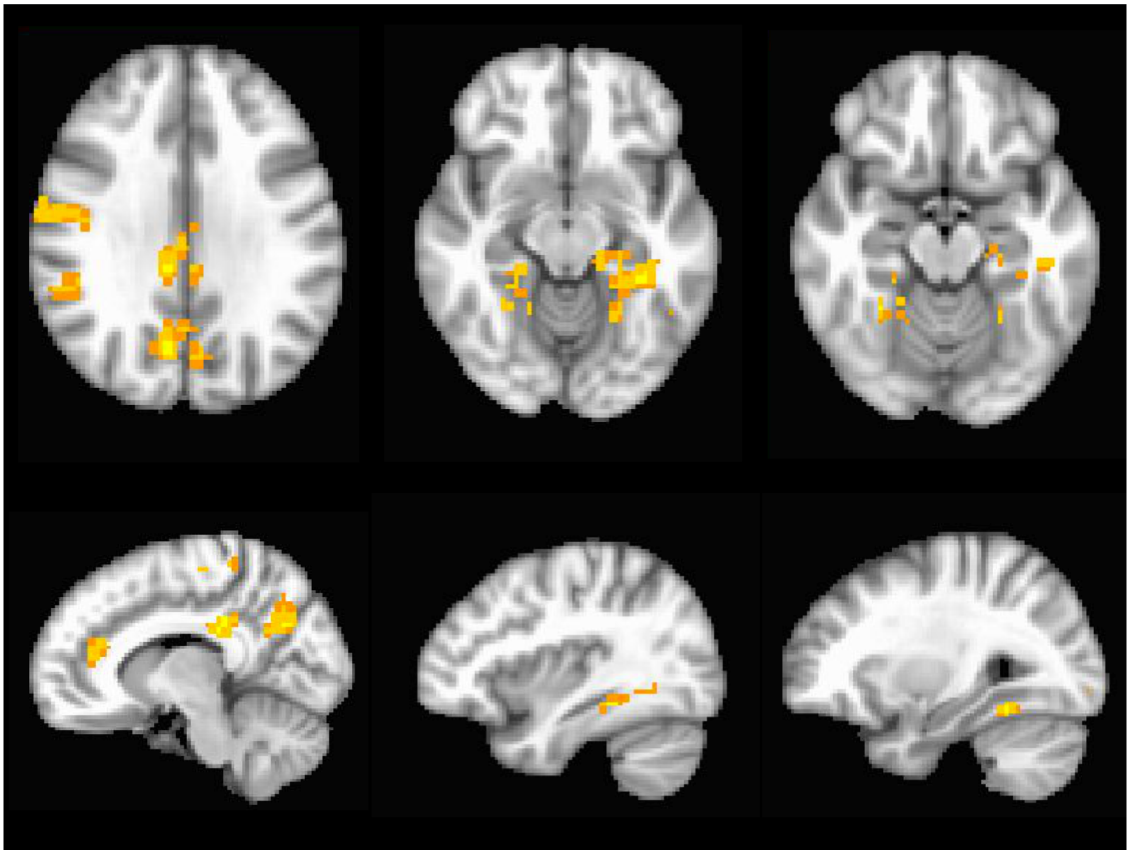


Figure 16. Clusters where termporal synchrony enhancement (AV Sync > AV Async) was greater for the group without ASD than the group with ASD (3cClustSim, AFNI, $p < 0.01$, $\alpha < 0.05$, $p < 0.05$, cluster size > 20). From left to right the axial slices (top row) were taken at z = 30, z = -12, and z = -16, and the sagittal slices (bottom row) were taken at x = -10, x = 50, and x = -30. Notable clusters that survived cluster correction displayed here include the bilateral posterior and anterior cingulate, bilateral parahippocampal gyrus, left precuneus, bilateral percentral gyrus, and left fusiform gyrus.

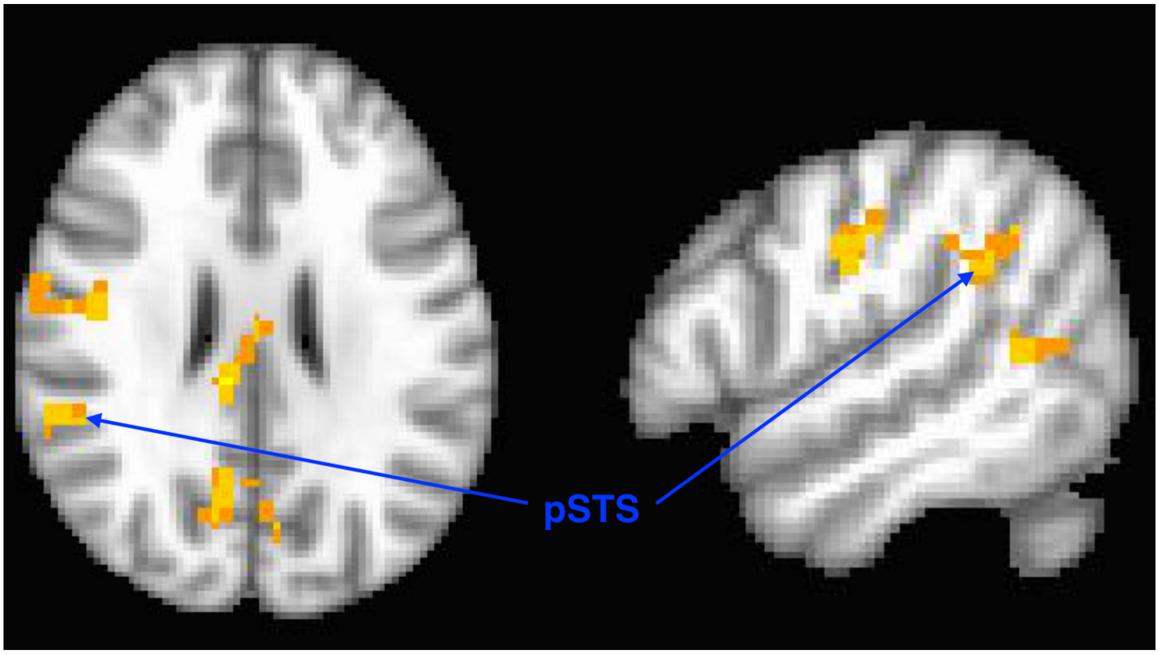


Figure 17. Clusters where temporal synchrony enhancement (AV Sync > AV Async) was greater for group without ASD than the group with ASD (3cClustSim, AFNI, $p < 0.01$, $\alpha < 0.05$, $p < 0.05$, cluster size > 20), highlighting the pSTS cluster. Axial slice (left) was taken at z = 26 and the sagittal slice (right) was taken at x = -54. Also observable in this figure are clusters in the left precentral gyrus, left middle temporal sulcus, and bilateral posterior cingulate.

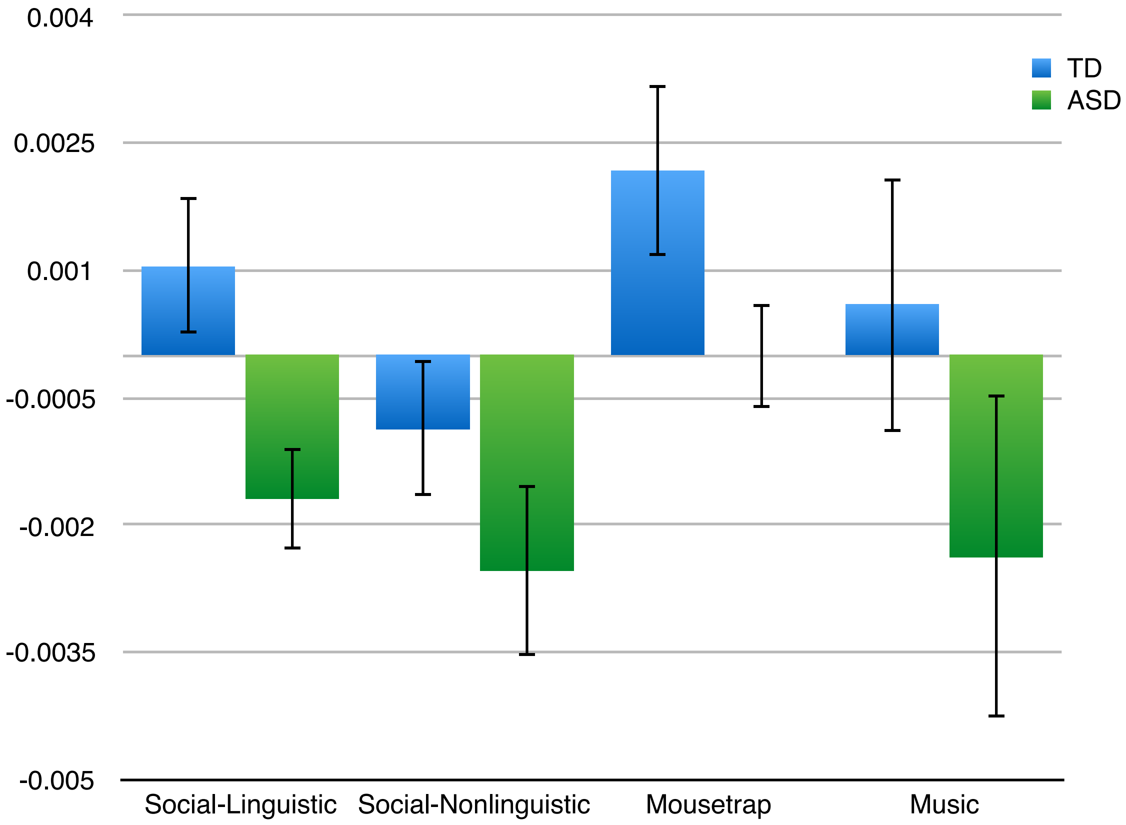


Figure 18. Mean beta values for the left pSTS ROIs for each content condition. Error bars represent standard error of the mean.

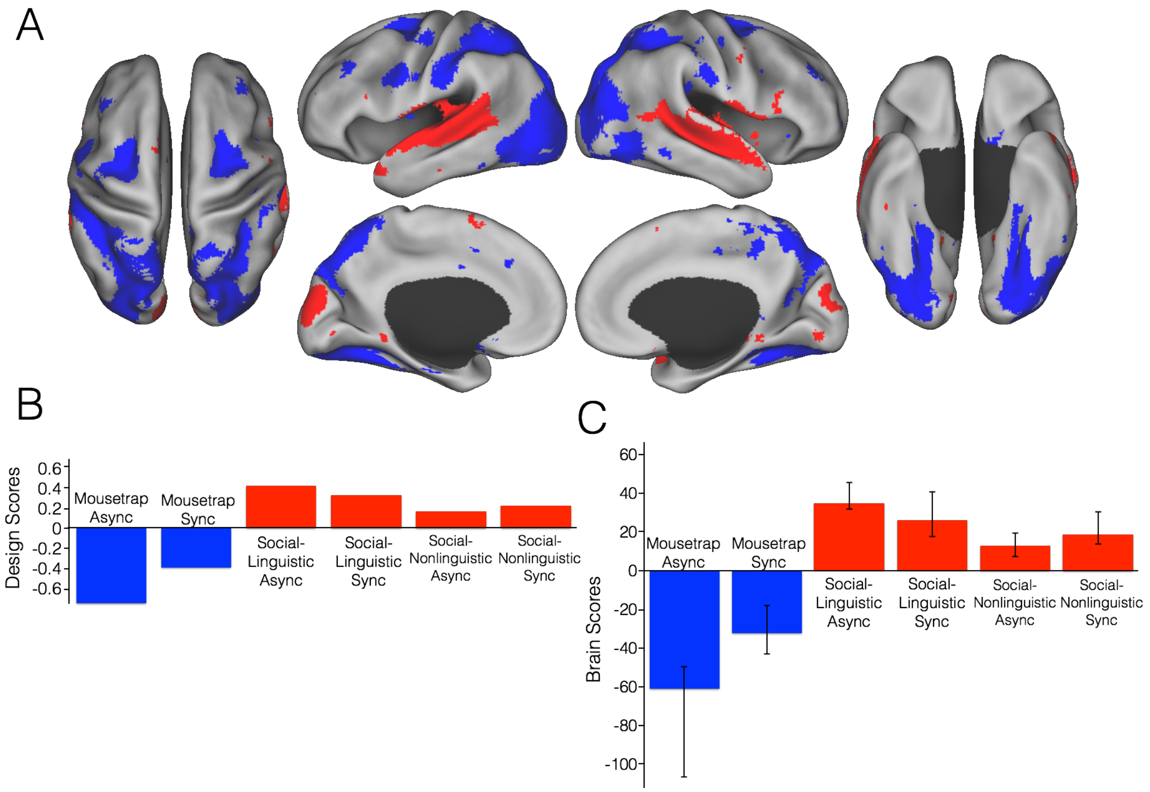


Figure 19. Task-related brain activity for the synchronous and asynchronous mousetrap, social-linguistic, and social-nonlinguistic content conditions in the group without ASD. LV 1: accounted for 63% of covariance, $p < 0.001$ (A) Activity associated with the mousetrap condition (blue) and the social-linguistic/social-nonlinguistic conditions (red). Data are displayed on the dorsal, lateral, medial, and ventral surfaces of the left and right hemispheres of a partially inflated surface map using Caret software (Van Essen, 2005). (B) Design scores for each condition on the LV. Mousetrap processing was maximally dissociated from social-linguistic and social-nonlinguistic processing, regardless of synchrony. (C) Mean brain scores across participants in each condition. Error bars represent 95% confidence intervals. Statistically significant differences between conditions are indicated by a lack of overlap of the confidence intervals.

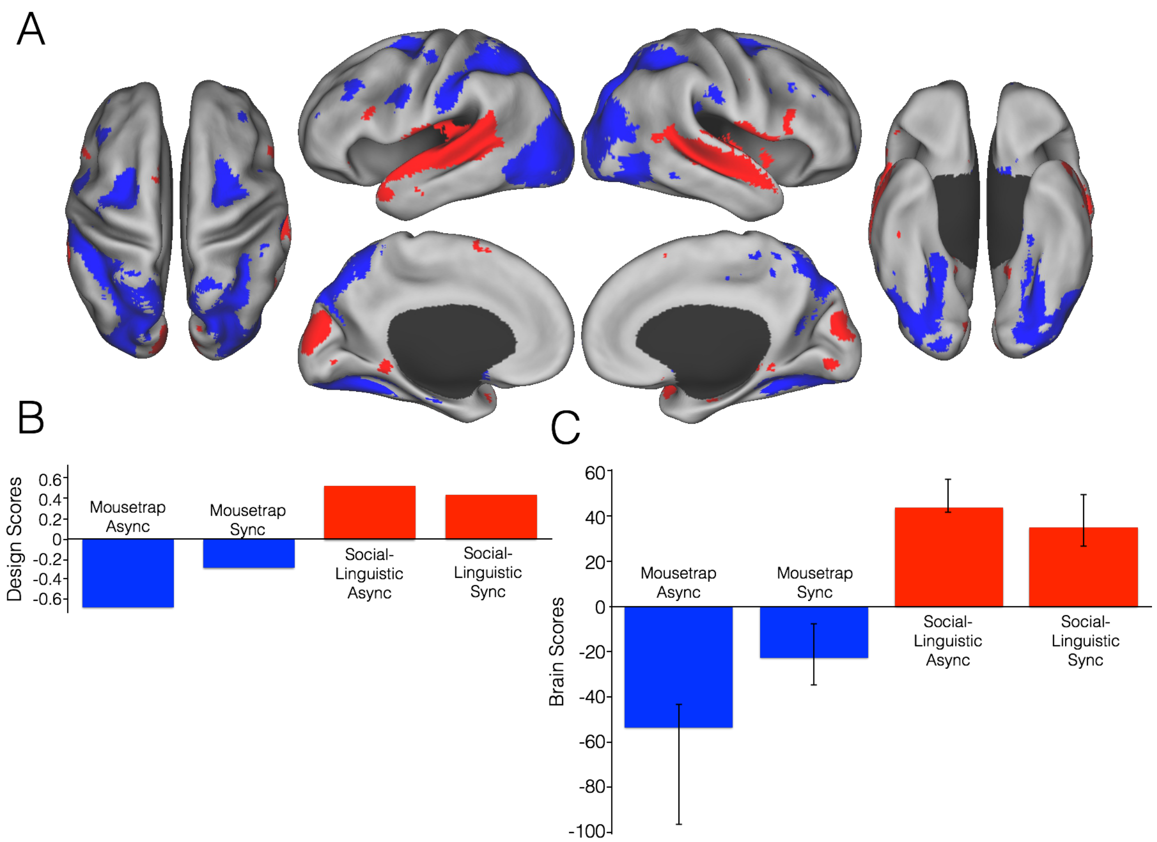


Figure 20. Task-related brain activity for the synchronous and asynchronous mousetrap and social-linguistic content conditions in the group without ASD. LV 1: accounted for 74% of the covariance, $p < 0.001$. (A) Activity associated with mousetrap condition (blue) and social-linguistic condition (red). Data are displayed on the dorsal, lateral, medial, and ventral surfaces of the left and right hemispheres of a partially inflated surface map using Caret software (Van Essen, 2005). (B) Design scores for each condition. Mousetrap processing was maximally dissociated from social-linguistic processing, regardless of synchrony. (C) Mean brain scores across participants in each condition. Error bars represent 95% confidence intervals. Statistically significant differences between conditions are indicated by a lack of overlap of the confidence intervals.

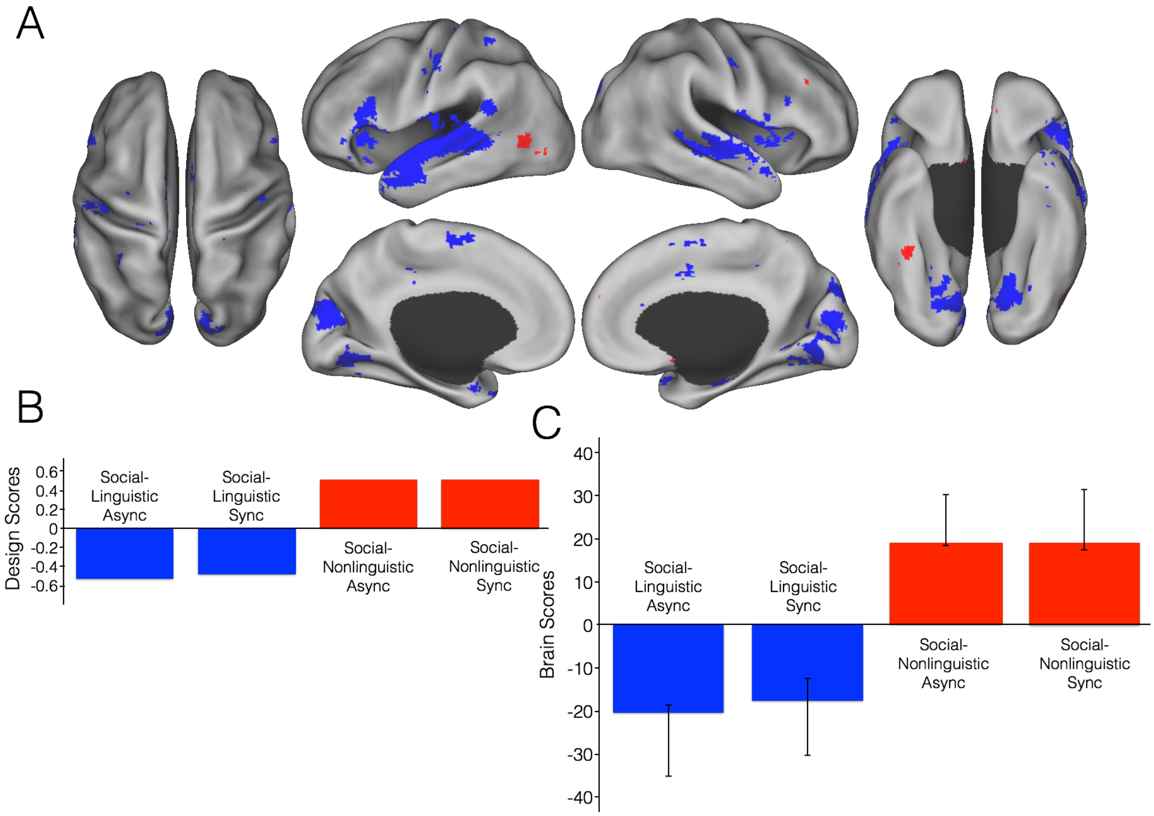


Figure 21. Task-related brain activity for the synchronous and asynchronous social-linguistic and social-nonlinguistic content conditions in the group without ASD. LV1: accounted for 43.32 % of the covariance, $p = .018$ (A) Activity associated with social-linguistic (blue) and social-nonlinguistic (red). Data are displayed on the dorsal, lateral, medial, and ventral surfaces of the left and right hemispheres of a partially inflated surface map using Caret software (Van Essen, 2005). (B) Design scores for each condition. Social-linguistic processing was maximally dissociated from social-nonlinguistic processing, regardless of synchrony. (C) Mean brain scores across participants in each condition. Error bars represent 95% confidence intervals. Statistically significant differences between conditions are indicated by a lack of overlap of the confidence intervals.

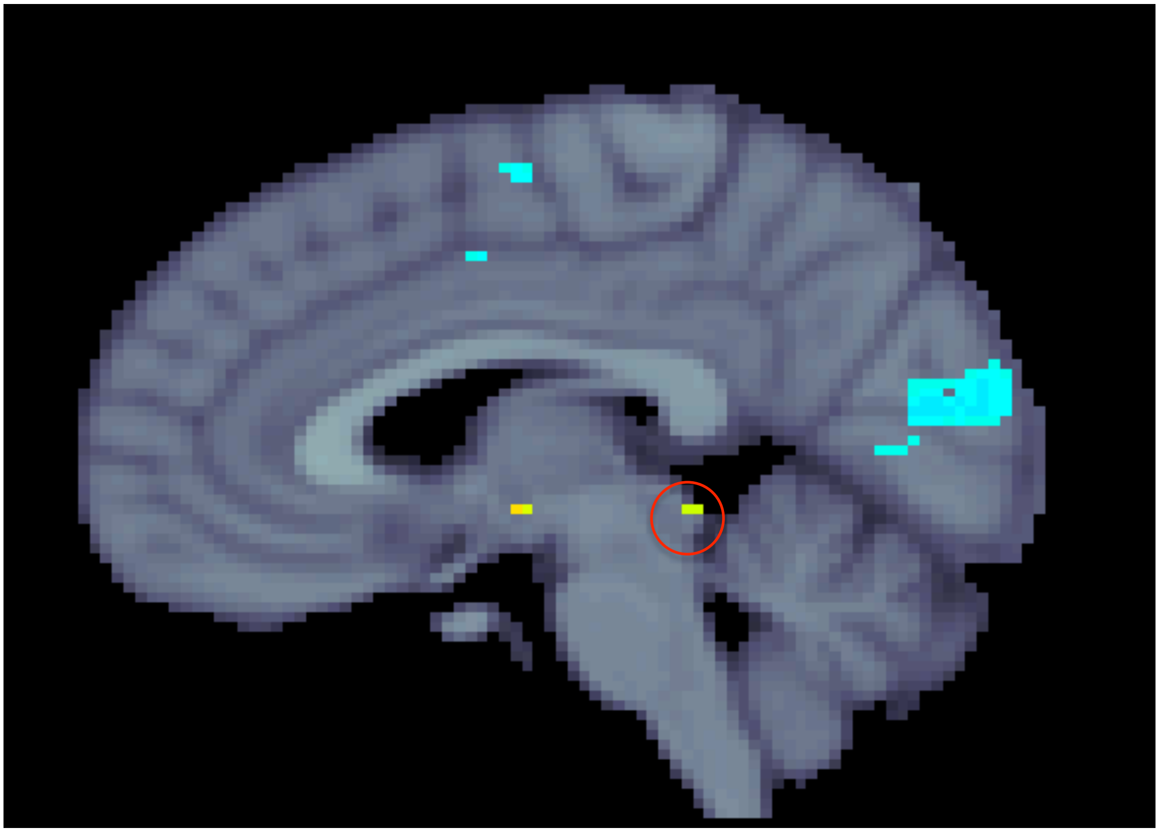


Figure 22. Task-related brain activity for the synchronous and asynchronous social-linguistic and social-nonlinguistic content conditions in the group without ASD. This figure shows clusters of activation associated with social-linguistic (blue) and social-nonlinguistic (yellow), $p < 0.01$, bootstrap threshold $\underline{+}$ 2.58. A notable cluster is the superior colliculus (circled) involvement in processing of the social-nonlinguistic information. Sagittal slice was taken at x = 4. MNI 152 space.

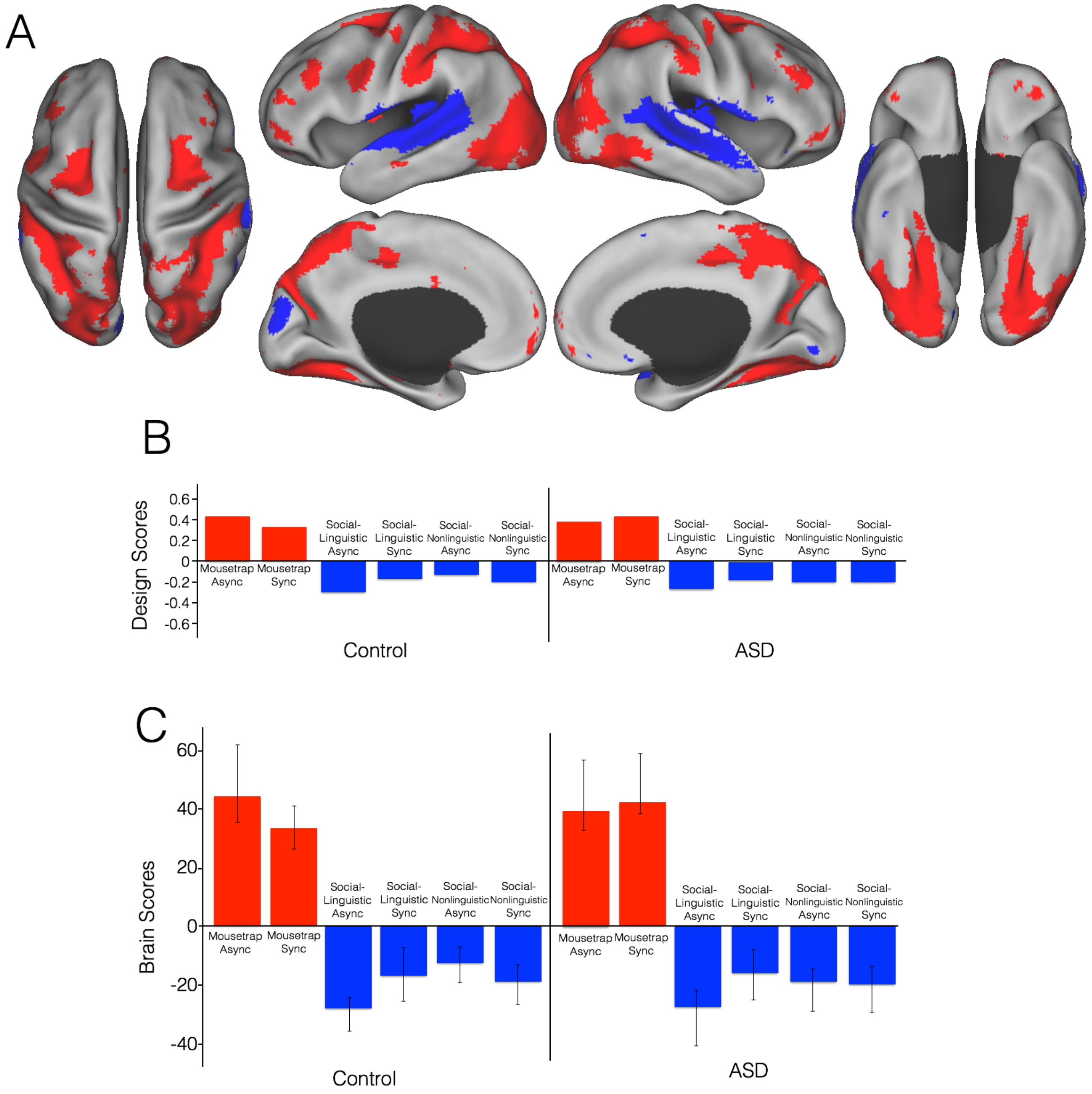


Figure 23. First latent variable for task-related brain activity for the synchronous and asynchronous mousetrap, social-linguistic, and social-nonlinguistic content conditions in the both the participants with and without ASD. LV1: accounted for 49.27% of the covariance in the model, $p < 0.001$ (A) Activity associated with mousetrap condition (red) and social-linguistic/social-nonlinguistic conditions (blue). Data are displayed on the dorsal, lateral, medial, and ventral surfaces of the left and right hemisphere of a partially inflated surface map using Caret software (Van Essen, 2005). (B) Design scores

for each condition. Mousetrap processing was maximally dissociated from social-linguistic and social-nonlinguistic processing, regardless of synchrony or group membership. (C) Mean brain scores across participants in each condition. Error bars represent 95% confidence intervals. Statistically significant differences between conditions are indicated by a lack of overlap of the confidence intervals.

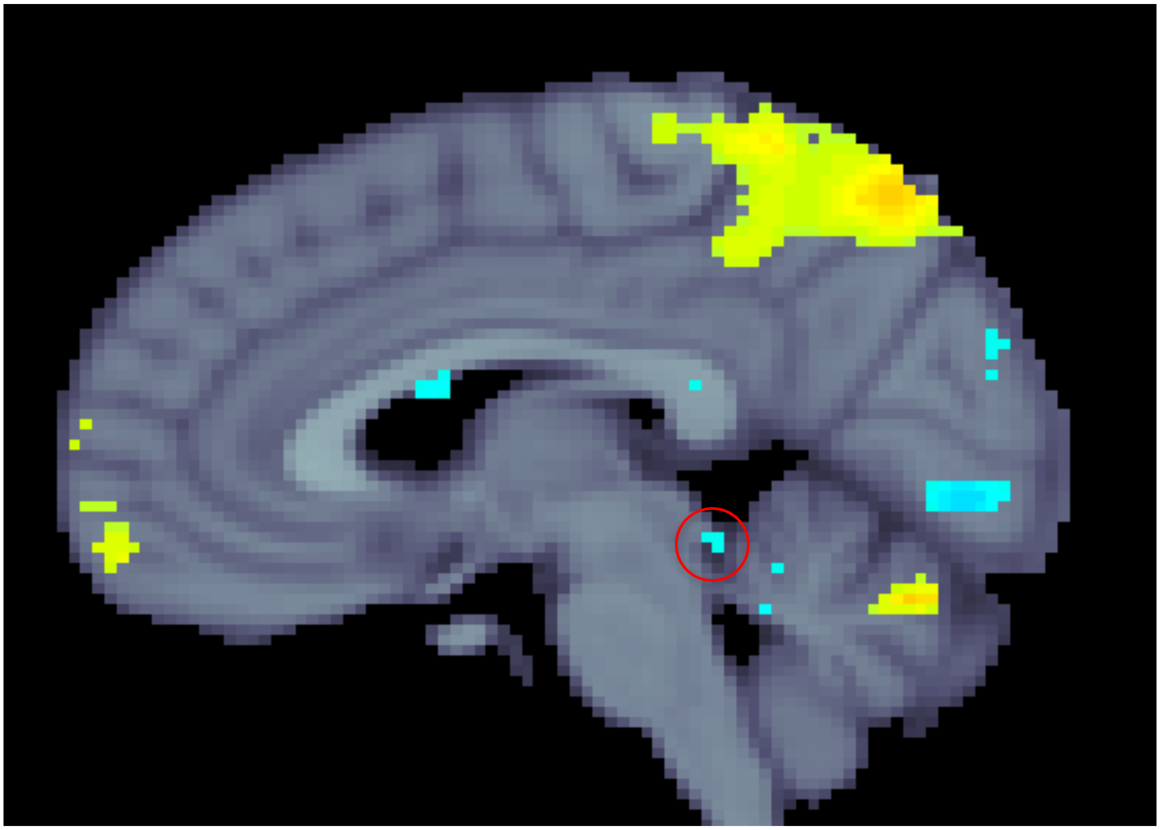


Figure 24. First latent variable for task-related brain activity for the synchronous and asynchronous mousetrap, social-linguistic, and social-nonlinguistic content conditions in the both the participants with and without ASD. A notable cluster is the superior colliculus' (circled) involvement in processing of the social information, regardless of linguistic or non-linguistic content. This figure shows clusters of activation associated with mousetrap (yellow) and social-linguistic (blue), $p < 0.01$, bootstrap threshold $\pm$ 2.58. Sagittal slice was taken at x = 4. Presented in MNI 152 space.

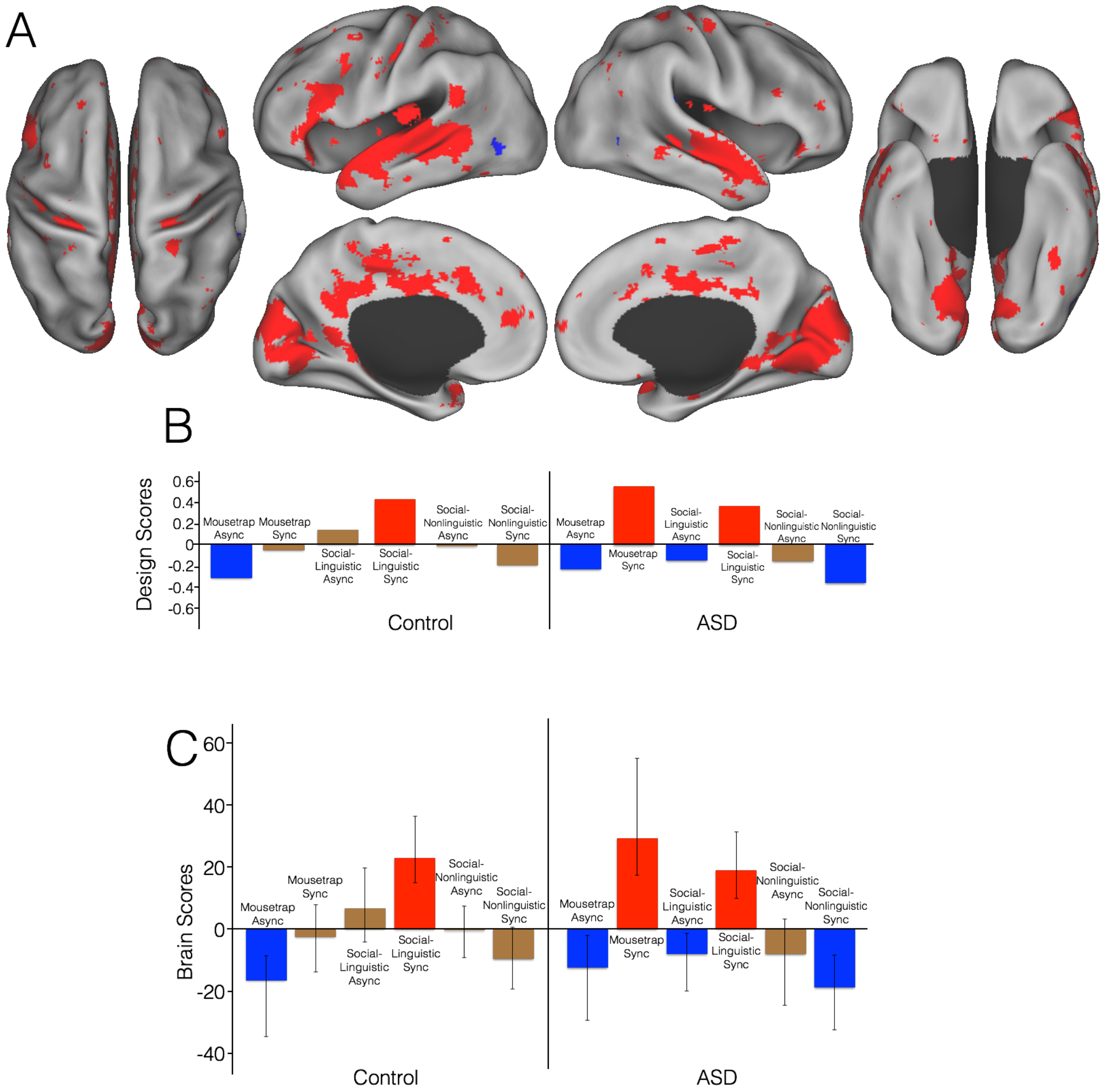


Figure 25. Second latent variable for task-related brain activity for the synchronous and asynchronous mousetrap, social-linguistic, and social-nonlinguistic content conditions in the participants with and without ASD. LV2: accounted for 14.35% of the covariance, $p <$ 0.001 (A) Activity associated with synchronous social-linguistic processing for the control group and both synchronous social-linguistic and mousetrap processing for the group with ASD is in red. Activity associated with asynchronous mousetrap processing for the control group and the asynchronous mousetrap, asynchronous social-linguistic,

and synchronous social-nonlinguistic processing in the group with ASD is in blue. Data are displayed on the dorsal, lateral, medial, and ventral surfaces of the left and right hemispheres of a partially inflated surface map using Caret software (Van Essen, 2005). (B) Design scores for each condition. In the group without ASD, the synchronous social-linguistic condition was maximally dissociated from asynchronous mousetrap condition. In the group with ASD, the synchronous social-linguistic and synchronous mousetrap conditions were maximally dissociated from asynchronous mousetrap, asynchronous social-linguistic, and synchronous social-linguistic conditions. Conditions that did not significantly contribute to the latent variable are in brown. (C) Mean brain scores across participants in each condition. Error bars represent 95% confidence intervals. Statistically significant differences between conditions are indicated by a lack of overlap of the confidence intervals.

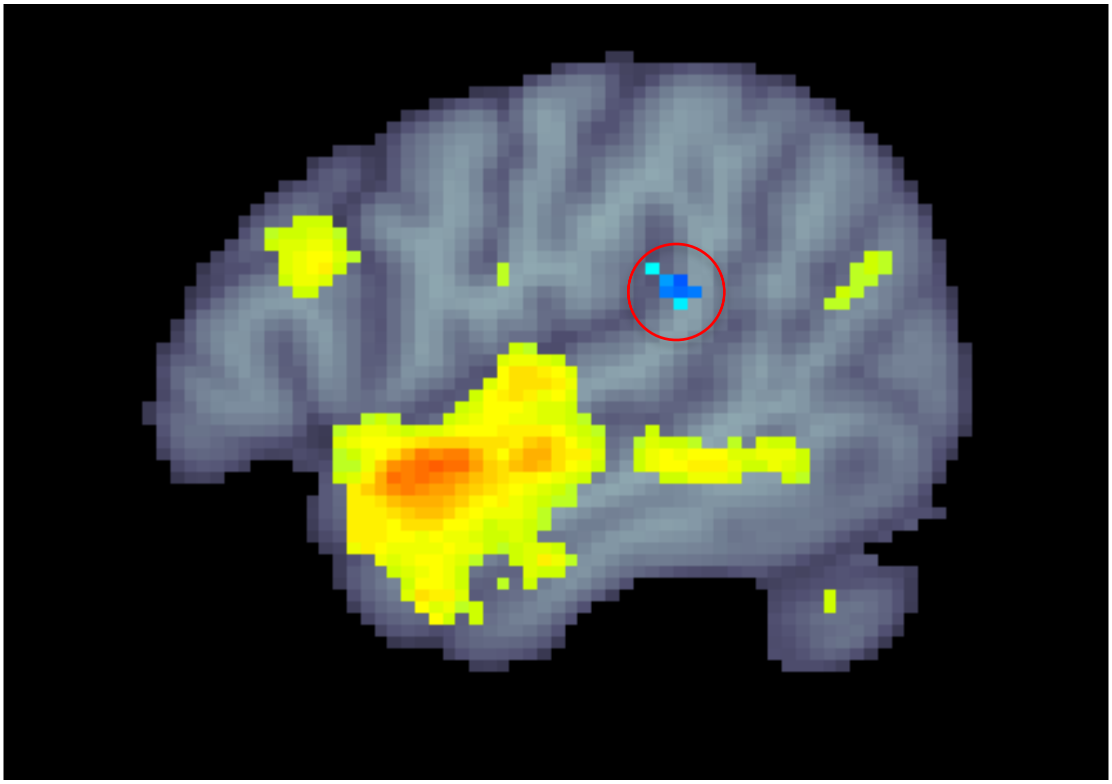


Figure 26. Second latent variable for task-related brain activity for the synchronous and asynchronous mousetrap, social-linguistic, and social-nonlinguistic content conditions in the participants with and without ASD. Activity associated with synchronous social-linguistic condition for the control group, and both synchronous social-linguistic and mousetrap conditions for the group with ASD, is in yellow. Activity associated with asynchronous mousetrap processing for the control group, and the asynchronous mousetrap, asynchronous social-linguistic, and synchronous social-nonlinguistic processing in the group with ASD is in blue, $p < 0.01$, bootstrap threshold ± 2.58. A notable cluster is the posterior STG (circled). Sagittal slice was taken at x = 54 (right hemisphere). MNI 152.

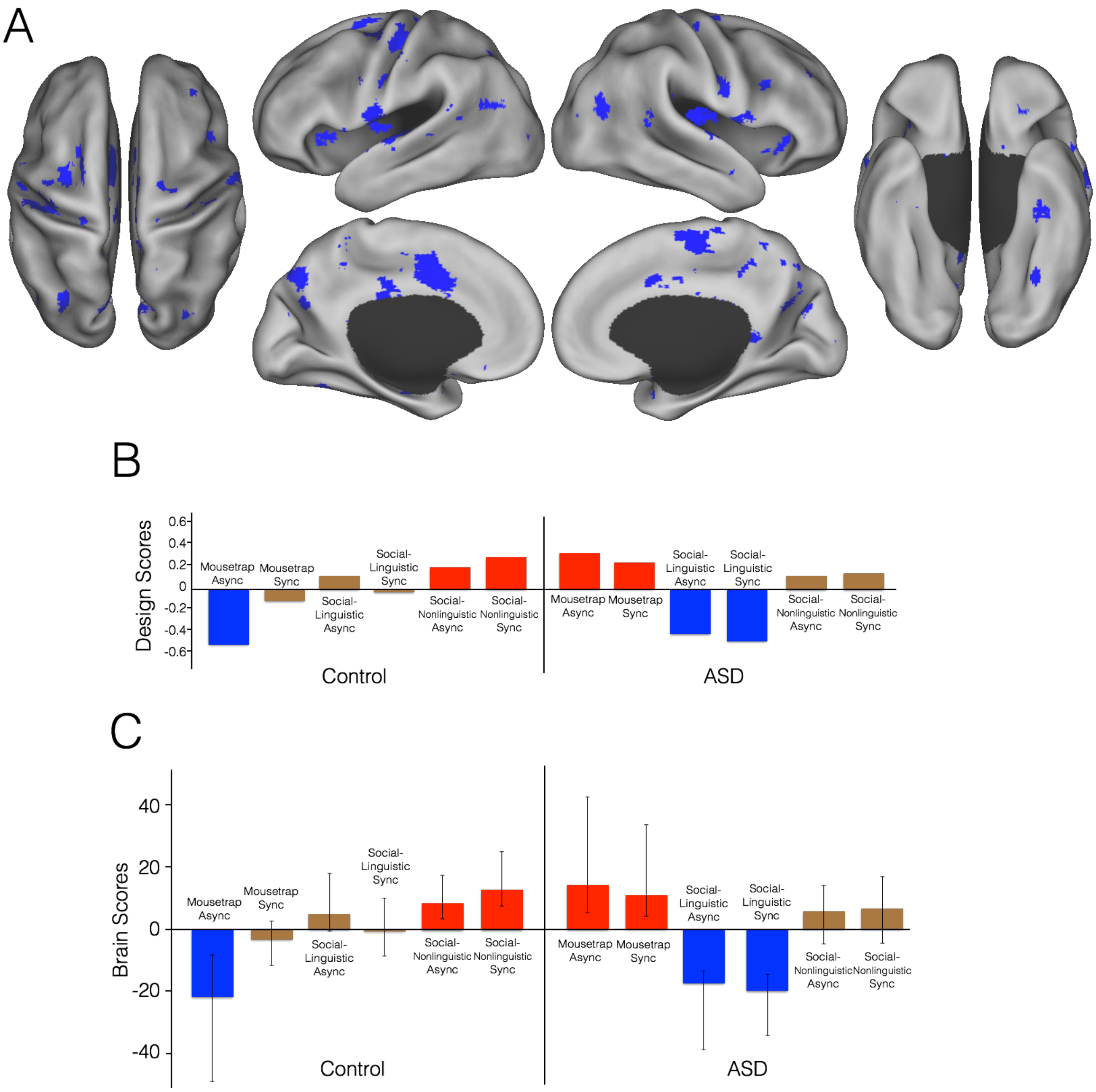


Figure 27. Third latent variable for task-related brain activity for the synchronous and asynchronous mousetrap, social-linguistic, and social-nonlinguistic content conditions in the participants with and without ASD. LV3: accounted for 9.21% of the covariance, *p* = 0.01 (A) Activity associated with synchronous and asynchronous social-nonlinguistic processing for the control group and synchronous and asynchronous mousetrap processing for the group with ASD is in red. Activity associated with asynchronous mousetrap processing for the control group, and synchronous and asynchronous social-

linguistic processing for the group with ASD is in blue. Data are displayed on the dorsal, lateral, medial, and ventral surfaces of the left and right hemispheres of a partially inflated surface map using Caret software (Van Essen, 2005). (B) Design scores for each condition. In the control group, the synchronous and asynchronous social-nonlinguistic conditions were maximally dissociated from asynchronous mousetrap condition. In the group with ASD, the synchronous and asynchronous mousetrap conditions were maximally dissociated from the synchronous and asynchronous social-linguistic conditions. Conditions that did not significantly contribute to the latent variable are in brown. (C) Mean brain scores across participants in each condition. Error bars represent 95% confidence intervals. Statistically significant differences between conditions are indicated by a lack of overlap of the confidence intervals.