Investigation of multisensory processing and structural brain differences in Autism Spectrum Disorder

Paula Regener

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

School of Psychology

College of Science and Engineering

University of Glasgow

October 2015

Abstract

This thesis is an investigation of structural brain abnormalities, as well as multisensory and unisensory processing deficits in autistic traits and Autism Spectrum Disorder (ASD). To achieve this, structural and functional magnetic resonance imaging (fMRI) and psychophysical techniques were employed.

ASD is a neurodevelopmental condition which is characterised by the social communication and interaction deficits, as well as repetitive patterns of behaviour, interests and activities. These traits are thought to be present in a typical population. The Autism Spectrum Quotient questionnaire (AQ) was developed to assess the prevalence of autistic traits in the general population. Von dem Hagen et al. (2011) revealed a link between AQ with white matter (WM) and grey matter (GM) volume (using voxel-based-morphometry). However, their findings revealed no difference in GM in areas associated with social cognition. Cortical thickness (CT) measurements are known to be a more direct measure of cortical morphology than GM volume. Therefore, Chapter 2 investigated the relationship between AQ scores and CT in the same sample of participants. This study showed that AQ scores correlated with CT in the left temporo-occipital junction, left posterior cingulate, right precentral gyrus and bilateral precentral sulcus, in a typical population. These areas were previously associated with structural and functional differences in ASD. Thus the findings suggest, to some extent, autistic traits are reflected in brain structure - in the general population.

The ability to integrate auditory and visual information is crucial to everyday life, and results are mixed regarding how ASD influences audiovisual integration. To investigate this question, Chapter 3 examined the Temporal Integration Window (TIW), which indicates how precisely sight and sound need to be temporally aligned so that a unitary audiovisual event can be perceived. 26 adult males with ASD and 26 age and IQ-matched typically developed males were presented with flash-beep (BF), point-light drummer, and face-voice (FV) displays with varying degrees of asynchrony and asked to make Synchrony Judgements (SJ) and Temporal Order Judgements (TOJ). Analysis of the data included fitting Gaussian functions as well as using an Independent Channels Model (ICM) to fit the data (Garcia-Perez & Alcala-Quintana, 2012). Gaussian curve fitting for SJs showed that the ASD group had a

wider TIW, but for TOJ no group effect was found. The ICM supported these results and model parameters indicated that the wider TIW for SJs in the ASD group was not due to sensory processing at the unisensory level, but rather due to decreased temporal resolution at a decisional level of combining sensory information. Furthermore, when performing TOJ, the ICM revealed a smaller Point of Subjective Simultaneity (PSS; closer to physical synchrony) in the ASD group than in the TD group.

Finding that audiovisual temporal processing is different in ASD encouraged us to investigate the neural correlates of multisensory as well as unisensory processing using functional magnetic resonance imaging fMRI. Therefore, Chapter 4 investigated audiovisual, auditory and visual processing in ASD of simple BF displays and complex, social FV displays. During a block design experiment, we measured the BOLD signal when 13 adults with ASD and 13 typically developed (TD) age-sex- and IQ- matched adults were presented with audiovisual, audio and visual information of BF and FV displays. Our analyses revealed that processing of audiovisual as well as unisensory auditory and visual stimulus conditions in both the BF and FV displays was associated with reduced activation in ASD. Audiovisual, auditory and visual conditions of FV stimuli revealed reduced activation in ASD in regions of the frontal cortex, while BF stimuli revealed reduced activation the lingual gyri. The inferior parietal gyrus revealed an interaction between stimulus sensory condition of BF stimuli and group. Conjunction analyses revealed smaller regions of the superior temporal cortex (STC) in ASD to be audiovisual sensitive. Against our predictions, the STC did not reveal any activation differences, per se, between the two groups. However, a superior frontal area was shown to be sensitive to audiovisual face-voice stimuli in the TD group, but not in the ASD group. Overall this study indicated differences in brain activity for audiovisual, auditory and visual processing of social and non-social stimuli in individuals with ASD compared to TD individuals. These results contrast previous behavioural findings, suggesting different audiovisual integration, yet intact auditory and visual processing in ASD.

Our behavioural findings revealed audiovisual temporal processing deficits in ASD during SJ tasks, therefore we investigated the neural correlates of SJ in ASD and TD controls. Similar to Chapter 4, we used fMRI in Chapter 5 to investigate audiovisual temporal processing in ASD in the same participants as recruited in Chapter 4. BOLD

signals were measured while the ASD and TD participants were asked to make SJ on audiovisual displays of different levels of asynchrony: the participants' PSS, audio leading visual information (audio first), visual leading audio information (visual first). Whereas no effect of group was found with BF displays, increased putamen activation was observed in ASD participants compared to TD participants when making SJs on FV displays. Investigating SJ on audiovisual displays in the bilateral superior temporal gyrus (STG), an area involved in audiovisual integration (see Chapter 4), we found no group differences or interaction between group and levels of audiovisual asynchrony. The investigation of different levels of asynchrony revealed a complex pattern of results indicating a network of areas more involved in processing PSS than audio first and visual first, as well as areas responding differently to audio first compared to video first. These activation differences between audio first and video first in different brain areas are constant with the view that audio leading and visual leading stimuli are processed differently.

Abbreviations

A . . . . . audio

AF . . . . . audio first

ANOVA . . . . . analysis of variance

ASD . . . . . Autism Spectrum Disorder

AV . . . . . audiovisual

BA . . . . . Brodmann area

BF . . . . . beep-flash

EEG . . . . . electroencephalography

BOLD . . . . . blood-oxygen-level dependence

BPM . . . . . beats per minute

DLPFC . . . . . dorsolateral prefrontal cortex

fMRI . . . . . functional magnetic resonance imaging

FV . . . . . face-voice

FWE . . . . . family wise error

FWHM . . . . . full width at half maximum

GA . . . . . genetic algorithm

GLM. . . . . general linear model

ICM . . . . . independent channels model

PLD . . . . . point light drumming

PSS . . . . . point of subjective simultaneity

RFX . . . . . random effects

SD . . . . . standard deviation

SJ . . . . . synchrony judgment

SOA . . . . . stimulus onset asynchrony

STC . . . . . superior temporal cortex

STS . . . . . superior temporal sulcus

TD. . . . . typically developed

TIW . . . . . temporal integration window

TOJ . . . . . temporal order judgment

V . . . . . video

VF . . . . . video first

List of Tables

List of Figures

Acknowledgements

During my PhD I have required and received help from a number of people, both in academic and personal matters. I would like to take this opportunity to thank as many of them as possible.

Firstly, I would like to thank all my participants, who enthusiastically took part in my research.

I would also like to thank the members of Frank's lab who over the years have helped with different aspects of my PhD, as well as supporting me in developing the skills I needed. In particular, I would like to thank Katie Noble for fMRI analysis support, as well as our numerous lunch break adventures and discussions. I'd like to thank Aso Muhamed and Salim Al-wasity for helping me by building the PACO PRT creator and Presentation programming.

I would also like to thank previous lab members Karin Petrini, Scott Love and Lawrie McKay, who I could always email with questions or run ideas by.

To my supervisors, Frank Pollick and David Simmons, I wish to offer special thanks for helping me obtain the funding I needed in order to carry out the PhD. I would also like to thank David Simmons for his enthusiasm for my research and always being there for good advice. Frank Pollick, I would like to thank you in particular for helping constantly throughout this PhD, and for enabling me to disseminate my research at many different interesting conferences across the world.

Outside of my lab, I would like to thank my office mates Ashley Robertson, Kirsty Ainsworth, Aliyah Rehman and Louise Beattie for being so patient with me and my messy desk. A big thank you also to all my amazing friends for all the help and support they gave me over the years.

Author's Declaration:

This thesis has been composed by the undersigned. It has not been accepted in any previous application for a degree. The work, of which this thesis is a record, has been completed by myself, unless otherwise indicated in the text. I further state that no part of this thesis has already been, or is concurrently, submitted for any such degree or qualification at any other university.

.....

Paula Regener

1. Introduction

In this thesis I will be investigating how different senses are combined by the brain and how that influences behaviour in both participants with ASD and typically developed (TD) participants. I will also be looking at how structural aspects of the brain correlate with autistic traits and how such underlying structural abnormalities can be related to audiovisual processing in ASD.

Integration of information across different sensory modalities is an important part of everyday experience, as we are constantly flooded with different sensory stimuli and have to decide which stimuli belong together and which are unrelated. Integration of audio and visual information is particularly important in speech perception (Massaro, 1998). It has been shown that we tolerate a degree of temporal asynchrony between sound and sight and still perceive it as one event; this is called the temporal integration window (TIW).

Autism spectrum disorders (ASD) are a range of neurodevelopmental conditions often characterized by widespread abnormalities in social interactions and communication, as well as severely restricted interests and repetitive behaviour (American Psychiatric Association, 2000). Kanner (1943) originally reports sensory abnormalities in his description of autism. Sensory abnormalities have consistently been reported in clinical literature (e.g., Leekam, Nieto, Libby, Wing & Gould, 2007) and the DSM-V has included sensory abnormalities as a central feature in ASD. Traits of ASD are said to lie on a continuum within the general population (Frith, 1991; Baron-Cohen, 1995). To measure the extent of autistic traits in the general population the Autism-Spectrum Quotient (AQ; Baron-Cohen et al., 2001), a self-administered questionnaire, has been developed. The AQ is used to predict performance on tasks that are impaired in ASD, for example inferring others' mental states from the eyes (Baron-Cohen et al., 2001).

1.1 Audiovisual temporal processing

A well-known example of how we integrate audio and visual information to form a single percept is the McGurk illusion (McGurk & MacDonald, 1976), during which the pairing of incongruent visual and auditory speech information results in a novel integrated percept. For example the syllable “ba-ba” is simultaneously spoken over the video of someone saying “ga-ga” produces the combined percept of “da-da”. However, this McGurk illusion gets weaker the bigger the temporal asynchrony

between the visual and the auditory cues (e.g., Jones & Jarick, 2006). A similar illusion is the flash-beep illusion, which is based on the phenomenon that auditory stimulation (beeps) can influence the perception of visual stimulation (flashes). For example when a single flash is presented simultaneously with two beeps people perceive two flashes (Shams, Kamitani, & Shimojo, 2000). The bigger the temporal asynchrony between the second beep and the flash, the less often the second flash is perceived (Shams, Kamitani & Shimojo 2002). The attenuation of these multisensory illusions is due to the degree of asynchrony between the two multisensory signal cues, showing the importance of temporal synchrony, but also that multisensory integration does not require exact temporal synchrony. In the aforementioned examples, a small degree of asynchrony had little effect.

1.2 Measuring audiovisual temporal processing

During this section, the different methods used to investigate audiovisual processing are introduced. There are a variety of tasks that claim to measure the same psychophysical parameters, but whether the different tasks tap into the same perceptual mechanisms is questionable (Love, Petrini, Cheng & Frank, 2013). The psychophysical parameters that are used when investigating audiovisual synchrony perception are the point of subjective simultaneity (PSS) and the width of the temporal integration window (TIW). The PSS is a value that corresponds to the participant's stimulus onset asynchrony (SOA) most often perceived as synchronous. This often deviates from the true synchronous point, i.e., when the SOA of the audio and the visual stimuli equals 0ms. Thus the PSS value presents the time difference between the audio and visual stimuli that is required for an individual to optimally perceive them as synchronous, and this is often a non-zero value as people are not perfect at detection asynchrony. For example, a negative PSS value indicates that the individual perceived synchrony when the audio information was presented before the visual information (audio-leading asynchrony). Furthermore, within a range of SOAs centred around the PSS, known as TIW, people are unable to reliably detect asynchrony between the audio and visual stimuli. The TIW could be described as a range of SOAs, during which we are not sensitive to certain levels of asynchrony. Thus the TIW width measures the sensitivity of task responses to changes in SOA, i.e., narrow TIW represent higher sensitivity to deviation from perceived audiovisual synchrony. The PSS and the TIW width can be measured using a range of different tasks

1. 2. 1 Synchrony Judgements

Synchrony Judgements (SJ) are commonly used to measure people's PSS and TIW width. During SJs, participants are presented with audiovisual displays at various levels of SOA and are asked to judge whether the audio and the visual information were displayed in synch or out of synch (e.g., Petrini et al., 2009a,b; Love at al., 2013). Commonly, the Gaussian probability density functions (e.g., Love et al., 2013, Petrini et al., 2009) or two cumulative Gaussians (e.g., van Eijk, Kohlrauch, Juola & van de Par, 2008, Stevenson et al 2014) are fitted to the proportion of synchronous responses at each SOA level from which the PSS and TIW width are derived. The PSS is the highest point of the fitted function, whereas the TIW width is either derived from the standard deviation or the full width at half maximum of the fitted function. SJ can lead to response biases, such as the equalisation bias which occurs because there is only one physically synchronous condition, but many asynchronous conditions. Participants might try to equalise frequency of the asynchronous and synchronous conditions, which has been reported before in other psychophysical experiments (Erlebacher & Sekuler, 1971). The width of the TIW might also depend on the participant's subjective criterion setting. A participant might have less stringent criteria and responds "in synch" more often than someone with more stringent criteria (Vroomen & Keetels, 2010). A cognitive bias of participants assuming that the audio and visual information naturally belong together and must therefore be synchronous might also occur (Vatakis & Spence, 2007).

1.2.1 Temporal Order Judgements

Temporal Order Judgements (TOJ) are also frequently used to investigate participants' PSS and TIW width. During TOJs, the participants are presented with audiovisual displays at a range of different levels of SOAs and their task is to decide whether the audio or the visual information was presented first (e.g. Vatakis and Spence, 2007; de Boer-Schellekens, Eussen & JeanVroomen, 2013; Love et al., 2013). The PSS and TIW width are commonly derived from fitting a cumulative Gaussian distribution function to the proportion of video first or audio first responses (e.g., Love et al., 2013). The PSS is the 50% point on the function and the TIW is either taken as the just noticeable difference (JND) or the standard deviation of the function. The PSS has been suggested to be influenced by a response bias, as participants may have a bias towards either responding audio first or visual first when they are guessing their

responses. However, Fujisaki and Nishida (2009) argued that the TIW width would not be influenced by this.

1.2.2 Three Choice Synchrony Judgements

Another popular task measuring the PSS and TIW width is the 3 choice Synchrony Judgement (SJ3) task, where participants are asked to indicate whether the audio stimulus or the visual stimulus was presented first or whether they were presented simultaneously. Two cumulative Gaussians are commonly fitted to the proportion of synchronous responses at each SOA (e.g., van Eijk et al., 2008).

Although SJ, TOJ and SJ3 tasks have been used almost interchangeably in the literature to investigate temporal processing, recent research shows that these tasks produce different PSS and TIW width (Love at al., 2013; Petrini et al., 2010; Van Eijk at al., 2008), thus suggesting they measure different processes. Van Eijk at al. (2008) found that PSS estimates of SJ and SJ3 tasks were similar and highly correlated, whereas the PSS estimates of TOJ were significantly different and were uncorrelated to those of SJ and SJ3 tasks. These differences suggested that TOJ might have different underlying processes compared to the SJ tasks (Spence & Parise 2010). Furthermore, Love at al., (2013) showed no correlation of between SJ and TOJ for neither PSS nor TIW. Further evidence for different underlying perceptual differences between the tasks is that training in one task does not influence performance of the other task (Mossbridge et al., 2006). Furthermore, the differences between these tasks have been argued to be due to decisional aspects and not due to sensory parameters of the tasks (Garcia-Perez & Alcala-Quintana, 2012).

1.2.3 Implicit measures of sensitivity of audiovisual asynchrony

While the above sections discussed explicit ways of measuring sensitivity to audiovisual asynchrony, it should be noted that there are also more implicit ways of measuring this sensitivity. For example, the perception of audiovisual illusions is often dependent on temporal synchrony and can therefore be used as an implicit measurement of the sensitivity to audiovisual asynchrony. The flash-beep illusion, described above, elicited when two beeps are presented simultaneous with a flash, which causes the participants to perceive two flashes. However, the bigger the temporal asynchrony between the second beep and the flash, the less often the second flash was perceived (Shams, Kamitani & Shimojo 2002).

1.3 Fitting procedures of synchrony judgements and temporal order judgements

As mentioned above, the response data for SJs is commonly fitted with a Gaussian probability density function (e.g., Love et al., 2013, Petrini et al., 2009) or two cumulative Gaussians (e.g., van Eijk et al., 2008, Stevenson et al 2014), whereas, TOJs are often fitted with a cumulative Gaussian distribution function (e.g., van Eijk et al., 2008; Love et al., 2013) or a linear function (de Boer-Schellekens et al., 2013). Fitting these psychometric functions to SJ and TOJ responses does not account for asymmetry and irregularities within the data. For example, an individual's proportion of synchronous responses in SJs are known to generally be asymmetric. Similarly, "video first" responses of an individual in TOJs often show a pronounced plateau midway along the range of SOAs. Once the data is averaged across individuals, however, these asymmetries and irregularities are likely to be averaged out, too, and information might be lost. Therefore, it is questionable whether presenting the participants response data by using best-fitting Gaussian functions is the most appropriate procedure (Maier et al., 2011; Garcia-Perez & Alcala-Quintana, 2012). Maier et al. (2011) suggested four new metrics (peak location, peak performance, width, asymmetry) to measure synchrony perception performance and replace the PSS and TIW, which do not require fitting the data to Gaussian functions. However, this approach has its draw backs and the authors demonstrated the inability to calculate the width and asymmetry when the peak location was significantly shifted towards the video leading side of the x-axis, which forced them to define another alternative metric to represent the TIW. Another alternative to fitting Gaussian functions to the response data of temporal judgement tasks was suggested by Garcia-Perez and Alcala-Quintana (2012). The authors proposed the Independent Channels Model (ICM) of timing judgements (Sternberg & Knoll, 1973) to fit the raw data of SJ and TOJ in a more flexible manner.

1.3.1 Independent Channels Model

Garcia-Perez and Alcala-Quintana (2012) used the Independent Channels Model (ICM) of timing judgements (Sternberg & Knoll, 1973) to fit the response data of SJ and TOJ in a more flexible way, allowing for individual asymmetries and irregularities in the data. Furthermore research suggests that SJ and TOJ tasks have different response biases. The nature of these response biases and the underlying processes have recently been modelled by Garcia-Perez and Alcala-Quintana (2012). In their

study they made use of the ICM of timing judgements to take into account the sensory processing factors and decisional aspects involved in SJs and TOJs (see Figure 1.1). The authors fitted the ICM to Van Eijk at al.'s (2008) data by both fitting the tasks together (assuming similar sensory parameters but different decisional parameters) and fitting the tasks separately (allowing for parameter differences across tasks). Garcia-Perez and Alcala-Quintana (2012) argued that by fitting the ICM to the data and estimating the arrival latencies, decisional factors, response bias and response errors they can infer underlying processes to the tasks. Their results showed that arrival latencies did not differ across tasks when stimulus conditions were identical, but that the resolution parameter was different across tasks. The ICM provides estimates similar to the TIW width and PSS comparable to the TIW width and PSS outcome measures obtained through the Gaussian fits. The parameters related to sensory and decisional factors of audiovisual processing that the ICM provides are Delta, the onset, Lambda, the rate parameter and Tau, the processing delay of the corresponding sensory information. Lambda Audio (A), Lambda Visual (V) and Tau describe the arrival latency in SJ and TOJ tasks, and Delta is the resolution parameter and it limits the observer's ability to detect small differences in arrival latencies. TOJ includes an additional response bias parameter called Xi, taking into account the tendency of participants to respond "audio first" or "video first" more often.

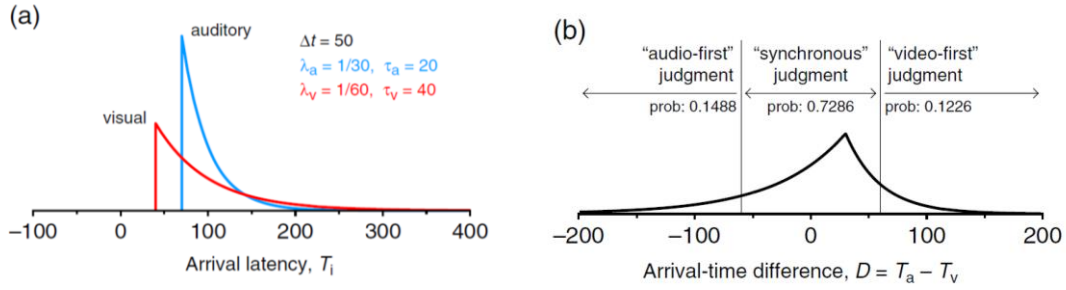


Figure 1.1 Model of timing judgments. a) Exponential distributions for the arrival latency of a visual stimulus (red curve) presented at time 0 and an auditory stimulus (blue curve) presented at time Δt 0 = 50 ms. Parameters as indicated in the inset. b) Bilateral exponential distribution of arrival-time differences and cutpoints on the decision space (vertical lines, at $D = \pm\delta$ with $\delta = 60$), determining the probability of each judgment (taken from Garcia-Perez & Alcala-Quintana, 2012).

1.4 Neuroanatomical differences in ASD

Previous research has shown that the trajectory of brain volume development is different in ASD compared to the typical population. Brains of new-borns with ASD tend to be comparable in volume to brains of typically developing new-borns (Courchesne, Carper & Akshoomoff, 2003; Dawso at al., 2007), but tend to be enlarged in early childhood (Anagnostou & Taylor, 2011; Courchesne et al., 2001; Levy, Mandel & Schulz, 2009). In adolescence and early adulthood, the results are less clear, while some researchers find that the increased brain volume is still present in individuals with ASD (e.g., Freitag et al., 2009; Hazlett, Poe, Gerig, Smith, & Piven, 2006) others find a normalisation of total brain volume (e.g., Aylward et al., 2002; Redcay & Courchesne, 2005; review: Courchesne et al., 2007; Hyde, Samson, Evans & Mottron, 2010). It is also elusive whether this putative increase in total brain volume is a result of grey matter (GM) volume (Hazlett et al., 2006), white matter (WM) volume (Herbert et al., 2004), or a combination of both.

1.4.1 Diffusion tensor imaging differences in ASD

Diffusion weighted magnetic resonance imaging (DW-MRI) has been suggested to be the most direct, non-invasive way of mapping white matter (WM) tracts in vivo (Le Bihan et al., 2001). Diffusion tensor imaging (DTI) is used to investigate the WM tracts by providing a measure of diffusion (most often of water molecules) within voxels of the brain (Assaf &Pasternak, 2008). Factional anisotropy (FA) is one of the four different measures used to investigate the diffusivity of tissue microstructure. FA provides a measure of coherence of diffusion directionality (diffusion anisotropy), which ranges from entirely isotropic (identical properties in all directions) to entirely anisotropic (directionally driven). Other measures to investigate diffusivity of tissue microstructure are fibre coherence (Le Bihan et al., 2001); mean diffusivity (MD) as well as, axial diffusivity (AD) and radial diffusivity (RD), which describe diffusivity that is parallel and perpendicular to the axonal fibres, respectively.

DTI findings in ASD are somewhat heterogeneous. However, overall it appears that WM abnormality in ASD is found throughout the entire brain. For example, a recent study by Roine et al. (2013) noted that their ASD group globally had increased FA compared with their TD group. With regards to individual structures and pathways,

WM abnormality has most reliably been found in the corpus callosum. More precisely, studies have found both reduction (e.g., Shukla et al., 2011; Walker et al., 2012; Gibbard et al., 2013) and increases (Billeci et al., 2012) of FA, and increased MD and RD (Shukla et al., 2011) in individuals with ASD compared with TD controls in different regions within the corpus callosum.

Findings of other WM in other structures measured by DTI are again rather inconsistent. A study found increased MD, RD and AD in individuals with ASD in posterior WM tracts (Walker et al., 2012), while another revealed increased MD and RD values in the frontal areas of the brain (Ameis et al., 2011) . Shukla et al. (2011) and Barnea-Goraly et al. (2010) noted that differences in ASD of MD, RD and AD are more extensive throughout the brain, and expand across association, commissural and projection fibres.

1.4.2 Cortical thickness and grey matter volume

In MRI studies, methods called cortical thickness analysis (CTA) and voxel-based morphometry (VBM) are commonly used to investigate structural properties of GM in ASD. CTA is said to directly measure cortical surface features, such as cortical thickness (CT; Jiao et al., 2010), while VBM gives a probabilistic measure of local GM and WM concentration (Ashburner & Friston, 2000). The VBM method has been reported to be restricted as it conflates information about morphology, size and position (Ashburner & Friston, 2001). CTA provides a more direct index of cortical morphology that is less susceptible to positional variance given that the extraction of the cortex follows the GM surface despite local variations in its position (Kim et al., 2005).

CT and GM volume are thought to reflect changes in myelination (Sowell et al., 2007) and neuronal loss in ageing (Salat et al. 2004), and CT procedures have been validated using post-mortem histological analysis (Rosas et al., 2002). CT measurements have also been used to provide a method of relating changes in brain structure to cognitive abilities, behaviour (Anagnostou &Taylor, 2011) and activation levels (Fusar-Poli et al., 2011), suggesting that functional and structural abnormalities share a common pathophysiology. However, it needs to be noted that the T1-weighted signal used to measure CT represent the degree of MRI visible water which is least visible in white matter, intermediately so in GM, and most visible in cerebrospinal fluid (Diwadkar &

Keshavan, 2002). MRI data has neither the resolution nor the specificity to explain the relationship between estimated CT and complex cellular processes including dendritic remodelling, cell death, synaptic pruning, or plausible encroachment from myelination (Toga, Thompson & Sowell, 2006). Diwadkar et al. (2011) investigated adolescent children of individuals with Schizophrenia and did not find a correlation between functional hypoactivity in frontal and parietal cortex and GM volume differences compared to control participants. Moreover, another study showed that the neural bases of GM estimates and blood oxygen level depletion (BOLD) appear to be independent or have a complex relationship (Kannurpatti, Motes, Rypma & Biswal, 2010).

1.4.3 Cortical thickness and grey matter volume differences in ASD

A recent study by Zielinski et al. (2014) examined CT from childhood to adulthood using a large mixed cross-sectional and longitudinal sample of autistic subjects and their controls, and found early accelerated growth in childhood followed by accelerated thinning in adolescence and decelerated thinning in early adulthood. Similarly, Osipowicz, Bosenbark & Patrick (2015) examined GM volume across the lifespan of people with ASD and their controls, as well as, correlated GM volume with autism severity. They showed bilateral decreases of GM volume in the ASD group in the thalamus, the cerebellum, anterior medial temporal lobes and the orbitofrontal regions. More severe autism was associated with decreased GM volume in the prefrontal cortex, inferior parietal and temporal cortex, as well as, temporal poles. No increases of GM volume were associated with ASD or its severity. GM volume and autistic symptomology and severity have also been found to be correlated in children (Pierce & Corchesne, 2001), whereas other researcher have found no relationship (Langen, Durston, Staal, Palmen & van Engeland, 2007).

Looking at GM and CT in adults with ASD, studies have reported GM volume or CT increases throughout the whole brain and specific regions, such as the frontal, temporal and parietal regions, lingual gyrus, insular regions, precentral gyrus postcentral and cingulate gyri, caudate nucleus, hippocampus, brainstem and midbrain (Hyde, Samson, Evans & Mottron, 2010; Ecker et al., 2010; Waiter et al., 2004; Doyle-Thomas et al., 2013; Ecker et al., 2013). Others have found specific reduction in GM volume or CT in temporal and parietal regions, sensory and motor cortex,

anterior cingulate, supramarginal gyrus, precentral and postcentral gyri, thalamus, corpus callosum, cerebellum, parahippocampal gyrus (Chung et al., 2005; Hadjikhani, Joseph, Snyder & Tager-Flusberg, 2006; Hyde et al., 2010; Scheel et al., 2011; Toal et al., 2010; Greimel et al., 2013; Ecker et al, 2013; Ecker et al., 2010; McAlonan et al., 2002; Wallace, Dankner, Kenworthy, Giedd & Martin, 2010; See Table 1.1 for an overview).

This heterogeneity of results across studies investigating cortical morphology in participants with ASD may be due to factors such as a variety of algorithms and techniques being used to compute CT and GM volume, differences in MRI image resolution across studies, as well as, sample heterogeneity (i.e., differences in diagnostic methods, participants' ages and IQ). In fact, sample heterogeneity is a common problem in studying ASD.

1.5 Neuroanatomical differences and autistic traits measured by AQ

To measure the extent of autistic traits in the general population, the Autism-Spectrum Quotient (AQ; Baron-Cohen et al., 2001), a self-administered questionnaire, has been developed. The AQ has been used to predict performance on tasks that are impaired in ASD, for example inferring others' mental states from the eyes (Baron-Cohen et al., 2001). The AQ has been shown to be associated with changes in brain structure, including GM volume, WM volume, SulcoGyral patterns and Diffusion Tensor Imaging (DTI) in typically developed brains (Iidaka et al., 2012; Kosaka et al., 2010; Geurts er al., 2013; Saito et al., 2013; Gebauer et al., 2015; Von dem Hagen et al., 2011). However, a recent exploration-validation study showed no association between AQ scores and brain structure, including analysis of GM volume, CT and DTI (Koolschijn, Greurts, Leij & Scholte, 2015).

Autistic traits in the general population have also been correlated with structural differences. Higher AQ scores correlate with smaller GM volume of right insula and inferior frontal gyrus, larger GM volume of left middle frontal gyrus (Kosaka et al., 2010; Geurts et al., 2013; Saito et al., 2013) and reduced CT in right medial orbitofrontal cortex, postcentral gyrus, lingual gyrus (Gebauer et al., 2015), whereas others showed no links between AQ scores GM volume and CT (Koolschijn et al., 2015).

Table 1.1 Studies investigating grey matter volume in ASD

Brain Region		Method	ASD			TD controls		Authors
↓*	↑*		n**	Age ***	Diagnosed by	n**	Age**	
Middle frontal gyrus L/R, Precentral gyrus R, Inferior frontal gyrus L/R, Amygdala R, Hippocampal gyrus L/R, Uncinate L/R, Inferior parietal lobe L/R, Superior parietal lobe L/R, Precuneus, Posterior cingulate gyrus L/R, Precuneus L/R, Cerebellar cortex L/R, Putamen R, Caudate nucleus R	Inferior frontal gyrus L, DLPFC L/R, Precentral gyrus L/R, Middle temporal gyrus L/R, Inferior temporal gyrus L/R, STS L/R, Fusiform gyrus L/R, Inferior parietal lobe L/R, Medial occipital gyrus L/R, Lingual gyrus L/R, Insular cortex L/R	SVM (support vector machine approach)	22 m	27 (7)	ADI-R, ADOS, AQ	22	28 (7)	Ecker et al., 2010
Cerabellar Crus L/R, Cerebellar Lobule L/R	Medial frontal gyrus L/R, Precentral gyrus L, Postcentral gyrus R, Fusiform gyrus R, Caudate nucleus L/R, Hippocampus L	VBM	24 m	21 (11)	DSM-IV, ADI, ADOS	23 m	21 (11)	Rojas et al., 2006
Postcentral gyrus, R, Precentral gyrus, L/R	Brainstem/midbrain, medial frontal gyrus R, Medial oribital frontal gyrus L, Middle frontal gyrus L/R	VBM	15 m	23 (6)	ADI-R, ADOS	15 m	19 (5)	Hyde et al., 2010
Paracentral gyrus R, Postcentral gyrus R, Precentral gyrus R	Anterior fusiform gyrus L/R, Anterior STS L, Dorsal posterior cingulate gyrus L/R, Heschl's gyrus L/R, Lingual gyrus L, Medial frontal gyrus L/R, Medial oribital frontal gyrus L, Middle frontal gyri L/R, Posterior fusiform gyrus L, pSTS, L, Ventral posterior cingulate gyrus L, Dorsal anterior cingulate gyrus R, Inferior frontal gyrus R, Inferior parietal lobule R, Middle occipital gyrus R, Superior frontal gyrus R, STS R	CTA						Hyde et al., 2010
Thalamus R	Fusiform gyrus R, Temporo-occipital region R, Frontal pole extending to the Medial frontal cortex L.	VBM	16 m	15 (2)	DSM-IV, ADI-R, ADOS-G	16 m	15 (2)	Waiter et al., 2004
Inferior orbital prefrontal cortex R, superior temporal sulcus L, Occipito-temporal gyrus L	None	CTA	16 m	16 (5)	ADI-R	12 m	17 (3)	Chung et al., 2005
Cerebellum R, Lenticular nucleus L/R, Cingulate gyrus R, Precuneus R, Medial frontal gyrus L/R, Superior frontal gyrus R	None	VBM	21 (19 m)	32 (10)	ICD 10, ADI-R	24 (22 m)	33 (7)	McAlonan et al., 2002
Cerebellum L/R, Parahippocampal gyrus L/R, Fusiform gyrus L/R, from Inferior temporal gyrus L/R to STS L/R	None	VBM	65 (57 m)	31 (10)	ICD-10, ADI-R, ADOS	33 (30 m)	32 (9)	Toal et al., 2010
Caudal middle frontal gyrus L, Paracentral frontal gyrus L, Superior temporal gyrus L, Inferior temporal gyrus L Entorhinal gyrus L, Fusiform, gyrus L, Superior Banks temporal sulcus L/R, Superior parietal gyrusL/R, Inferior parietal gyrus L/R, Supramarginal gyrusL, Postcentral parietal lobe L	None	CT measurement	41 m	17 (3)	ADI, ADOS	40 m	17 (3)	Wallace et al.,, 2010
Posterior STS L, Middle temporal gyrus L, Supramarginal gyrus L	None	CTA	28 (18 m)	33 (10)	AQ	28 (18 m)	33 (9)	Scheel et al., 2011
Inferior frontal gyrus L/R, inferior parietal lobe L/R, the STS R, precentral Gyrus L/R, inferior occipital gyrus L/R, orbitofrontal cortex L/R, Anterior cingulate R, supramarginal gyrus R, middle occipital gyrus L, superior parietal lobule R/L, medial parietal cortex L, superior parietal lobule L/R	none		14m	33 (12)	ADI-R, ADOS	14 m	31 (9) (1.4)	Hadjikhani, Joseph, Snyder & Tager-Flusberg, 2006;
None	Superior frontal gyrus L/R, Middle frontal gyrus L/R, Inferior frontal gyrus L/R, Medial frontal gyrus L/R, Orbitofrontal gyrus L/R, Precental Gyrus L/R, Postcentral Gyrus L/R, Superior parietal lobule L/R, Inferior parietal lobule/Middle occipital gyrus L, Inferior parietal lobule R, pSTG L/R, Middle temporal gyrus L/R, Inferior temporal gyrus L/R, Medial orbitofrontal gyrus L, Posterior cingulate L/R, Precuneus L, Parietoccipital fissure R.	CTA	28 (21m)	22.5 (7.9)	DSM-IV, ADI-R, ADOS	26 (17m)	22.7 (9.1)	Doyle-Thomas et al., 2013
Anterior cingulate cortex L/R, pSTS L/R, middle temporal gyrus.R.	None	VBM	47m	21.4 (10.1)	DSM-IV,ADOS-G, ADI-R, AQ	51m	18.3 (7.5)	Greimel et al., 2013
Frontal anterior cingulate L, Medial prefrontal cortex R, Middle temporal gyrus L/R, Inferior temporal gyrus R, Parahippocampal gyrus L, Superior parietal L/R, Supramarginal gyrus L, Precuneus L, Pericalcarine fissure R, Lingual gyrus R.	Superior frontal L, Causal middle frontal L/R, Rostral middle frontal L/R, Pars opercularis L, Pars triangularis R, Medial orbitofrontal R, Middle temporal gyrus R, STG R, Inferior parietal L, Supramarginal gyrus L/R, Superior parietal R, Lateral occipital cortex L/R, Postcentral gyrus L/R, Posterior cingulate L.	CT measurement	84m	26 (7)	ICD-10, ADI-R	84m	28 (6)	Ecker et al, 2013

* Decreased (↓) and increased (↑) gray matter or cortical thickness in ASD compared to TD controls, ** Number of participants (m = males) *** Mean age, standard deviations are in brackets,
For all studies the IQ of the ASD group was not different to that of the TD controls

1.6 Neuroscience of audiovisual processing

From studies investigating multisensory integration, superior temporal cortex (STC) plays a major role in integrating audio and visual cues of social and non-social stimuli (Watson et al., 2014; Steveson & James, 2009; Allison, Puce, & McCarthy, 2000). Moreover, parts of the posterior superior temporal sulcus (pSTS; a sub region of the STC) have been shown to respond more to social signals, compared to non-social control stimuli in both the visual and auditory modalities, although the relative location of face- and voice-sensitive regions in pSTS remains unclear (face: Haxby, Hoffman, & Gobbini, 2000; Hoffman & Haxby, 2000; voice: Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Ethofer, Van De Ville, Scherer, & Vuilleumier, 2009; Grandjean et al., 2005; Latinus, Crabbe, & Belin, 2011). Moreover, evidence for the pSTS to be involved in audiovisual integration is currently accumulating. Functional magnetic resonance imaging (fMRI) studies investigating audiovisual integration have been searching for brain areas which are involved in the processing of unisensory audio and visual information, but show an even stronger activation response to the information when presented together. The different statistical criteria of audiovisual integration regions are summarised below. Researchers find the STC to be involved in integrating audiovisual integration of social and non-social stimuli and have revealed sub-regions to be specific for audiovisual object and face-voice processing (Stevenson & James, 2009; Watson, et al., 2014). More specifically, Stevenson and James (2009) measured the super-additive changes in BOLD for multisensory and unisensory information and revealed different regions specific for audiovisual tool and speech stimuli within the STC. Moreover, they showed that these regions elicited identical patterns of neuronal convergence across a range of stimulus saliencies. Similarly, Watson et al., (2014) noted that the right STS contained a heteromodal people selective region (activated by face and voice), a separate region in the STS, with preference for audiovisual face-voice stimuli as compared to objects. These findings suggest a dedicated social information processing role of the STS.

1.7 Statistical criteria used to classify audiovisual brain regions

Audiovisual integrative effects can be modelled many different ways using different statistical criteria, ranging from conservative to liberal: supra-additive, max criterion and mean criterion. All criteria, however, define more activation of audiovisual stimuli

than unisensory stimuli as enhancement, where the unisensory stimulus are binding together, and less activation as suppression, assuming that stimuli are not binding together, and no difference between audiovisual and unisensory activation is interpreted as no integration. Below is a summary of the super-additivity, the max criterion and the mean criterion (for more detailed reviews see Beauchamp, 2005; Laurienti et al., 2005; Goebel & van Atteveldt, 2009; Love et al., 2011; James & Stevenson, 2012).

1.7.1 The super-additive criterion

The super additive criterion is assuming that brain regions involved in multisensory integration show greater responses to multisensory stimuli that exceeds the sum of the responses to the unisensory stimuli (i.e., Audiovisual > Audio + Visual). By employing this criterion, regions of the temporal, occipital, parietal and frontal lobes have been found to be involved in integration face- voice information (Joassin et al., 2011a, 2011b). However, this technique was adopted from electrophysiology measuring the responses of single neurons and might not be the most appropriate method when recording the BOLD activation, as BOLD activation is used to measure the blood flow to a heterogeneous group of neurons. In an audiovisual integration area, the proportion of audiovisual neurons might be small compared to unisensory neurons. While only a small proportion of these neurons might respond in a super-additive manner, and super-additive neurons have lower impulse counts compared to other neurons, the average impulse count of multisensory neurons is used to determine whether the response it super-additive (Laurienti et al., 2005). This suggests that BOLD activation may never exceed the super-additive criterion. Therefore, the super-additive criterion is overly strict and is likely to lead to false-negative errors (Beauchamp, 2005). Interestingly, as found by Love et al., (2011) the super-additive criterion can also lead to false-positive errors due to a negative response in only one of the unisensory modalities. Thus the super-additive criterion is only the strictest of multisensory criteria if the brain regions show increased activity for both unisensory conditions compared to baseline. Otherwise, regions that are defined as super-additive and multisensory only are falsely defined as those only because in unisensory condition it caused a deactivation (Goebel & van Atteveldt, 2009).This could be avoided by utilising a heterosensory contrast (audio> baseline, visual > baseline), which guarantees significant unisensory activation.

1.7.2 The max-criterion or conjunction analysis

In fMRI, a conjunction analysis is commonly used to investigate brain regions that show a significantly stronger response to audiovisual information than to unisensory information of both sensory modalities (audiovisual > audio) ∩ (audiovisual > visual). This approach has been utilised to identify, for example, the superior colliculus, which is a well-recognised multisensory structure, as well as the bilateral STC, as regions of face-voice integration (Kreifelts et al., 2010; Szycik et al., 2008). Although, qualitatively the max-criterion is less stringent than the super-additive criterion (if there is no deactivation in one sensory modality), it may lead to loss in sensitivity. This is important when two different contrasts that are predicted to have small effects are submitted to such an analysis. However, there are different ways to improve the sensitivity of the max-criterion: restricting the analysis to a smaller number of voxels by defining anatomical regions (regions of interest; ROI), or to define separate conjunction analyses for specific comparisons, for example in emotion research (audiovisual happy > audio happy) ∩ (audiovisual happy > visual happy) (Pourtois et al., 2005). Due to the level of stringency, we chose the max-criterion or conjunction analysis to define audiovisual regions in Chapter 4.

1.7.3 The mean criterion

The mean criterion defines audiovisual regions by testing for a stronger response to audiovisual information than the average of the two unisensory responses to audio and visual information, i. e., audiovisual > (audio + visual)/2. This provides an index of the degree of audiovisual integration in brain regions. It is a more liberal criterion than the super-additive criterion and the max-criterion, and therefore is able to identify presumed multisensory regions, such as the STC. However, it has been argued to be too liberal, especially when one of the unisensory responses is weak or a deactivation, as this reduces the mean of the unisensory responses. Thus it is possible for the audiovisual response to exceed the mean even when the response is weaker than the largest unisensory response. Thus, similar to the super additive criterion, the mean criterion could lead to misinterpretations. Using face-voice stimuli, Love et al. (2011) used the mean criterion and found the regions in the occipital and temporal lobe involved in audiovisual integration. At closer inspection, the response profiles to the

face-voice stimulus of those audiovisual regions showed little difference to those of the unisensory regions.

1.8 Neuroscience of temporal audiovisual processing

Several studies have investigated the brain areas involved in synchrony perception, (e.g., Calvert, Campbell & Brammer, 2000; Werner & Noppeney, 2010) however, few studies have specifically looked at audiovisual SJ tasks (Miller & D'Esposito, 2005; Stevenson et al., 2010; Stevenson, Mullin, Wallace & Steeves, 2013; Love, 2011; Love et al., in preperation) and the role of the STC is less clear when the temporal aspects of audiovisual integration is introduced. A network of regions responding more to synchronous than asynchronous speech, including the right middle STC, and bilateral superior colliculus, fusiform gyrus, lateral occipital cortex, and extrastriate visual cortex has been found (Stevenson et al., 2010). Similarly, Love (2011) examined participants' brain activation during SJ tasks on physical and perceptual synchronous (group mean PSS) as well as asynchronous (±400, ±320, ±240, ±160, ±80 ms between the audio and the video information) audiovisual speech displays. Like Stevenson et al., (2010), he showed an asynchrony network and a synchrony network, but only for perceptual synchrony (audio preceding visual information by about 90 ms) and not for physical synchrony (SOA = 0). He also distinguished two regions of the STC: a middle region of STC, responding to synchronous speech, and a posterior region, responding to asynchronous speech. Moreover, he defined the right posterior STC as a neural correlate of the fact that people are better at detecting asynchrony in audio-first (audio leading visual information) stimuli. These results reveal that investigating perceptually rather than physically defined contrasts disclose more activation for asynchronous stimuli compared to synchronous. Furthermore, Love et al. (in preparation) compared the neural mechanism underlying SJs and TOJs, using audiovisual point-light drumming displays of audio-first condition, video-first (visual leading audio information) condition, physically synchronous condition as well as a condition that showed the participants PSS. Their results showed that the two judgements use different brain areas. The middle occipital cortex was found to show sustained activation during SJ and deactivation during TOJ. Whereas, transient activation was greater in TOJ than in SJ, in regions of the left middle occipital, middle frontal, precuneus and medial superior frontal lobe. Moreover, they showed that only during TOJ the right anterior cingulate showed more deactivation to audio- and visual-

first conditions than to PSS and physical synchrony. This can be taken as evidence that the SJ and TOJ measure different aspects of audiovisual synchrony perception. These results are mainly supported by a recently published study also comparing SJ and TOJ, demonstrating that TOJ recruits additional brain regions compared to SJ (Binder, 2015).

A recent repetitive transcranial magnetic stimulation (rTMS) study by Stevenson et al. (2013) measured the contribution of the STS to audiovisual temporal processing. rTMS stimulation, prior to making SJs on beep-flash stimuli of the multisensory region (STS), caused an overall widening of the TIW (increased tolerance for visual-first stimuli). Whereas, stimulation of auditory (Hechl's gyrus) and visual (striate cortex) regions caused a broadening within the audio-first stimuli and video-first stimuli, respectively. The broadening of the TIW to the more ecologically valid visual-first stimuli with STS disruption advocates that audiovisual temporal processing in STS reflects learned environmental information.

1.9 Audiovisual processing in Autism Spectrum Disorder

The research in audiovisual integration in ASD is not entirely consistent. Several studies looking at lip reading (Smith and Bennetto, 2007; de Gelder et al., 1991) and the McGurk illusion (Irwin, Brancazio, Tornatore & Whalen, 2008) suggested abnormal audiovisual integration in ASD. For example, Smith and Bennetto (2007) revealed that individuals with ASD benefit less from additional visual information when perceiving speech and that they are worse at lip reading. They concluded that these findings could only be explained by a deficit in audiovisual integration. Irwin et al. (2011) made use of the McGurk illusion (McGurk & MacDonald, 1976) in which participants were visually presented with a video of a person saying the syllable "ga", while being simultaneously presented with the voice of the person saying a different syllable "ma". This led participants to report a fused percept such as "na". They found that even when children with ASD looked at the speakers, they reported the fused percept less often than TD children. Therefore, children with ASD were influenced less by the incongruent visual information.

Although behavioural performance of individuals with pervasive developmental disorder (PDD; a slightly broader diagnosis than ASD) was the same, electroencephalography (EEG) data argued that audiovisual integration of complex

phonological information is impaired in PDD, while low-level audiovisual integration is intact (Magnee, de Gelder, van Engeland and Kamner, 2008). In line with this idea, Mongillo et al. (2008), showed that behaviourally children with ASD performed differently from their control only in tasks involving the human speech (e.g., McGurk stimuli), but performed the same on a non-human audiovisual task for which they had to determine whether the sound of a bouncing ball matched it's physical appearance. Similarly, utilising the flash-beep illusion, Van der Smagt, van Engeland, and Kemner (2007) also showed evidence that multisensory integration in ASD is preserved. The flash-beep illusion is based on the phenomenon that auditory stimulation (beeps) can influence the perception of visual stimulation (flashes), i. e., when a single flash is presented simultaneously with two beeps, people perceive two flashes (Shams, Kamitani, & Shimojo, 2000). They reported that adults with ASD also perceived the illusion of a second flash, and that the number of second flashes reported is the same between the ASD and TD group.

However, a study by Williams et al. (2004) looked at children's complex speech perception and showed that those with ASD performed comparable in audiovisual syllable identification tasks when controlling for unisensory deficits. In their experiment the ASD group and the TD group had to identify spoken syllables. In the audiovisual condition, the audio and the visual information was either congruent or incongruent (as in the McGurk stimulus). Children with ASD performed poorer at recognising the stimuli in unisensory (visual or auditory) conditions compared to their controls. Controlling for these lower performances in the unisensory conditions, both groups performed comparably in the audiovisual conditions. This suggests that while children with ASD have difficulties in the unisensory conditions, they still show normal audiovisual integration. The authors also showed that training improved the children's ability to utilise visual information in their processing of speech.

More recently, researchers looked into the development of audiovisual integration in ASD. A cross-sectional study (Foxe et al., 2015) explored how seen and heard speech was integrated in ASD, from childhood to adolescence, when background noise was manipulated. Profound integration deficits were revealed in ASD, which were increasingly evident as background noise increased. These deficits were present in children with ASD from the age of 5 to 12 years old, but were resolved in teenage

children with ASD (13-15 year olds). The severity of the deficit in childhood and its amelioration in teenage years let the authors suggest that multisensory processing differences would be responsive to intervention in earlier childhood, with possibly great consequences for the development of social communication abilities in ASD. Interestingly, Ross, Del Bene, Molholm, Frey and Foxe (2015) highlighted the importance of considering sex differences in ASD research, as they recently revealed sex differences in audiovisual speech perception in children with ASD. More specifically they showed that girls both with and without ASD outperformed boys at recognising words under audiovisual listening conditions, however this sex difference was absent in their adult TD sample. The authors concluded that audiovisual integration is delayed in boys, compared to girls, and that in adulthood, females reach their performance maximum and males catch up.

1.10 Temporal audiovisual processing in Autism

The temporal relationship between the incoming information of sight and sound effects the way sensory information is integrated across these senses. As previously introduced, the TIW is used to measure how tolerant we are to temporal asynchrony between sound and sight and still perceive it as one event. Temporal processing has been shown to be altered in ASD (Brock, Brown, Boucher & Rippon, 2002; Szelag, Kowalska, Galkowski & Poppel, 2004). Szelag et al. (2004) showed that children with ASD had deficits in reproducing the durations of both auditory and visual unisensory stimuli. Brock et al. (2002) proposed the temporal binding hypothesis of ASD. This theory is based on the idea originally formulated by Frith (1989), and termed as weak central coherence, which hypothesises that individuals with ASD mostly focus on local rather than global aspects of information. That is, individuals with ASD perceive sensory information in isolation (e.g. a voice), rather than as a meaningful whole (e.g. a person speaking). The temporal binding hypothesis of ASD proposes that the deficits in global processing are linked to impairments in temporal processing. In other words, individuals with ASD cannot exploit the temporal correspondence of different sensory inputs to the same extent as TD individuals. This claim is supported by recent evidence showing decreased sensitivity to audiovisual asynchrony in children, adolescents and young adults with ASD in low-level and complex speech stimuli (Bebko, Weiss, Demark & Gomez, 2006; Foss-Feig et al., 2010; Kwakye, Foss-Feig, Cascio, Stone & Wallace, 2011; de Boer-Schellekens,

Eussen & JeanVroomen, 2013; Stevenson et al., 2014). This decreased sensitivity to audiovisual asynchrony was demonstrated by a broadened TIW in the ASD group compared to the TD group.

Foss-Feig et al. (2010) investigated audiovisual temporal processing in ASD by taking advantage of the beep-flash illusion and its dependency on the SOA between the flash and the beeps; the bigger the SOA between the beeps and the flashes, the weaker the illusion. Their results showed that children with ASD successfully perceived the flash-beep illusion over a wider TIW than TD controls. This finding suggests that individuals with ASD may show more extensive, but less temporally precise audiovisual integration. These findings were supported by Kwakye et al. (2011), who investigated audio, visual and audiovisual temporal acuity in children with ASD by measuring individuals' thresholds on TOJ tasks under visual, auditory and audiovisual conditions. Their multisensory task included two circles, presented successively above and below the central fixation point. The first circle was always presented simultaneously with a beep, while the second beep was presented at different SOAs (0–500 ms) after the second circle. The additional auditory information is known to increase performance compared to the visual task. However, the increase in performance is SOA dependent (Hairston et al., 2005, 2006). Whereas, no differences in thresholds for the visual TOJ task were seen between children with ASD and their controls, thresholds were higher in ASD on the auditory TOJ task. On the multisensory TOJ task, children with ASD showed performance improvements over a wider range of SOA than TD children, supporting the idea of an extended TIW. This potentially suggests that the extended multisensory is due to auditory processing differences. It would have been interesting to see whether the performance difference between the two groups would have reduced if they controlled for the auditory processing differences.

More recently, de Boer-Schellekens et al., (2013) and Stevenson et al. (2014) investigated the TIW over a whole range of audiovisual stimuli, including simple beep-flash and speech displays, as well as a complex non-speech stimulus. In both studies the audiovisual displays were presented at different SOA between the audio and visual information, thus the visual information preceded or followed the corresponding audio information at different intervals. In de Boer-Schellekens et al.'s

(2013) study, adolescents and young adults (16-24 years of age) with ASD and TD controls were asked to make TOJ, and the researcher fitted a linear function to the response data to estimate the width of the TIW and PSS. This revealed a wider TIW in ASD across all display types, but no difference in the PSS. Whereas, Stevenson et al. (2014) utilised SJ and fitted psychometric sigmoid functions to the response data to estimate the TIW width. Their results demonstrated that children with ASD (aged 6-18 years) only had an extended TIW for speech displays (a face saying: "ba" or "ga"), but not for the other, non-social and simpler displays (beep-flash and hammer-hitting-a-nail displays). Contrary to the results by Kwakye et al. (2011), Stevenson and colleagues showed that audiovisual temporal processing deficits in ASD were not accompanied by unisensory processing deficits in unisensory TOJ tasks. The researchers also showed that the ASD participants reported less often a fused percept of the McGurk illusion. Furthermore, Stevenson et al. (2014) found that wider TIWs in the beep-flash, hammer and speech displays were correlated to weaker precepts of the McGurk illusion, but only in the ASD group. This demonstrates that difficulties in audiovisual integration in ASD are associated with reduced precision of detecting audiovisual asynchrony. However, if you consider the TIW and the McGurk effect to be a measure of ability to integrate audio and visual information, the result could seem surprising. A wider TIW would suggest that the person is able to integrate audio and visual information across an extended temporal gap between the two senses, and therefore a weaker percept of the McGurk illusion at the point of synchrony might seem contradictory. Since the participant samples' ages ranged from 6 to 18 years old, it would have been interesting to see how the developmental trajectory of the TIW looks like in ASD and controls. Furthermore, the TIW reported in this sample were bigger than what is commonly reported in the literature. It would have been important to see how well their psychometric sigmoid functions fitted the response data. Previous literature regularly excluded participants from the final analysis when their response data could not be fitted by psychometric functions, as we can assume that they were unable to do the task. Moreover, the use of a unisenosry TOJ control task for their main audiovisual SJ experiment is questionable, as accumulating behavioural and fMRI evidence suggests that these tasks are actually measuring different processes (van Eijk et al., 2008; Love et al., 2013; Petrini et al., 2010; Vatakis et al., 2008; Vroomen & Stekelenburg, 2011; Love et al., in preperation; Binder, 2015).

Thus far the research is pointing towards temporal audiovisual processing differences in ASD. However, Grossman, Schneps & Tager-Flusberg (2009) have shown that adolescents with ASD perform comparably to their TD controls when doing SJs on meaningful phrases. The researchers looked at how adolescents with ASD integrate audiovisual information of meaningful phrases because it was previously suggested that the cognitive processing of meaningful phrases might be different to the processing of simple non-word syllables (Grant & Seitz, 1998). The accuracy of the onset asynchrony detection was no different between the ASD group and TD group. Grossman et al. (2009) suggested that these findings are due to the meaningful nature of the stimuli in combination with a non-distracting environment. Another reason for these findings could be the larger SOA intervals (ranging from 120ms to 500ms).

1.11 Neuroimaging evidence for audiovisual integration differences in ASD

Only a few neuroimaging studies have been conducted to investigate audiovisual integration in ASD. Much of the neuroimaging evidence for deficits in audiovisual integration comes from EEG studies. EEG records event-related potentials (ERPs), which provide a direct measure of the brain's response to incoming sensory information. ERP components are defined as component waves of the more complex ERP waveform. More specifically, ERP components are defined by their polarity (positive or negative going voltage), scalp distribution, timing, and sensitivity to task manipulations. The temporal resolution of ERPs allows for the measurement of brain activity from one millisecond to the next. This permits one to describe the response in terms of early cortical sensory registration, sensory-perceptual processing, and later cognitive stages of processing (Foxe & Simpson, 2002; Lucan et al., 2010). EEG studies have revealed differences in audio (e.g., Dunn et al., 2008; Lepisto et al., 2005) and visual (Frey et al., 2013) sensory processing, as well as decreased integration of audiovisual information (e.g., Brandwein et al., 2013; Russo et al., 2010) in ASD compared to TD participants. Furthermore, the neurophysiological indices of sensory processing differences have recently been suggested to reflect neuropathology underlying clinical symptoms of ASD. This was demonstrated by the correlation of severity of ASD symptoms (as measured by the autism observation schedule; ADOS) and neural indices of early audio processing, as well as audiovisual integration (Brandwein et al., 2015). The authors proposed that these sensory processing differences might be a strong candidate for biomarkers of the clinical ASD phenotype.

Using fMRI, Doyle-Thomas, Goldberg, Szatmari and Hall (2013) showed that adolescents with ASD employ different cortical areas when processing audiovisual emotion stimuli compared to TD adolescents. More specifically, when presented with audiovisual emotional displays and asked to match the emotions to an emotional label on the screen, both groups activated regions in the frontal and temporal lobe, however, fewer regions were activated in the ASD group than the TD group. In the frontal lobe, the ASD group showed higher activation in regions of the medial frontal gyrus and middle frontal gyrus, whereas the TD group revealed more activation in regions of the superior frontal gyrus, precentral gyrus, posterior cingulate and distinct area of the middle frontal gyrus. In the temporal lobe, the ASD participants showed higher activation in the middle temporal gyrus, whereas the TD participants revealed higher activations in the superior temporal gyrus and middle temporal gyrus. Furthermore, areas in the parahippocampal gyrus and inferior occipital gyrus were activated stronger in the TD group. Similar activation patterns have previously been shown in a pilot study (including five ASD and four TD participants) examining the neural correlates of a similar audiovisual emotion matching task (Loveland et al., 2008). They showed that TD participants had more activation compared to the ASD participants in the STC, orbitofrontal cortex, posterior cingulate, parahippocampus and occipital regions (left fusiform gyrus, and bilateral lingual gyrus extending into the left cuneus). However, since both studies (Doyle-Thomas et al., 2013; Loveland et al., 2008) employ emotional stimuli and ask the participants to make emotion judgements, it is likely that the studies assess the underlying neural correlates of emotion processing and not audiovisual processing itself. Interestingly, thus far audiovisual synchrony has never been investigated in ASD using fMRI.

1.12 Audiovisual temporal processing in other clinical populations

Understanding audiovisual temporal processing and their underlying neural correlates is also important in other clinical populations. Several researchers have revealed deficits in audiovisual temporal processing in individuals with developmental dyslexia (Hairston, Burdette, Flowers, Wood, & Wallace, 2005), schizophrenia (Fourcher et al., 2007) as well as people with synaesthesia (Neufeld, Sinke, Zedler, Emrich & Szycik, 2012; Brang, Williams & Ramachandran, 2012). More specifically, individuals with dyslexia and schizophrenia seem to have a wider TIW compared to TD individuals

(Hairston et al., 2005; Foucher et al., 2007), whereas, in synaesthesia, the literature seems more contradicting (e.g., Neufeld et al., 2012; Brang et al., 2012).

This shows that deficits in audiovisual temporal perception are a common discovery across different clinical populations, as well as in typical ageing. Therefore, understanding the audiovisual perception in clinical populations and TD individuals could enhance our understanding of the symptoms shared across these clinical populations. Furthermore, Kwakye et al. (2011) speculated that it could provide the foundation of a diagnostic tool, as well as becoming the basis of new intervention methods.

1.12.1 Audiovisual temporal processing in dyslexia

Although most of the audiovisual integration research has been done in ASD, evidence suggests that audiovisual integration differences are not unique to ASD and are also present in developmental dyslexia. Dyslexia is a disability in which affected individuals have reading difficulties, but have normal or above-normal intelligence. Both sensory and multisensory changes have been found to accompany dyslexia. Indeed, original clinical descriptions of dyslexia interventions predominantly refer to multisensory approaches (Henry, 1998).

Hairston et al. (2005) adapted the audiovisual version of the visual TOJ task. In typical participants, the introduction of a pair of task-irrelevant sounds during performance of the visual TOJ task improved performance, especially when the second sound was presented after the appearance of the second light (Morein-Zamir, Soto-Faraco, & Kingstone, 2003). Hairston et al. (2005) showed that the dyslexic individuals received performance benefits from this second sound over a much wider range of SOAs compared to the TD controls, revealing an extended TIW. It was speculated that this extended TIW could lead to profound difficulties when constructing strong reading representations, as it would cause ambiguity as to which auditory component of a written word (i.e., phonemes) belongs with which visual component (i.e., graphemes). Consequently, it would decline the speed and accuracy of reading (Hairston et al., 2005; Wallace & Stevenson, 2014). Supporting this, EEG studies have revealed that when people start to read fluently, letters and speech-sounds are integrated early and automatically in the auditory cortex, a process heavily reliant on the relative timing of the paired stimuli (Froyen, van Atteveldt, Bonte & Blomert,

2008; Froyen, Bonte, van Atteveldt & Blomert, 2009), and that this progression to early and automatic processing does not seem to take place in dyslexia (Froyen, Willems & Blomert, 2011). Furthermore, an fMRI study showed that dyslexic individuals under activate regions of the STC when integrating audio and visual components of speech stimuli (Blau, van Atteveldt, Ekkebus, Goebel & Blomert, 2009). As discussed above, STC plays an important role in audiovisual integration.

1.12.2Audiovisual temporal processing in Schizophrenia

Schizophrenia is characterised by cognitive deficits and processing abnormalities at the behavioural level of different sensory modalities (Williams, Light, Braff, & Ramachandran, 2010). As audiovisual integration requires the senses to work together in a cooperative fashion, it is likely that audiovisual integration is impaired in Schizophrenia. It has been suggested that deficits in the integration of audiovisual information in individuals with Schizophrenia are specific to social or speech displays (de Gelder et al., 2005; de Jong et al., 2009; Szycik et al., 2009). Other research has shown a wider TIW in people with Schizophrenia, compared to the TD people (Martin, Giersch, Huron, & van Wassenhove, 2013). De Boer-Schellekens et al. (2014), however, found evidence for intact audiovisual temporal perception using TOJ task. They found that individuals with Schizophrenia were less sensitive to judging the temporal order of two successively presented visual stimuli than TD controls. However, their performance improved as to the level of the control group when two accessory sounds were added (temporal ventriloquism). This suggested that individuals with Schizophrenia are less sensitive to visual temporal order, but have no deficits when integrating auditory and visual information. Evidence from EEG studies, however, supports the view that individuals with Schizophrenia have deficits in audiovisual integration. Recently, Stekelenburg, Maes, Van Gool, Sitskoorn and Vroomen (2013) reported that in TD controls, visual information that predicts the onset of a sound reduces the auditory-evoked N1, compared to the N1 elicited in their audio-only condition. However, this reduction of the N1 was absent in individuals with Schizophrenia, proposing a deficit in audiovisual temporal prediction of sound. This supports the view that individuals with Schizophrenia did not integrate multisensory stimuli as well as controls.

1.12.3 Audiovisual temporal processing in synaesthesia

Synaesthesia is the rare ability to perceive an internally generated perception in one sensory modality, triggered by an external stimulus from another sensory modality or sub-modality (Grossenbacher & Lovelace, 2001). Utilising the flash-beep illusion, as a tool to measure audiovisual integration, inconsistent findings have been reported that synaesthetes are either more (Neufeld et al., 2012) or less (Brang et al., 2012) susceptible to the illusion at short SOAs between the beeps. In other words, this means that they either integrate audio and visual information over a wider TIW or over a narrower TIW, respectively. Both these results are in contrast to Whittingham et al. (2014) and Bargary's (2008) results, reporting no differences in perception of the flash-beep illusion in people with synaesthesia and TD people, across a wide range of SOAs. The basis of the discrepancies of these findings is not entirely clear. However, the results could be due to differences of the characteristics of the synaesthetes across the studies. For example, Whittingham et al (2014) and Neufeld et al. (2012) included not only grapheme-colour synaesthetes, but also included colour-hearing synaesthetes in their sample. It might be that different types of synaesthesia affect multisensory integration in different ways. Furthermore, age differences between the samples could lead to potentially different results. It is possible, on the basis of previous findings (e.g. Hillock et al., 2011), that the ageing process has an effect on multisensory function. Indeed, when Bargary (2008) investigated age effects, they showed that older synaesthetes had a reduced susceptibility to the flash-beep illusion than the younger synaesthetes. However, this finding is inconsistent with studies of typical ageing which report an increase in multisensory integration in older adults (Setti, Burke, Kenny, Newell, 2011).

1.12.4 Audiovisual temporal processing in typical aging

Moreover, the TIW in older adults have been found to be wider, showing that they struggle to separate temporally distinct audio and visual information, as measured by SJ tasks (Chan, Pianta & McKendrick, 2014a) and the flash-beep illusion (Setti et al., 2011). Furthermore, the researchers argue that this observation cannot only be explained by age-linked decreases in unisensory detection thresholds (decline in peripheral vision or hearing). Older adults still had a wider TIW compared to younger adults, when making SJ, even when visual contrasts and auditory pip intensity of the stimuli were based on individuals' audio and visual detection thresholds. Moreover, it

has been found that audiovisual synchrony perception is less likely to be adapted in older age (Chan, Pianta & Mckendrick, 2014b).

1.13 Objectives of this thesis

The aim of Chapter 2 is to see how autistic traits correlate with cortical thickness of the brain in a typical population. Chapter 3 aims to investigate whether temporal audiovisual integration in ASD is different to that of TD participants, and whether this integration difference is dependent on the type of stimulus presented, or on the type of audiovisual synchrony task used. Moreover this chapter aims to understand the underlying process behind the atypical audiovisual integration in ASD. The objective of Chapter 4 is to investigate activation differences between the ASD and TD participants when perceiving audiovisual, auditory and visual, social and non-social displays, with a particular emphasis on investigating audiovisual sensitive areas. Chapter 5 aims to examine neural correlates of SJs in ASD and TD participants through using social and non-social audiovisual displays.

2 Cortical thickness investigation of autistic traits

2.1 Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterised by social communication and interaction deficits, as well as repetitive patterns of behaviour, interests and activities. These traits are thought to be present in a typical population. The Autism Spectrum Quotient questionnaire (AQ) was developed to assess the prevalence of autistic traits in the general population. Von dem Hagen et al. (2011) revealed a link between AQ with white matter (WM) and grey matter (GM) volume (using voxel-based-morphometry), as well as Blood Oxygen level-dependent (BOLD) response. Findings revealed no difference in GM areas associated with social cognition. Using cortical thickness analysis in the same sample of participants, this study showed that AQ scores were correlated with cortical thickness (CT) in the left temporo-occipital junction, left posterior cingulate, right precentral gyrus and bilateral precentral sulcus, in a typical population. These areas were previously associated with structural and functional differences in ASD. Thus the findings suggest, to some extent, autistic traits are reflected in brain structure - in a typical population.

2.2 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterised as a variety of deficits in social communication and interaction (DSM V), as well as repetitive patterns of behaviour, interests and activities (American Psychiatric Association, 2013). Traits of neurodevelopmental conditions, such as ASD, are said to lie on a continuum within the general population (Frith, 1991; Baron-Cohen, 1995).

Previous research has shown that the trajectory of brain volume development is different in ASD compared to typically developed individuals. Brains of new-borns with ASD tend to be comparable in volume to brains of typically developing new-borns (Courchesne, Carper & Akshoomoff, 2003; Dawson, Munson, Webb, Nalty, Abbott & Toth, 2007). The brains tend to be enlarged in early childhood (Courchesne et al., 2001; Hazlett et al., 2005; Stanfield et al., 2008). In adolescents and adults it is less clear; while some researchers have reported that this increased total brain volume is still present (e.g., Freitag et al., 2009; Hazlett, Poe, Gerig, Smith, & Piven, 2006), others have found normal total brain volume in adolescents and adults with ASD (e.g., Aylward et al., 2002; Redcay & Courchesne, 2005; review: Courchesne et al., 2007;

Hyde, Samson, Evans & Mottron, 2010). It is also elusive whether this putative increase in total brain volume is a result of grey matter (GM) volume (Hazlett et al., 2006), white matter (WM) volume (Herbert et al., 2004), or a combination of both. Findings of structural brain differences of individual brain regions in ASD are even more inconsistent with regard to the localisation and direction (increases or decreases) with in those regions (for reviews see Amaral et al., 2008; Stanfield et al., 2008).

In MRI studies, methods called voxel-based morphometry (VBM) and cortical thickness analysis (CTA) are commonly used to investigate grey matter morphometric changes in ASD. VBM gives a probabilistic measure of local GM and WM concentration (Ashburner & Friston, 2000), whereas cortical thickness analysis (CTA) directly measures cortical surface features, such as cortical thickness (CT; Jiao et al., 2010). It is important to mention that VBM conflates information about morphology, size and position (Ashburner & Friston, 2001). CTA is less susceptible to positional variance because the extraction of the cortex follows the GM surface regardless of positional variance (Kim et al., 2005). Thus CTA provides a more direct index of cortical morphology. Furthermore, since CT is measured in vertices rather than voxels it measures CT with sub-voxel precision compared to voxel-based measures (Fischl & Dale, 2000).

CT and GM volume make use of the T1-weighted signal, representing the degree of MRI visible water, which is least visible in WM, intermediately so in GM, and most visible in cerebrospinal fluid (Diwadkar & Keshavan, 2002). MRI measurements of cortical morphology can reflect neuronal loss in ageing (Salat et al. 2004) and have been validated using post-mortem histological analysis (Rosas et al., 2002), showing that CT measured by MRI strongly correlates with post-mortem CT measurements. Moreover, CT changes have been related to cognitive function (Shaw et al., 2006; Narr et al., 2007), behaviour (Anagnostou &Taylor, 2011) and activation levels (Fusar-Poli et al., 2011; Schmitz et al., 2005), suggesting that functional and structural abnormalities share a common pathophysiology. However, it needs to be noted that MRI data have neither the resolution nor the specificity to explain the relationship between estimated CT and complex cellular processes, including dendritic remodelling, cell death, synaptic pruning, or plausible encroachment from myelination (Toga, Thompson & Sowell, 2006). Other research suggests that neural bases of GM

estimates and BOLD are independent or have a more complex relationship (Kannurpatti, Motes, Rypma & Biswal, 2010; Diwadkar et al., 2011).

A recent study by Zielinski et al. (2014) examined CT from childhood to adulthood using a large mixed cross-sectional and longitudinal sample of autistic subjects and their controls. They found early accelerated growth in childhood, followed by accelerated thinning in adolescence, and decelerated thinning in early adulthood. Similarly, Osipowicz, Bosenbark & Patrick (2015) examined GM volume across the lifespan of people with ASD and controls, as well as correlated GM volume with autism severity. They showed bilateral decreases of GM volume in the ASD group in the thalamus, the cerebellum, anterior medial temporal lobes and the orbitofrontal regions. More severe ASD was associated with decreased GM volume in the prefrontal cortex, inferior parietal and temporal cortex, as well as temporal poles. No links between increased GM volume and ASD or symptom severity were found. Although this is in agreement with other research that found non relationship of GM morphology and autistic symptom severity (Langen, Durston, Staal, Palmen & van Engeland, 2007; Webb et al., 2009), others have found correlations between GM volume and autistic symptomology and severity in children (Hadan, et al., 2009; Pierce & Corchesne, 2001). Studies using CTA and VBM have reported both increased and decreased regional GM volume and CT in ASD. Some studies have reported GM volume or CT increases throughout the whole brain and in specific regions, such as the frontal, temporal and parietal regions, lingual gyrus, insular regions, precentral gyrus postcentral, cingulate gyri, caudate nucleus, hippocampus, brainstem and midbrain (Hyde, Samson, Evans & Mottron, 2010; Ecker et al., 2010; Waiter et al., 2004; Doyle-Thomas et al., 2013; Ecker et al., 2013). Others have found specific reductions in GM volume or CT in temporal and parietal regions, sensory and motor cortex, anterior cingulate, supramarginal gyrus, precentral, postcentral gyri, thalamus, corpus callosum, cerebellum, parahippocampal gyrus (Chung et al., 2005; Hadjikhani, Joseph, Snyder & Tager-Flusberg, 2006; Hyde et al., 2010; Scheel et al., 2011; Toal et al., 2010; Greimel et al., 2013; Ecker et al, 2013; Ecker et al., 2010; McAlonan et al., 2002; Wallace, Dankner, Kenworthy, Giedd & Martin, 2010; See Table 1.1 for an overview).

Traits of ASD are thought to lie on a continuum within the general population (Frith, 1991; Baron-Cohen, 1995). The Autism-Spectrum Quotient (AQ; Baron-Cohen et al., 2011), a self-administered questionnaire, is used to measure the extent of autistic traits in the general population. The AQ has also been shown to distinguish between individuals with ASD and individuals with other psychiatric disorders (Woodbury-Smith et al. 2005). Moreover, the AQ scores predict performance on tasks commonly associated with superior performance in individuals with ASD. For example, on a variety of psychometric tests, better disembedding is observed in high-AQ scores. These include: tests of block design (Stewart, Watson, Allcock & Yaqoob, 2009), the Embedded Figures Task (Grinter et al., 2009) and faster target detection in a visual search task (Almeida et al., 2010). The AQ also predicts performance in tasks that are impaired in ASD, such as inferring others' mental states from the eyes (Baron-Cohen et al., 2001), face processing as measured using the face inversion effect (Wyer, Martin, Pickup, & Neil Macrae, 2012), and spontaneous facial mimicry (Hermans, Van Wingen, Bos, Putman, & Van Honk, 2009).

Autistic traits measured by AQ have also been shown to be associated with changes in brain structure and activation patterns in typically developed brains (Iidaka et al., 2012; Kosaka et al., 2010; Geurts er al., 2013; Saito et al., 2013; Gebauer et al., 2015; Von dem Hagen et al., 2011; see Table 2.1 for an overview of studies investigating the relationship between AQ scores and brain structure). However, a recent exploration-validation study showed no association between AQ scores and brain structure, including analysis of GM volume, CT, Division Tensor Imagining (DTI; Koolschijn, Greurts, Leij & Scholte, 2015).

Looking specifically at the relationship between AQ scores and GM volume or CT, studies show that higher AQ scores correlate with smaller GM volume of the right insula and inferior frontal gyrus; larger GM volume of left middle frontal gyrus and superior frontal sulcus (Kosaka et al., 2010; von dem Hagen et al., 2011; Geurts et al., 2013; Saito et al., 2013) and reduced CT in right medial orbitofrontal cortex, postcentral gyrus, lingual gyrus (Gebauer et al., 2015). However Koolschijn et al. (2015) showed no links between AQ scores and GM volume or CT. See Table 2.1 for an overview of recent findings of studies investigating associations between autistic traits and brain structure.

Table 2.1 Studies investigating the relationship between brain structure and autistic traits measured by AQ.

Brain Region	Method	N	Age	Autistic traits measurement	Authors
Higher AQ scores correlated with: - smaller GM volume of right insula and inferior frontal gyrus	VBM	32 PDD-NOS 40 (m)	23.8 (4.2)	Full AQ	Kosaka et al., 2010
Higher AQ scores correlated with: -smaller WM volume in right posterior superior temporal sulcus -larger GM volume in left superior frontal sulcus	VBM	91 (m)	25 (5)	Full AQ	Von dem Hagen et al., 2011
Higher AQ scores correlated with: -larger GM volume of left middle frontal gyrus; - smaller GM volume in left inferior frontal gyrus central gyrus, posterior cingulate, inferior and superior parietal lobe	VBM	85 (m=53)	21.5 (2.4)	Full AQ [a]	Geurts er al., 2013
Lower AQ prosociality scores correlated with: - smaller right insula in males - reduced structural coupling of right insula with ventral anterior cingulate in males	VBM	79 (m) 56 (f)	29.4 (4.2) 28.1 (4.4)	Full AQ [a]	Saito et al., 2013
No correlation between AQ scores and sulcal subtype	SulcoGyral patterns	ASD: 51 (m) TD: 55 (m)	30.9 (8.2) 32 (7.1)	Full AQ	Watanabe et al., 2014
Higher AQ scores correlated with: -reduced CT in right medial orbitofrontal cortex, postcentral gyrus, lingual gyrus	CT	ASD: 25 (m=18) TD:26 (m=20)	28.4 (6.4) 25.2 (4.4)	Full AQ	Gebauer et al., 2015
None	VBM, CT, DTI	204 (m=105) 304 (m=155)	22.85 (1.7) 22.82 (1.73)	AQ 28 (Hoekstra et al., 2011)	Koolschijn et al., 2015
Higher AQ scores correlated with -larger volume of connectivity between the superior temporal sulcus and amygdala	DTI	30 (m= 14)	22.5 (3.0)	Full AQ	Iidaka et al., 2012

n=Number of participants (m = males), Age= Mean age, standard deviations are in brackets, AQ= autism spectrum quotient, PDD-NOS= pervasive developmental disorder not otherwise specified, VBM = voxel-based morphometry, CT = cortical thickness, TD = typically developed, GM = grey matter, WM = white matter

[a] 4-point scale of AQ

Von dem Hagen et al. (2011) investigated the WM and GM volume of the same participants using VBM. Changes in blood oxygen level dependent (BOLD) response were measured in 19 of the participants at rest, as well as when performing a Stroop task. Their results revealed that higher AQ scores were correlated with lower volumes of WM in the posterior superior temporal sulcus (pSTS); an area related to social processing such as attentional cueing from eye gaze (Bayliss & Tipper, 2005). The pSTS has also been found to have structural differences in GM (Scheel et al., 2011; Greimel et al., 2013; Doyle-Thomas et al., 2013; Hyde et al., 2010) and WM (Barnea-Goraly et al ., 2004), as well as functional differences in ASD (Gusnard & Raichle, 2001; Buckner et al., 2008). Furthermore, Von dem Hagen et al. (2011) showed that the AQ correlated with the degree of cortical deactivation in an area neighbouring the pSTS whilst performing a Stroop task compared to baseline. However, using VBM, AQ and GM volume only correlated in the left superior frontal sulcus. Contrary to their hypothesis, no correlations between AQ and GM volume were found in areas involved in social cognition and mentalising, such as the pSTS, temporal parietal junction/angular gyrus and medial prefrontal sulcus.

VBM potentially conflates information about morphology, size and position (Ashburner & Friston, 2001), while CTA is less susceptible to positional variance, providing a more direct index of cortical morphology (Kim et al., 2005; Jiao et al., 2010). Therefore, the current study made use of the semi-automatic, surface-based CTA tools in Brainvoyager to further investigate the relationship between CT and AQ in the same sample previously investigated by von dem Hagen et al. (2011).

2.3 Methods

2.3.1 Participants

91 right-handed participants were included in this study (mean age = 25 ± 5 years, range: 18-42; 53 females). 95 participants were originally recruited through the volunteer panel of the MRC cognition & Brain Sciences unit at the University of Cambridge. Four participants were excluded; two were excluded due to excessive head movement and two due to poor image intensity distributions. None of the participants reported a history of psychiatric or physical illness. For 31 participants, two structural scans were taken and were then averaged to improve image quality.

All participants completed the Autism Quotient (AQ) questionnaire developed by Baron-Cohen, Wheelwright, Skinner, Martin and Clubley (2001) (mean score: 16 ± 7, range: 2-33). The AQ contains 50 items measuring the degree of autistic traits within the general population, as well as in individuals with high functioning autism and in Asperger's Syndrome. Examples of items include: "When I'm reading a story, I find it difficult to work out the characters' intentions", and "I am often the last to understand the point of a joke." A higher AQ score indicates a greater extent of autistic traits. The total AQ has been shown to have good test-retest reliability as well as good internal consistency (Baron-Cohen et al., 2001). Moreover, the AQ has been reported to have high sensitivity and specificity in individuals referred for a diagnosis of ASD: at a cut-off score of 26, 83% of people with an ASD diagnosis were correctly identified (sensitivity 0.95, specificity 0.52), whereas a cut-off score of 32 correctly identified 76% of people diagnosed with ASD (sensitivity 0.77, specificity 0.74) (Austin, 2005; Woodbury-Smith, Robinson, Wheelwright, Baron-Cohen, 2005;). Therefore, evidence suggests that AQ is a sensitive measure of autistic traits in the general population.

One participant scored 33 on the AQ and therefore scored above the cut-off point for Asperger's and high-functioning autism (Baron-Cohen et al., 2001). However, the AQ is not a diagnostic measure, and none of the participants were reported to have a clinical ASD diagnosis. In all cases, written informed consent was obtained from all participants. The study was approved by the research ethics committee at Cambridge University (see von dem Hagen et al., 2011 for more detail).

2.3.2 Procedure

2.3.2.1 MRI Acquisition parameters

A Siemens 3T Tim Trio scanner was used to acquire the anatomical scans, and all analyses were performed in BrainVoyager QX 2.4 and 2.6. (Brain Innovation, Maastricht, The Netherlands, http://www.BrainVoyager.com). A high-resolution structural magnetisation, resulting in rapid gradient echo scans (voxel size = 1 x 1 x 1 mm, repetition time = 2250 ms, echo time = 2.99 ms, inversion time = 900 ms, flip angle = 9°, total scan time = 4 min 16 s), was acquired for all participants.

2.3.2.2 Data preprocessing

BrainVoyager QX 2.6 was used for processing all stages of the data. The structural data of all participants was converted from NIFTY files to VMR files (BrainVoyager's own file format). The structural scans' intensities were inhomogeneity corrected, the brains were extracted from the skull, and the scans were transformed into ACPC and Talairach space (Talairach and Tournoux, 1988).

2.3.2.3 Advanced segmentation analysis

Before the CTA was performed, several advanced segmentation steps had to be carried out. The data set was resampled from 1 x 1 x 1 mm to 0.5 x 0.5 x 0.5 mm iso-voxels using sinc interpolation. The "brain peeling" step and manual removal of dura was performed on some scans in which the previously performed brain extraction step was not satisfactory. The subcortical structures and ventricles were labelled as white matter and the cerebellum was manually removed. Lastly, the tissue contrast and homogeneity was enhanced using a sigma filter.

The segmentation started with the white matter-grey matter (WM-GM) border followed by the GM-cerebrospinal fluid border. The results of this automatic step

were visually inspected and manually corrected by authors blinded to which AQ score the scan belonged to. CT maps were computed using the Laplace method (Jones, Buchbinder & Aharon, 2000) implemented by BrainVoyager QX.

Since a good match between corresponding brain areas is important for group-level statistical data, analysis cortex-based alignment (CBA) was performed. It has been shown that a cortical matching approach substantially improves statistical analysis across participants by reducing anatomical variability (Fischl et al., 1999a;b; Dale et al.,1999). The cortical mapping approach by BrainVoyager QX aligns the brains using curvature information of the cortex, reflecting the gyral and sulcal folding pattern of the brain (Goebel, Staedtler, Munk, & Muckli, 2002; Goebel, Hasson, Harel, Levy, & Malach, 2004). CBA contains several steps. The input for CBA is the reconstructed cortex of a properly segmented brain hemisphere (without topological errors, e.g. "bridges", otherwise the morphing and subsequent alignment will fail). Any topological errors were manually corrected. BrainVoyager's atlas brain of Colin was included. Then, the folded cortex meshes were transformed into spherical mesh representations (for each hemisphere separately), which provided a parameterizable surface for across-subject non-rigid alignment. Each vertex on a sphere corresponded to a vertex of the corresponding folded cortex and vice versa. The spheres also contained the curvature information which was computed from the folded cortices. This curvature information was smoothed along the surface to provide spatially extended gradient information driving intercortex alignment. This minimised the mean squared differences between the curvature of a source and a target sphere. The reconstructed cortices were aligned using curvature information of the cortex, reflecting the gyral and sulcal folding pattern. Fischl, Sereno & Dale (1999) have shown that this method has been shown to reduce anatomical variability.

Voxel-wise regressions between CT and AQ were computed for every participant. Multiple comparisons were controlled for using cluster-based threshold estimation. The computation of minimum cluster threshold was accomplished via Monte Carlo simulation. After 1000 iterations, the minimal cluster size threshold that yielded a cluster-level false positive of 5% was applied to the statistical map ($p<0.05$, minimal cluster size = 1.5 cm). Moreover, the range of cortical thickness was limited to up to 7 mm and anything above was regarded as artefact (Jones et al., 2000).

2.4 Results

Making use of the semi-automatic, surface-based cortical thickness analysis (CTA) tools in Brainvoyager, we computed the CT in 91 healthy adults. The whole-brain voxel-wise correlation showed positive correlations between CT and AQ scores in the left temporo-occipital junction (r = .323, p=.0018), left precentral sulcus (r = .336, p=.0011), left posterior cingulate (r = .364 p=.0004), right precentral sucus (r=.306, p=.00032), and right precentral gyrus (r= .355, p=.0006). No negative correlations were found.

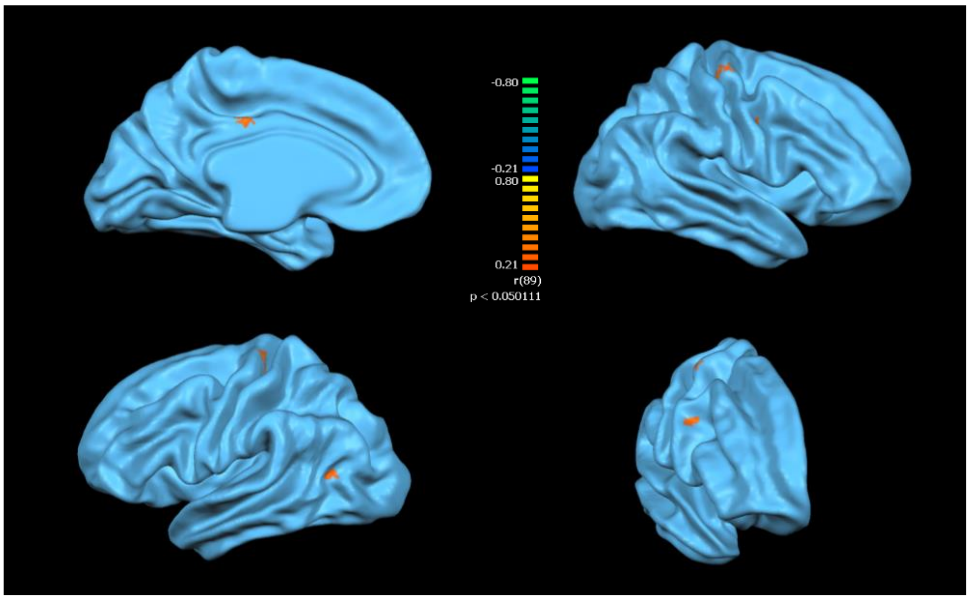


Figure 2.1 Whole brain correlation of cortical thickness and AQ. In yellow are the clusters corrected for multiple comparisons by cluster size threshold estimation to determine minimum cluster sizes for each contrast, based on a significance of p<0.05. Clusters were defined based on those that survived cluster-size threshold of 1.5mm. Clusters are projected on a surface reconstructed from the average curvature patterns from all participants and Colin's brain.

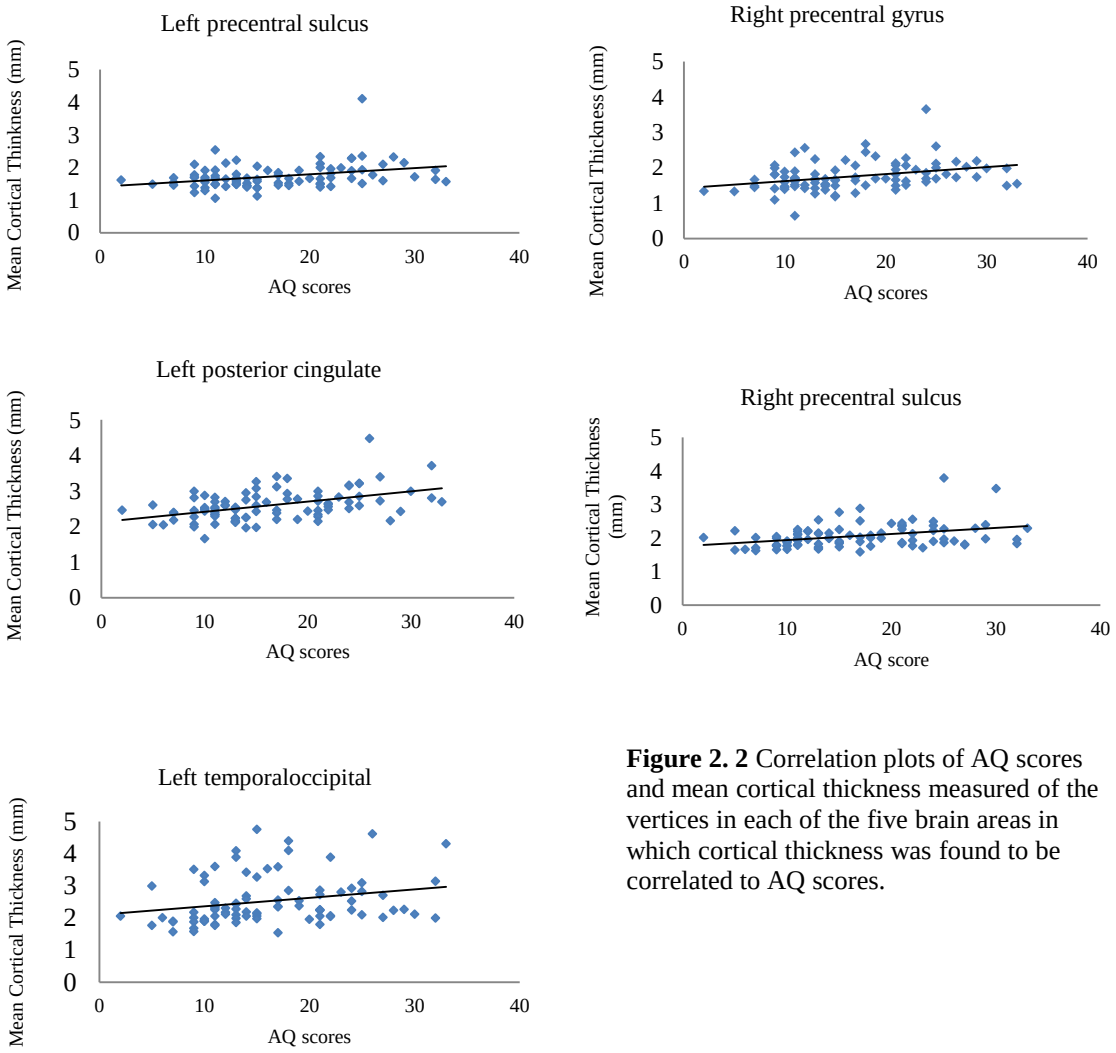


Figure 2. 2 Correlation plots of AQ scores and mean cortical thickness measured of the vertices in each of the five brain areas in which cortical thickness was found to be correlated to AQ scores.

2.5 Discussion

Making use of the semi-automatic and surface-based CTA tools in Brainvoyager, we examined the relationship between CT and autistic traits, measured by AQ, in 91 healthy adults. The whole-brain analysis revealed positive correlations between CT and AQ scores in areas previously reported to have atypical structure in ASD, including the left temporo-occipital junction, left posterior cingulate, right precentral gyrus and bilateral precentral sulcus. These findings suggest that the prevalence of autistic traits in a typical population can be associated with thickening of cortical regions. Interestingly, this study did not reveal an association between higher AQ scores and cortical thinning.

The current study supports previous findings associating higher AQ scores with areas of larger GM volume, particularly in the right posterior cingulate (Geurts et al., 2013). However, the majority of studies investigating the relationship between AQ and brain structure show that higher AQ scores are linked to a thinner cortex or smaller GM volume (Kosaka et al., 2010; Geurts et al., 2013; Gebauer et al., 2015), a relationship we fail to show in this study.

In the same sample of participants, von dem Hagen et al. (2011) previously revealed AQ to be related to GM and WM volume, as well as BOLD responses. Using VBM, the authors found that AQ and GM volume correlated in the left superior frontal sulcus, but not in areas associated with social cognition and mentalising, such as the pSTS, temporal parietal junction/angular gyrus and medial prefrontal sulcus, as they had previously predicted. Although the current study did not find correlations between AQ and CT directly in those areas, it found AQ correlates with CT in two adjacent areas; the tempoccipital area and the posterior cingulate. Furthermore von dem Hagen et al. (2011) showed that AQ correlated with the degree of cortical deactivation in an area neighbouring the pSTS, and that higher AQ scores were correlated with lower volumes of WM in the pSTS. The pSTS is an area commonly associated with structural (Scheel et al., 2011; Greimel et al., 2013; Doyle-Thomas et al., 2013; Barnea-Goraly et al ., 2004) and functional (Gusnard & Raichle, 2001; Buckner et al., 2008) differences in ASD.

Temporo-occipital area/ Angular gyrus

An area between the occipital and temporal cortex revealed a correlation between AQ scores and CT. This temporo-occipital area is located right next to the angular gyrus, which is associated with shifting of attention (Gottlieb, 2007), a characteristic that the AQ measures. Furthermore, the angular gyrus has been suggested to be involved in multisensory integration of audio and visual information (Ramachandran, Azoulai, Stone, Srinivasan, & Bijoy, 2005), which people with ASD have been shown to have deficits in (e.g., Chapters 3 and 5; Smith & Bennetto, 2007). Moreover, our results are supported by previous literature suggesting an increased GM volume and CT in individuals with ASD in areas of the temporo-occipital/inferior parietial lobule (Waiter et al., 2004; Doyle-Thomas et al., 2013 Ecker et al, 2013). However, other research has shown decreased CT and GM volume in the temporo-occipital/ inferior parietal gyrus (Chung et al., 2005; Wallace et al., 2010 ; Hadjikhani et al, 2006).

Posterior cingulate

The CTA revealed that the CT in the posterior cingulate (Brodmann area 23) was correlated with AQ scores. The posterior cingulate has been associated with social information processing, such as the processing of emotionally salient stimuli (Maddock & Buonocore, 1997; Maddock, Garrett, Buonocore, 2003). Impairments in social skills are a common characteristic in ASD and are also measured by the AQ. Thus, finding CT of the posterior cingulate to correlate with AQ scores could be linked to a social processing difficulty. Moreover, studies have shown that the abnormalities in cingulate responses during interpersonal interaction correlate with the severity of autistic symptoms (Chiu et al., 2008). Using Positron Emission Tomography (PET), Haznedar et al. (2014) found decreased metabolism in both the anterior and posterior cingulate gyri in ASD. Taken together, the our results of AQ scores correlating with CT in the posterior cingulate is supported by research associating this area with social information processing, as well as observed abnormal activation levels in ASD. In close agreement with our findings are studies showing CT and GM volume increases in ASD in the posterior cingulate (Hyde et al., 2010; Doyle-Thomas et al., 2013). Contradictory to our findings, Ecker et al. (2010) found decreased GM volume in the posterior cingulate gyrus.

Precentral gyrus & sulcus

We found a positive correlation between AQ and CT in areas of the precentral gyrus and sulcus. The precentral gryrus and sulcus are part of the primary motor cortex. Motor impairments associated with ASD are commonly observed in infants (Brian et al., 2008) and persist throughout childhood and adulthood (Hallett et al., 1993; Freitag et al., 2007). Motor abnormalities in ASD are also shown to be heritable and part of the broader ASD phenotype. More precisely, early motor delays are more commonly observed in infant siblings of children with ASD than in infants without ASD siblings (Bhat et al., 2012). Although the AQ does not tap into motor deficits, individuals with higher autistic traits might be more likely to have more motor deficits, which could potentially explain our results. Moreover, looking at the ASD literature, these areas have been found to have increased CT or GM volume in adults with ASD, compared to typically developed adults (Ecker et al., 2010, Rojas et al., 2006; Doyle- Thomas et al., 2013). However, contradictory to our findings are results showing decreased CT and GM volume in the precentral gyrus in individuals with ASD (Ecker et al., 2010; Hyde et al., 2010; Hadjikhani, Joseph, Snyder & Tager-Flusberg, 2006).

Studies investigating GM volume and CT in participants with ASD or their relationship with AQ scores show heterogeneous results. This heterogeneity of results across studies investigating cortical morphology may be due to factors such as: a variety of algorithms and techniques being used to compute CT and GM volume; differences in MRI image resolution across studies; MRI sequences (e.g., MPRAGE sequence versus ADNI sequence), as well as, sample heterogeneity. Compared to MPRAGE, ADNI sequences provide an improved contrast between GM and WM, and therefore improve the segmentation process (Jack et al., 2009). Studies have since investigated what underlies this heterogeneity by looking at the different techniques used to measure cortical morphology. For example, Hazlett et al. (2011) examined GM volume, CT and surface area (SA) in ASD and suggested that increased GM volume might be associated with increased SA rather than CT. Moreover, Raznahan et al. (2010), in a cross-sectional study in ASD, reported a altered neurodevelopmental trajectories for GM volume and CT, but not SA. These results were supported in a recent study by Ecker et al. (2013) which investigated GM volume, SA, and CT, as well as their relationship in a large sample of men with ASD and well matched typically developed controls. These results suggest that GM volume is made of SA and CT, which are measurements associated with different developmental pathways.

These pathways are likely to be controlled by different underlying neurobiological mechanisms.

Moreover, heterogeneity in ASD samples is an important discussion point in all studies investigating ASD. However, more specifically, sample heterogeneity due to differences in diagnostic methods, participants' ages IQ and sex are also likely to contribute to the heterogeneity of cortical morphology results (Anagnostou & Taylor, 2011). In particular, IQ has been found to correlate with CT (Narr et al., 2007; Choi et al., 2008) and age has been linked with GM volume (Osipowicz, Bosenbark & Patrick, 2015) and CT (Zielinski et al., 2014). A study by Sowell et al. (2007) found sex differences in CT across their sample, a difference also reported in the AQ literature (Baron-Cohen et al., 2014; Ruzich et al., 2015). Therefore, future experiments studying the association of AQ scores and structural differences in the general population should control for age, IQ and sex differences in order to better isolate this specific relationship. Moreover, using only ADNI sequence would improve the segmentation process.

The total AQ has been shown to have good test-retest reliability, as well as good internal consistency (Baron-Cohen et al., 2001). Moreover, the AQ has been reported to have suitably high sensitivity and specificity in individuals referred for diagnosis (Austin, 2005; Woodbury-Smith, Robinson, Wheelwright & Baron-Cohen, 2005). However, it needs to be mentioned that the AQ is not the only measure of autistic traits. For example, the Broad Autism Phenotype Questionnaire (BAPQ) was developed by Hurley, Losh, Parlier, Reznick and Piven (2007), while the adult Social Responsiveness Scale (SRS) was originally developed by Constantino and Todd (2005). A study by Brooke, Hopwood, Wainer and Donnellan (2011) compared these three self-report measures of autistic traits and showed that the BAPQ and SRS clearly demonstrated sex differences and had better internal consistency than the AQ.

Furthermore, Gregory and Plaisted-Grant (2013) recently suggested that using AQ scores as a substitution for ASD participants requires unverified assumptions about high-AQ scoring individuals and their relationship to individuals with an ASD. Further, research has not fully explained the endophenotypes related to ASD, and thus the AQ can only function as an approximation of these. The researchers make an important point, which should be considered in all AQ research.

Conclusion

In conclusion, the present findings provide further evidence that the autistic traits (measured by the AQ) and CT are correlated in the left temporo-occipital junction, left posterior cingulate, right precentral gyrus and bilateral precentral sulcus in a typical population. These areas have previously been associated with functions often impaired in ASD, such as social processing, attention switching and motor skills. Additionally, these areas have previously been related to have structural and functional brain differences in ASD. This supports our findings that autistic traits of individuals are reflected in the brain structure in a typical population. Moreover, the discrepancy between the results by von dem Hagen et al., (2011) and our results reveals that GM volume and CT results are not necessarily comparable. Furthermore, our results suggest that CT measurements are more sensitive to cortical grey matter differences than GM volume measurements.

3 Psychophysical investigations of audiovisual processing differences in Autism Spectrum Disorder measured using Simultaneity and Temporal Order Judgements

3.1 Abstract

The ability to integrate auditory and visual information is crucial to everyday life, and results are mixed regarding how Autism Spectrum Disorder (ASD) influences audiovisual integration. To investigate this question, we examined the Temporal Integration Window (TIW), which indicates how precisely sight and sound need to be temporally aligned so that a unitary audiovisual event can be perceived. 26 adult males with ASD and 26 age and IQ-matched typically developed males were presented with flash-beep (BF), point-light drummer, and face-voice (FV) displays with varying degrees of asynchrony and asked to make Synchrony Judgements (SJ) and Temporal Order Judgements (TOJ). Analysis of the data included fitting Gaussian functions, as well as using an Independent Channels Model (ICM) to fit the data (Garcia-Perez & Alcala-Quintana, 2012). Gaussian curve fitting for SJs showed that the ASD group had a wider TIW, but for TOJ no group effect was found. The ICM supported these results, while model parameters indicated that the wider TIW for SJs in the ASD group, compared to the TD group, was not due to sensory processing at the unisensory level, but rather due to decreased temporal resolution at a decisional level of combining sensory information. Furthermore, when performing TOJ the ICM revealed a smaller Point of Subjective Simultaneity (PSS; closer to physical synchrony) in the ASD group than in the TD group.

3.2 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterised by a variety of deficits in social communication and interaction, as well as repetitive patterns of behaviour, interests and activities (American Psychiatric Association, 2013). Recent reports show that 1 in 88 children in USA have ASD (Centers for Disease Control and Prevention, 2012), revealing the pressing need to understand this condition better. In addition to the aforementioned features of autism, scientific and clinical research using questionnaire and sensory discrimination methods has repeatedly described differences in sensory processing between ASD and typically developing (TD) children (Lane, Young, Baker, & Angley, 2010). Robertson & Simmons (2013) revealed a strong correlation between autistic traits and sensory sensitivities in the general population. Yet, only recently the relevance of these sensory impairments as diagnostic criteria of ASD has been recognised as reflected in their inclusion in the DSM-V. This highlights the importance of developing sensory processing interventions.

The recent scientific interest in ASD sensory perception and behaviour has coincided with a shift in cognitive neuroscience to try to explain human perception and behaviour by examining multisensory perception, rather than each of the senses separately (Love, Pollick & Petrini, 2012). The ability to behave appropriately in the environment and to conduct everyday tasks relies on the brain's ability to decide which sensory information should be combined and which should be kept separated. Combining multiple sensory cues can reduce uncertainty and enhance our ability to make better estimates of the situation (Ernst & Banks, 2002). For example, when crossing a road we are most likely using both sight and sound to estimate the position of approaching cars. Similarly, in a crowded and noisy environment, we can better understand another person's speech by looking at his/her face and lip movements. Accumulating evidence highlights that in ASD the efficiency gained from processing multiple sensory signals as a single percept could be lost, resulting in less efficient sensory processing overall. For example, people with ASD perceive audiovisual illusions such as the McGurk effect (McGurk & MacDoland, 1976) less often than their TD controls (de Gelder et al., 1991; Irwin et al., 2011; Mongolli et al., 2008),

benefit less from information from an additional sensory modality (Smith & Bennetto, 2007), rely more on one sensory modality (Stevenson et al., 2014), and show less effective neural integration during audiovisual tasks (Brandwein et al., 2013). Interestingly, Foxe et al. (2015) recently showed that multisensory processing differences ameliorate in teenage years.

These findings are in line with the temporal binding hypothesis of ASD (Brock, Brown, Boucher & Rippon, 2002). This theory is based on the idea originally formulated by Frith (1989), termed as weak central coherence, that individuals with ASD mostly focus on local rather than global aspects of information. That is, individuals with ASD perceive the sensory information in isolation (e.g. a voice) rather than as a meaningful whole (e.g. a person speaking). Different internal and external factors can determine whether two sensory cues would be combined in a meaningful whole. Meanwhile, the temporal binding hypothesis of ASD proposes that the deficits in global processing are linked to impairments in temporal processing. In other words, individuals with ASD cannot exploit the temporal correspondence of different sensory inputs to the same extent as TD individuals. This claim is supported by recent evidences showing decreased sensitivity to audio-visual asynchrony for individuals with ASD (Bebko, Weiss, Demark & Gomez, 2006; Foss Feig et al., 2010; Kwakye, Foss-Feig, Cascio, Stone & Wallace, 2011; de Boer-Schellekens, Eussen & JeanVroomen, 2013; Stevenson et al., 2014). Sensitivity to asynchrony has commonly been measured using video clips of simple beeps and flashes, complex audiovisual human actions and audiovisual speech (de Boer-Schellekens, Eussen & JeanVroomen, 2013; Stevenson et at., 2014). Participants are presented with these stimuli at different stimulus onset synchronies (SOAs) and are asked to make Synchrony Judgements (SJ) (Grossman, Schneps & Tager-Flusberg, 2009; Stevenson et at., 2014) or Temporal Order Judgement (TOJ) (e.g., de Boer-Schellekens, Eussen & JeanVroomen, 2013). In SJs, participants are asked to judge the synchrony between the audio and the visual information, whereas in TOJs they are asked to determine whether the auditory or the visual information came first.

Unisensory temporal processing differences in ASD have also been found in audio and vision (Kwakye et al., 2011; Szelag, Kowalska, Galkowski & Poppel, 2004). Szelag et al. (2004) showed that children with ASD had deficits in reproducing the durations of both auditory and visual unisensory stimuli. Moreover, Williams at al.

(2004) found that when controlling for unisensory processing abnormalities, the audiovisual processing differences are eliminated. Conversely, Stevenson et al. (2014) showed that audiovisual temporal processing differences were not due to unisensory processing differences.

Thus far, most studies show that children and adolescents with ASD have a wider audiovisual temporal integration window (TIW), which implies that they are less sensitive to audiovisual asynchrony than their age-gender-IQ-matched controls. However, it is unclear whether this reduced sensitivity in ASD persists later in life. One study included a few young adults up to the age of 24 in their sample (de Boer-Schellekens, Eussen & Vroomen, 2013). Studies have shown that adolescents and adults with ASD often develop compensatory strategies and eliminate behavioural differences in perceptual tasks (McKay et al., 2012; Fox et al., 2015). Thus it is essential to investigate whether the lower audiovisual sensitivity still persists in adults with ASD. It is also of importance to understand how audiovisual temporal integration differs in ASD, and therefore we included two of the most commonly used tasks, TOJ and SJ, as well as different audiovisual stimuli, ranging from simple beep and flashes, complex human action to complex speech. Due to the inherent task-related differences between SJs and TOJs (Love et al., 2013; Love et al., in preparation; Binder, 2015), these are a useful mean to examine the reasons behind the audiovisual temporal binding differences in ASD. If a wider TIW in ASD is due to difficulties in processing of global information (i.e., difficulties in combining the audio and the visual cues), then one would expect to see a more pronounced performance difference between ASD and TD in SJs compared to TOJs. This is because SJs require estimation of the temporal correspondence of the audio and visual cue, and thus depend on more global level processing (considering the stimulation as a whole). TOJs, however, could in principle be performed by focusing on only one sensory cue to detect whether it came first or not, thus depending on more local level processing (i.e., considering only the sound).

To investigate the underlying perceptual processes of temporal audiovisual integration in ASD, the study is taking advantage of the Independent Channels Model (ICM) (Garcia-Perez & Alcala-Quintana, 2012), which provides estimates of sensitivity to asynchrony in SJs and TOJs across a range of stimuli, as well as their estimates of unisensory and decisional factors needed to make those judgements.

3.3 Methods

3.3.1 Participants

Twenty-six high-functioning adults with Autism Spectrum Disorder (aged between 18 and 40) and 26 age-, sex- and IQ-matched control participants (aged between 18 and 39) took part in the study (Table 3.1). All participants in the ASD group reported to have a diagnosis of having an ASD according to DSM-IV criteria from a qualified clinician. All were native English speakers, had normal or corrected to normal vision and reported no hearing difficulties. The Autism Quotient (AQ), a 50 item autism traits questionnaire developed by Baron-Cohen, Wheelwright, Skinner, Martin and Clubley (2001), with the cut off score for Asperger's being 26, supported the diagnoses of the ASD group (M= 36.64, SD = 8.80) and reinforced the assumption that no-one in the TD group had an ASD (M=12.57, SD= 3.70). The participants were matched pair-wise on age (t(50)=.448, p=.656) and group-wise on full scale IQ (FSIQ) (t (50)-.557, p=.580) as measured using the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999).

Group	Age		FSIQ	
	Mean	SD	Mean	SD
ASD	26.62	7.01	117.54	11.14
Control	25.81	5.93	119.08	8.63

Table 3.1 Mean and Standard Deviation of the ages and Full Scale IQs of the ASD and TD group separately.

The experimental procedures were approved by the School of Psychology at the University of Glasgow and also the Greater Glasgow and Clyde National Health Service ethics board.

3.3.2 Stimuli

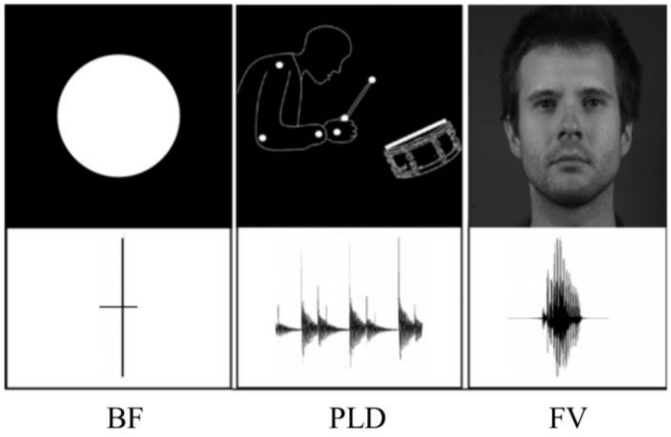


Figure 3.1 The top panel shows the visual information participants were presented with. The bottom panel shows the auditory waveform for each type of stimulus. The beep-flash (BF) stimulus consisted of a flash of a white dot on a black background and a beep. Point-light-drumming (PLD) shows a movie frame and the waveform drumbeat. The outlines of the drum and drummer are for illustrative purposes only. In the Face-voice (FV) stimulus a movie frame is shown and the waveform represents the word "tomorrow". Please note that the images are not to scale, the area of the point-light-drummer and the white flash dot are approximately the same size as the area of the mouth in FV.

Three stimulus types were used: beep-flash (BF), point-light-drumming (PLD) and face-voice (FV). These three different stimuli were used because they varied in the amount of social information and complexity. While the BF stimuli are very simple and do not have a strong prior (formed through experience) and do not contain social information, FV speech stimuli are much more complex and are based on prior social experience and situations. PLD stimuli, in contrast, are similar to BF in the way of presentation (white dots on a black background), while representing a more complex human action. In other words, they nicely collocate themselves for complexity and level of social information between BF and PLD. For more detailed descriptions of these stimuli see Love et al. (2013).

For the BF stimuli, the beep was a pure tone with 2000 Hz and 84 dB mean intensity, and the flash was a white dot with a luminance of 85 cd/m^2 presented on a black background with a luminance of 0.12 cd/m^2 (see Figure 3.1 for an illustration). The size of the white dot (with a visual angle of the diameter being 4.4 degrees) was of the same size as the of the drummer and the speaker's mouth in the PLD and FV displays, respectively. To produce the BF audiovisual movies (60 Hz), the pure tone and white

dot were imported in Adobe Premiere 1.5. The duration was cut to 33 ms with a Stimulus Onset Asynchrony (SOA) level of 0 ms. The audio and visual timelines were separated in 4 frame increments to create 11 SOA levels: 5 audio-leading (-333, -267, -200, -133, -67 ms), 5 video-leading (+333, +267, +200, +133, +67 ms) and 1 synchronous. In the ten asynchronous conditions, the space between the beep and the flash was filled with a black screen and no sound. The synchronous condition was 33 ms long. The duration of the asynchronous conditions increased with increasing SOA such that 67, 133, 200, 267, 333 ms SOA conditions were 100, 166, 233, 300, 366 ms long, respectively.

The PLD displays have been used and described previously by Love et al. (2013), Petrini, Holt & Pollick (2010), Petrini et al. (2009) and Petrini, Russell and Pollick (2009). The stimuli were dynamic audiovisual displays (60 Hz) of a point-light drummer (Figure 3.1) drumming a swing groove at 120 beats per minute with an accent on the second beat. The image of the drummer covered a visual angle of 4.8 degrees width and 2.8 degrees height. All PLD stimuli were cut from a 15 s long original recording and contained 9 audio and visual impacts (Petrini et al., 2009). The audio and visual information of the longer drumming sequence were first separated in time by each SOA level (333, 267, 200, 133, 67, 0 ms), and then the stimuli sequence was cut from that. This enabled the creation of equally long asynchronous stimuli (3 seconds) and made it possible to have an audio and video sequence at the beginning and end at all SOAs.

The FV stimuli were dynamic audiovisual displays (25 Hz) showing a native English male saying the word ‘‘tomorrow’’. The visual speech stimulus contained the full face and covered an approximate visual angle of 12.7 by 18.9 degrees (Figure 3.1), and the mouth region had a visual angle of approximately 3.2 by 2.5 degrees. The asynchronous conditions were produced by separating the audio and visual streams along the movie timeline using a method similar to that described by Vatakis and Spence (2006). This separation created gaps at the beginning and end of the movie timeline and these were filled with the first and last frame of the auditory or the visual stream in order to have a non-speaking still face image. Previous research (e.g., Van Wassenhove, Grant & Poeppel, 2007; Stevenson, Altieri, Kim, Pisoni & James, 2010) looking at speech displays used a wider range of SOA levels than that of the BF and PLD displays described above. Therefore, a wider range of SOAs was used for the FV

displays. Just as in the BF and PLD displays, ten asynchronous versions were produced for the FV, but the audio stream was shifted to either begin before the video stream (-400, -320, -240, -160, -80 ms) or after (+400, +320, +240, +160, +80 ms), in 80 ms (2 frames) increments. The synchronous condition was 1.6s long, and similar to the BF displays, the durations of FV displays became longer with increasing SOA levels, with the 400ms SOA condition lasting 2 seconds.

Part of this study's aim was to see whether audiovisual integration in ASD is dependent on the type of stimulus presented. These stimuli were chosen as they represent a variety of the types of stimuli generally used in audiovisual synchrony perception research. The stimuli chosen also ranged in complexity: the BF is a simpler stimulus than PLD, containing a point-light representation of the natural motion and FV, which contains the audiovisual information of a natural video recording of the talking human face, and can be described as the most complex. The complexity of a stimulus has been shown to cause differences in PSS and TIW (e.g. Vatakis & Spence, 2006a; Petrini et al., 2009; Arrighi, Alais & Burr 2006).

3.3.3 Apparatus and Procedure

Stimuli were presented via an Apple Macintosh MacPro 3.1 desktop computer running OS 10.5 and an NVIDIA GeForce 8800GT video card. The visual cues were displayed on a 21-inch ViewSonic Graphics Series G220f CRT monitor running at 1024 X 768 screen resolution and 60Hz refresh rate. Auditory cues were presented through high quality headphones (Bayerdynamic DT770). Presentation was achieved using MATLAB 2007b (MATHWORKS Inc., Natick, MA) and the Psychophysics Toolbox (PTB3) extensions (Brainard, 1989; Pelli, 1997). The experiment was split into 3 sub-experiments, one for each stimulus type. The order of these was pseudo randomised for each participant, with an attempt to have a similar number of the six possible order of stimuli presented. The order BF, FV, PLD and FV, PLD and BF were completed each by 5 participants in each group, the other 4 possible orderings of sub-experiments were each completed by 4 participants in each group. The participants were allowed breaks between the sub-experiments and all in all the experiment took a minimum of 1 hour and 15 minutes with BF taking ~15 min, PLD taking ~ 25 min and FV taking ~20 min.

Each sub-experiment presented one stimulus type and consisted of 24 blocks: half of the blocks were SJ blocks and the other half were TOJ and they were presented in a randomised order. After each sub-experiment participants completed a debrief questionnaire, which asked them to rate the difficulty of the two tasks by circling one of five answers ranging from easy to very difficult. More specifically participants were asked: “Please rate how difficult you found the Synchrony Judgement task. Please circle your choice”, and “Please rate how difficult you found the Temporal Order Judgement task. Please circle your choice”, with the choices ranging from: 1, easy; 2, not very difficult; 3, somewhat difficult; 4, difficult; 5, very difficult. In case participants gave the two tasks the same difficulty rating, the questionnaire also included a forced choice question: “Which task did you find more difficult?”.

The experiments took place in a quiet and dimly lit room. The viewing distance from the monitor displaying the stimuli was approximately 90 cm. At the start of the experiment, the participants read through the instructions and before each sub-experiment they had the chance to complete 3 practice trials of SJs and TOJs and ask any questions to clarify the experiment. The experimenter then left the room and the participants began the experiment by pressing any key. Task instructions, telling the participants whether the block that followed was a SJ or a TOJ block, appeared on screen for 4 seconds for every block. Within a block there were 11 trials: one presentation of each SOA level of the current stimulus type. Participants could only make a response once they had watched the entire stimulus. Therefore participants could base their SJs and TOJs on the entire stimulus duration. After each stimulus the current task question and possible responses were displayed on screen until the participant responded, which triggered the next trial. During SJ blocks participants were asked to press ‘1’ on the keyboard when they believed the audio and visual cues were synchronous and ‘2’ if they perceived them as being asynchronous. During blocks of TOJ they were asked to press ‘1’ if they perceived the video first and ‘2’ if they believed the audio came first. Feedback was never given. Participants were presented with 11 trials per SOA level for each combination of task and stimulus type. This is a similar number of trials used in previous research (Vatakis & Spence, 2006 a b) and Petrini et al. (2010) showed that results are comparable when 10 or 20 trials are used per SOA level.

3.4 Results

In the current study, a group of participants diagnosed with ASD (N=26) and their age, sex and IQ- matched TD controls (N=26) made either SJs or TOJs, in separate blocks, to three different audiovisual display types, ranging from simple beep-flash stimuli, biological motion stimuli and speech stimuli (Figure 3.1) that were presented in separate experimental runs. Furthermore, this study used traditional Gaussian fits as well as an adapted Independent Channels Model (ICM) by Garcia-Perez and Alcala-Quintana (2012) to fit the response data and provide estimates of PSS and TIW width as well as parameters describing unisensory and decisional factors.

3.4.1 Fitting Gaussian and cumulative Gaussian functions to the response data

The data of all combinations for tasks (SJ, TOJ) and stimulus (BF, PLD, FV) from each participant were separately fitted with a psychometric function. For SJs, a Gaussian probability density function was fitted to the proportion of synchronous responses at each SOA level, while for TOJs a cumulative Gaussian distribution function was fitted to the proportion of video first responses. The TIW width and PSS obtained from Gaussian fits will be referred to as TIWGF width and PSSGF.

		SJ			TOJ		
		BF	PLD	FV	BF	PLD	FV
ASD							
	N	26	26	25	24	10	21
	Excluded (%)	0	0	3.85	7.69	61.54	19.23
	Mean TIW_{GF} (ms) [s.e.m]	174.13 [9.57]	160.11 [9.62]	209.21 [12.26]	146.61 [17.75]	249.92 [88.99]	313.30 [42.58]
	Mean PSS_{GF} (ms) [s.e.m]	49.03 [8.88]	54.76 [6.26]	-13.37 [16.85]	-48.58 [15.92]	-60.13 [19.03]	-120.48 [39.59]
TD							
	N	26	26	26	25	8	24
	Excluded (%)	0	0	0	3.85	69.23	7.69
	Mean TIW_{GF} (ms) [s.e.m]	152.77 [7.95]	128.44 [5.94]	171.56 [8.12]	112.38 [10.74]	192.31 [54.82]	236.83 [17.98]
	Mean PSS_{GF} (ms) [s.e.m]	49.31 [7.65]	51.52 [4.77]	1.64 [11.09]	-35.39[12.40]	-39.23 [36.31]	-77.72 [15.47]

Table 3.2 ASD= autism spectrum disorder, TD= typical developed, TIW_{GF}= temporal integration window, PSS_{GF}= point of subjective simultaneity, N= number of participants included in the analysis, s.e.m= standard error of mean

Mixed-effects analysis of variance (ANOVA) tests were conducted on mean TIW_{GF} width and PSS_{GF} data (Table 3. 2, Figure 3.2) independently for SJs and TOJs. Significant main and interaction effects were identified at the $p<0.05$ level, using Greehouse-Geisser sphericity correction when appropriate. The TIW_{GF} width is derived from the standard deviation of fitted functions, and therefore measures the

sensitivity of task responses to changes in SOA, i.e., narrow TIW_{GF}s represent higher sensitivity to deviation from perceived audiovisual synchrony. Examining the individual fitted data indicated that some participants could not successfully make TOJs for BF, PLD and FV and one ASD participant could not successfully do SJs for FV. R^2 was calculated to indicate the goodness-of-fit between data and the fitted function. R^2 values below 0.5 were regarded as indicating that participants were unable to achieve a task/stimulus combination (e.g., SJ/FV, TOJ/PLD etc.). This criterion was applied to the data of each participant and task/stimulus combination separately, and each data set with R^2 below 0.5 was excluded from the group analysis (for similar exclusion criteria see: Love et al., 2013; Petrini et al., 2010; Boenke, Deliano & Ohl, 2009; Zampini, Shore & Spence, 2003 a b).

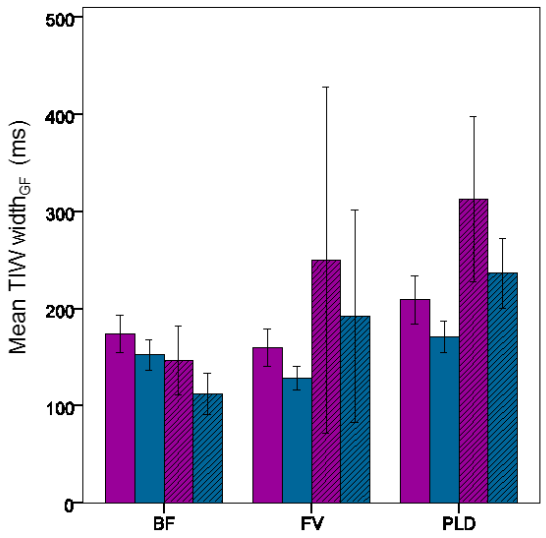


Figure 3.2 a Group-averaged TIW width (ms) optioned from Gaussian fits in ASD and TD for 3 stimulus types (BF, PLD, FV). Error bars represent standard error.

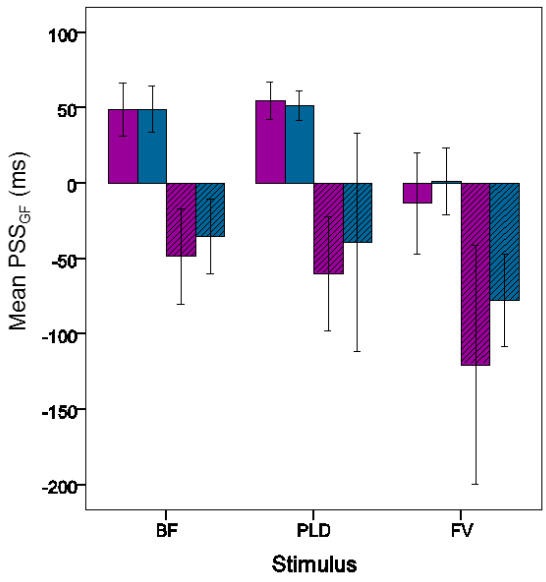


Figure 3.2 b Group-averaged PSS (ms) optioned from Gaussian fits in ASD and TD for 3 stimulus types (BF, PLD, FV). Error bars represent standard error.

A 2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVA was run on TIW_{GF} widths from the SJ tasks (Figure 3.2a). The ANOVA revealed a main effect of group ($F(1, 49)= 8.38$, $p=.006$ $\eta^2p= .146$) and a main effect of stimulus ($F(1.78, 87.61)=23.20$, $p< .001$, $\eta^2p= .529$), but no group by stimulus interaction ($F(1.7,87)=.977$, $p=.372$, $\eta^2p=.020$). Mauchly's Test of Sphericity indicated that the assumption of Sphericity were violated ($\chi^2(2) = 6.06$, $p = .048$.), thus the degrees of freedom were adjusted using Greenhouse-Geisser adjustments. For the SJ task, the ASD group had a wider TIW_{GF} (M= 180.82) in general than the control group (M=150.93). Post-hoc Bonferroni corrected t-tests revealed that FV TIW_{GF} (M= 190.39) was significantly wider than BF and PLD TIW_{GF} (M= 162.46, $p = .002$ M= 144.76, $p<.001$, respectively) and BF TIW was significantly wider than PLD TIW_{GF} ($-p =.015$).

A 2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVA on PSS_{GF} revealed a main effect of stimulus ($F(1.60, 78.23)=25.38$, $p<.001$, $\eta^2p=.34$), but no significant main effect of group ($F (1,49)=.20$, $p=.66$, $\eta^2p=.004$) or interaction effect ($F(1.60,78.23)=.594$, $p=.518$, $\eta^2p=.012$). Post-hoc Bonferroni corrected t-tests showed that PSS for BF (M=48.12) and PLD (M=53.99) did not differ, but that the PSS_{GF} for both displays was significantly greater than that of FV (M=-5.87, both $p<.001$). Two-tailed one-sample t-tests were run to show whether the mean PSS values for each display type were different from zero, i.e., physical synchrony. The one-sample t-tests showed that the mean PSS for BF and PLD were significantly different from zero (both $p<.001$), while the PSS for FV was not different from zero ($p=.569$)

As TOJ on PLD displays lead to a high proportion of participants being excluded (61.54 % of ASD and 69.23% of TD participants), this condition was analysed separately. A Mann-Whitney U test was performed in the TOJ PLD data as the Shapiro-Wilk test revealed that the TIW width data was not normally distributed for the ASD group ($W(10)=.800$, $p=.005$). This showed no group difference, $U=39.00$, $p= .929$. An independent t-test did not reveal any PSS difference in ASD and TD for the PLD stimulus ($t(16)=-.54$, $p=.597$). Due to the high exclusion rate for TOJ PLD, the means shown in Figure 3.2 a and b need to be viewed with caution, as they are likely to be a biased representation of the group averages.

A 2 (ASD, TD) x 2 (BF, FV) mixed-effects ANOVA on an TIW_{GF} width (Figure 3.2 a) obtained from TOJs revealed no main effect of group, $F(1,41)=2.69$, $p=.109$,

η^2p=.062, no interaction effect, F(1,41)=.699, p=408, η^2p=.017 but a main effect of stimulus, F(1,41)=49.33, p<.001, η^2p=.546, with only the TIW_{GF} width being smaller for BF (M=124.58) than for FV (M=268.81). For PSS_{GF} (Figure 3.2 b) the 2 x 2 ANOVA revealed no main effect of group, F(1,41)=.89, p=.351, η^2p=.022, no interaction effect, F(1,41)=.932, p=.340 η^2p=.022, but a main effect of stimulus, F(1,41)=6.74, p=.013, η^2p=.141, with BF (M=-47.98) having a less negative PSS_{GF} than FV (M=-96.82).

3.4.2 Fitting an adapted independent channels model to the response data

Fitting Gaussian functions to SJ and TOJ response data is argued to be not the best option, since they are symmetric and smooth, whereas an individual's proportion of synchronous responses in SJ are known to generally be asymmetric. Similarly, Video First responses of an individual in TOJ often show a pronounced plateau midway along the range of SOAs. Once the data averaged across individuals, these asymmetries and irregularities are likely to be averaged out, however, information might be lost. Garcia-Perez & Alcala-Quintana (2012) adopted the Independent Channels Model (ICM) to enable a more flexible to fit to the response data of SJ and TOJ.

Therefore, SJ and TOJ data were also fitted to using ICM, model-based psychometric functions. The data of SJ and TOJ were fitted jointly to the model as well as separately. The ICM provides estimates of sigma and theta, where sigma is the distance between the 15.87% and the 84.13% points, so half of this value would be the standard deviation if a cumulative Gaussian would be fitted. Theta is the 50% point on the psychometric function for "audio first" judgements if the observer had infinite resolution. Therefore, sigma and theta are somewhat comparable to the TIW width and PSS outcome measures we obtained through the psychophysical fits, respectively. So, for simplicity, we will refer to them as TIW_{ICM} width and PSS_{ICM}. The ICM also provides parameters related to sensory and decisional factors of audiovisual processing.

Delta is the onset, Lambda is the rate parameter and Tau is the processing delay of the corresponding sensory information. Lambda Audio (A), Lambda Visual (V) and Tau describe the arrival latency in SJ and TOJ tasks, and Delta is the resolution parameter which limits the observers' ability to detect small differences in arrival latencies. TOJ

included an additional response bias parameter called Xi, taking into account the tendency of participants to respond “Audio First” or “Video First” more often.

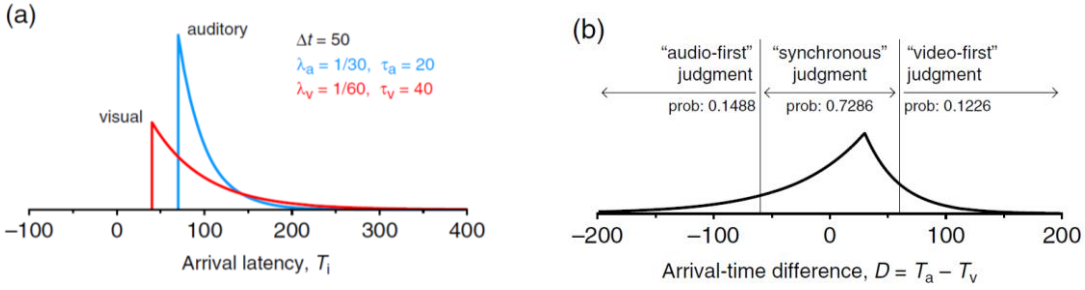


Figure 3.3 Model of timing judgments. a Exponential distributions for the arrival latency of a visual stimulus (red curve) presented at time 0 and an auditory stimulus (blue curve) presented at time Δt 0 = 50 ms. Parameters as indicated in the inset. b Bilateral exponential distribution of arrival-time differences and cutpoints on the decision space (vertical lines, at $D = \pm\delta$ with $\delta = 60$), determining the probability of each judgment (taken from Garcia-Perez & Alcala-Quintana, 2012)

3.4.2.1 Joint fit of SJ and TOJ data

Examination of individuals’ data for each stimulus indicated that, for some participants, the data could not be successfully fitted to the model. The exclusion criterion was a significant result of the Chi-square test (i.e., p<.05), as this indicated that the model had been rejected. In those cases, visual inspection also clearly showed a bad fit (for an example see Figure 3.5). These cases were excluded from the group analysis (see Table 3.3 for exclusion rates).

		SJTOJ		
		BF	PLD	FV
ASD				
	N	25	25	25
	Excluded (%)	3.85	3.85	3.85
	Mean TIW_{ICM} (ms) [s.e.m]	187.12 [22.92]	121.43 [16.38]	191.16 [22.36]
	Mean PSS_{ICM} (ms) [s.e.m]	60.93 [25.59]	51.67 [4.95]	-16.17 [15.95]
TD				
	N	26	24	25
	Excluded (%)	0	7.69	3.85
	Mean TIW_{ICM} (ms) [s.e.m]	154.39 [19.49]	81.42 [7.03]	160.37 [12.98]
	Mean PSS_{ICM} (ms) [s.e.m]	36.42 [18.51]	48.05 [4.92]	1.49 [8.50]

Table 3.3 ASD= autism spectrum disorder, TD= typical developed, N= number of participants included in the analysis, s.e.m= standard error of mean

A 2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVA on TIW_{ICM} (see Figure 3.4 a) showed a main effect of stimulus, $F(2,90)= 14.37$, $p < .001$, $\eta^2p=.242$) a marginally significant main effect of group $F(1,45)= 3.55$, $p= .066$, $\eta^2p=.073$, with ASD (M= 168.04) having a wider TIW_{ICM} than TD (M=132.44), but no interaction ($F(2,90)=.210$, p=811, $\eta^2p=.0050$. The Bonferroni corrected pair-wise comparisons showed that TIW_{ICM} of BF (M=174.61) and FV (M=174.18) were comparable and that they were both significantly wider than that of PLD (M=101.92; both p<.001).

PSS_{ICM} data (see Figure 3.4 b) showed a main effect of stimulus, $F(1.40, 63.16)= 9.01$, p< .001, $\eta^2p=.167$, but no main effect of group, $F(1,45)= 0.048$, p= .83, $\eta^2p=.001$ and no interaction ($F(2,90)=1.14$, p=.326, $\eta^2p=. 025$). Bonferroni adjusted pairwise comparisons revealed that the PSS_{ICM} for BF (M=45.84) and PLD (M=49.61) were similar, but that both were significantly larger than FV (M=-9.79), p=.015, p<.001, respectively. Two-tailed one-sample t-tests revealed that the mean PSS_{ICM} obtained from BF and PLD were significantly different from zero (both p< 0.01), while the PSS_{ICM} of FV was not (p=.420).

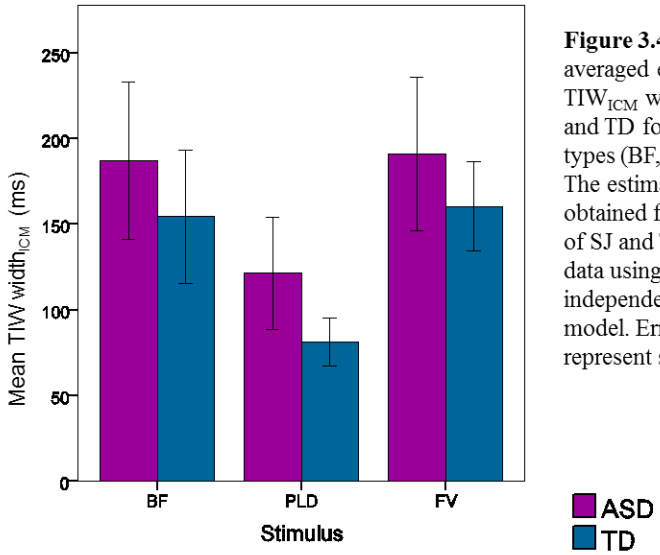


Figure 3.4 a Group-averaged estimates of TIW_{ICM} width (ms) ASD and TD for 3 stimulus types (BF, PLD, FV). The estimates were obtained from a joint fit of SJ and TOJ responds data using an independent channels model. Error bars represent standard error.

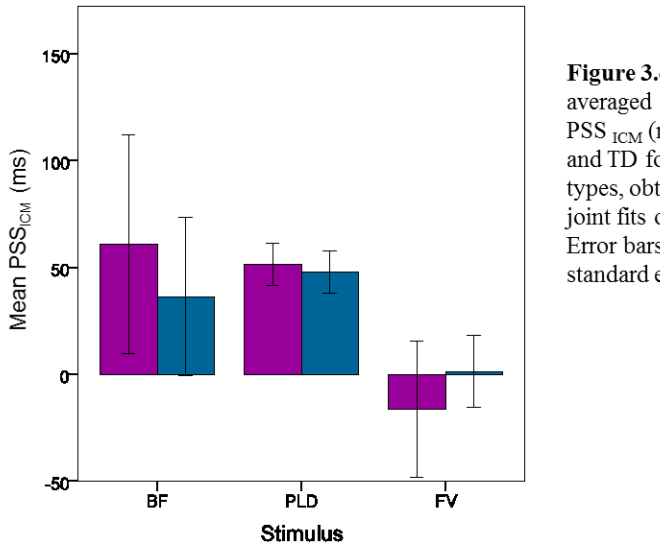


Figure 3.4 b Group-averaged estimates of PSS_{ICM} (ms) in ASD and TD for 3 stimulus types, obtained from joint fits of SJ and TOJ. Error bars represent standard error.

3.4.2.2 Fitting SJ and TOJ data separately

The ICM was also fitted separately to SJ and TOJ responses to see how each task influences audiovisual integration in people with ASD and their controls. We excluded individual data for each stimulus and task combination when the significant Chi-square test (i.e., $p<.05$) indicated that the data could not be successfully fitted to the model. In those cases, visual inspection also clearly showed a bad fit (for an example see Figure 3.5). We also excluded cases with impossibly wide TIW and big PSS (see Table 3.4 for detail). These cases were excluded from the group analysis.

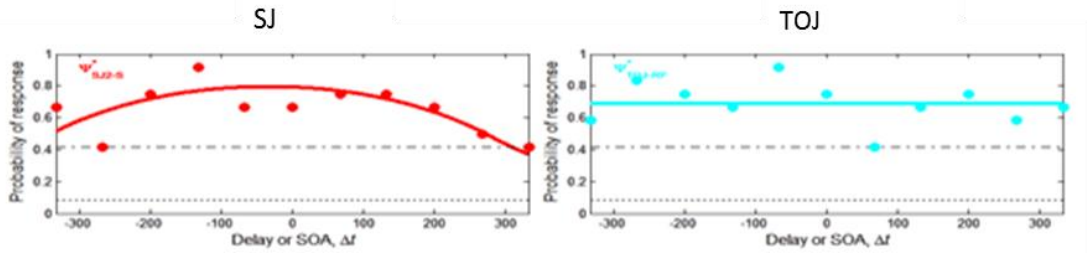


Figure 3.5 Example of unsuccessful Independent Channels model (ICM) fits of an individual's SJ and TOJ data when presented with FV stimuli.

		SJ			TOJ		
		BF	PLD	FV	BF	PLD	FV
ASD							
	N	25	25	23	20	17	22
	Excluded (%)	3.85	3.85	11.54	23.10	34.62	15.38
	Mean TIW_{ICM} (ms) [s.e.m]	147.91 [12.20]	108.86 [15.95]	164.44 [18.00]	125.88 [21.74]	100.60 [28.28]	139.76 [22.01]
	Mean PSS_{ICM} (ms) [s.e.m]	32.84 [11.81]	54.43 [4.91]	-20.01 [14.85]	16.40 [18.40]	48.50 [22.67]	-65.17 [27.99]
TD							
	N	25	25	26	25	17	23
	Excluded (%)	3.85	3.85	0	3.85	34.62	11.54
	Mean TIW_{ICM} (ms) [s.e.m]	115.31[9.33]	79.767 [15.95]	150.78 [12.87]	146.69 [14.01]	134.94 [23.89]	135.68 [21.97]
	Mean PSS_{ICM} (ms) [s.e.m]	34.76 [6.68]	47.12 [4.93]	-3.01 [8.58]	83.72 [20.37]	100.29 [47.37]	.633 [17.96]

Table 3.4 ASD= autism spectrum disorder, TD= typical developed, TIW_{ICM}= temporal integration window, PSS_{ICM}= point of subjective simultaneity, N= number of participants included in the analysis, s.e.m= standard error of mean

A 2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVA of TIW_{ICM} obtained from SJ responses revealed a marginally significant main effect of TIW_{ICM} width for group, F (1, 44)=3.82, p=.054, η^2p=.082 and significant main effect for stimulus, F (2, 88)= 16.56, p<.001, η^2p=.273 but no interaction, F(2,88)=.468, p=.628, , η^2p = .011. Bonferroni corrected post-hoc t-tests revealed that TIW_{ICM} width did not differ for BF (M=133.37) and FV (M=158.83) but that TIW_{ICM} for both stimuli was wider than for PLD (M=88.47), for both p<.001 (see Figure 3.6 and Table 3.4).

A 2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVA on PSS_{ICM} revealed no main effect of group, F (1, 44) = .049, p=.826, η^2p=.001 nor was there an interaction, F(2,88)=.684, p=.507, η^2p =.015. However a main effect of stimulus was found, F (1.58, 69.71) = 28.97, p< .001, η^2p=.397, where the PSS_{ICM} of BF (M= 37.49) and PLD (M= 50.15) were not significantly different from each other (p= .171) but both differed significantly from FV PSS_{ICM} (M= -12.97), p< .001. Two-tailed one-sample t-tests showed that PSS_{ICM} of BF and PLD were significantly different from zero (both p< 0.001), while the PSS_{ICM} of FV did not show a difference (p=.193).

A 2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVA on TIW_{ICM} computed from TOJ responses fitted using the model revealed neither a main effect of stimulus, F (2, 42) =.261, p=.771, η^2p=.012, group, F (1, 21) =.04, p=.842, η^2p=.002, nor an interaction, F(2,42)=.533, p=.591, η^2p =.025. However, PSS_{ICM} revealed main effects of group, F (1, 21) =4.83, p=.039, η^2p=1.87, and stimulus, F(1.59, 33.37)=5.207, p=.010, η^2p= .199, but no interaction F(2,42)=.419, p=.660, η^2p =.020. PSS in ASD (M=.784) was smaller than in TD (M=69.50). Contrasts of the different stimulus conditions showed that PSS_{ICM} of BF (M= 48.90) and PLD (M= 84.19) were not significantly different from each other (p= 1.00) but both differed significantly from FV PSS_{ICM} (M= -27.67), p= .017, p=.033, respectively. Again, two-tailed one-sample t-tests showed that PSS_{ICM} of BF and PLD were significantly different from zero (both p< 0.01), while the PSS_{ICM} of FV only showed a marginal difference (p=0.71).

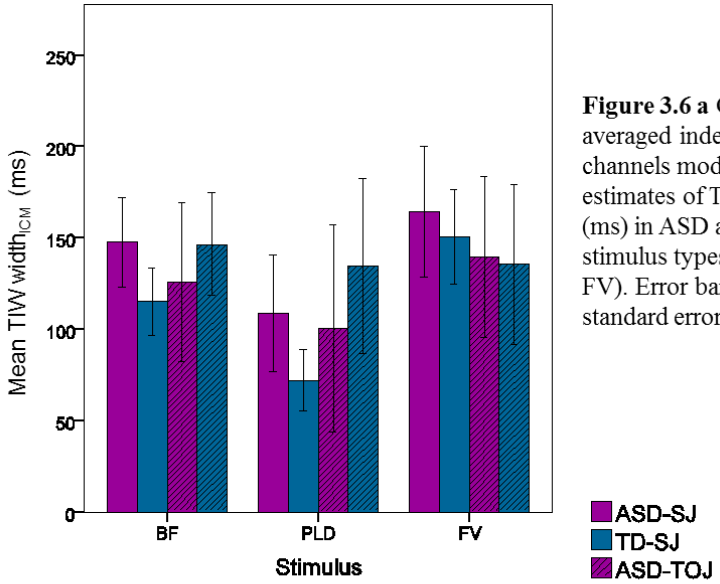


Figure 3.6 a Group-averaged independent channels model based estimates of TIW width (ms) in ASD and TD for 3 stimulus types (BF, PLD, FV). Error bars represent standard error.

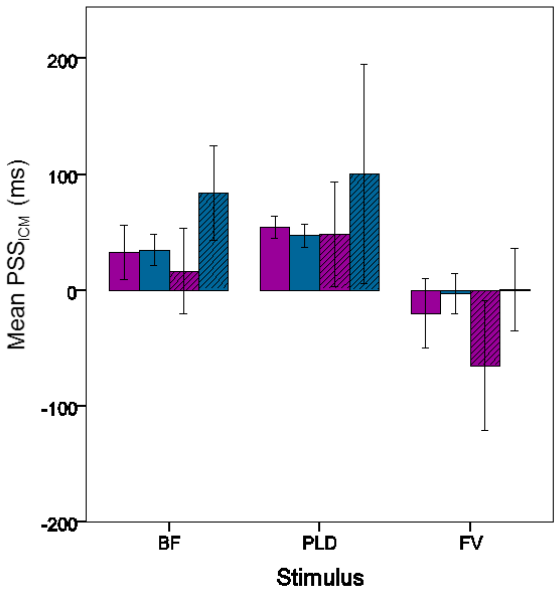


Figure 3.6 b Group-averaged independent channels model based estimates of PSS (ms) in ASD and TD for 3 stimulus types (BF, PLD, FV). Error bars represent standard error.

3.4.2.3 Estimated unisensory and decisional factors

Group	ICM parameters	SJ			TOJ		
		BF	PLD	FV	BF	PLD	FV
ASD							
	λ_a (ms) [s.e.m]	.095 [.136]	.099 [.125]	.074 [.102]	.115 [.137]	.215 [.143]	.127 [.146]
	λ_v (ms) [s.e.m]	.110 [.129]	.108 [.137]	.078 [.127]	.164 [.135]	.128 [.138]	.133 [.151]
	τ (ms) [s.e.m]	-42.59 [91.39]	-54.825 [39.43]	37.12 [88.17]	-30.72 [91.56]	-17.74 [99.08]	47.60 [100.77]
	δ (ms) [s.e.m]	199.6 [49.04]	199.82 [55.55]	210.45 [77.23]	160.18 [63.7]	236.87 [89.46]	161.09 [81.76]
	Xi (ms) [s.e.m]	-	-	-	.685 [.213]	.714 [.247]	.488 [.306]
TD							
	λ_a (ms) [s.e.m]	.096 [.122]	.146 [.142]	.107 [.124]	.083 [.121]	.163 [.155]	.165 [.140]
	λ_v (ms) [s.e.m]	.125 [.138]	.161 [.140]	.034 [.066]	.073 [.099]	.117 [.138]	.133 [.149]
	τ (ms) [s.e.m]	-36.05 [61.18]	-46.63 [40.09]	17.47 [73.95]	-88.47 [140.95]	-1447.09 [5837.38]	-116.99 [334.45]
	δ (ms) [s.e.m]	174.28 [53.97]	169.71 [39.00]	170.02 [46.35]	202.07 [101.30]	2268 [5671.88]	341.10 [305.26]
	Xi (ms) [s.e.m]	-	-	-	.685 [.312]	.694 [.178]	.613 [.197]

Table 3.5 The means (ms) and standard error of mean (s.e.m) of ICM parameters (λ_a, λ_v, τ, δ, Xi) for both groups (ASD, TD), both tasks (SJ, TOJ) and stimulus types (BF, PLD, FV). ICM= Independent Channels model, ASD= autism spectrum disorder, TD= typical developed, SJ= Synchrony Judgements, TOJ = Temporal order Judgements, λ_a = auditory lambda, λ_v = visual lambda, τ = tau, δ=delta.

2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVAs were used to look at the estimated parameters describing unisensory and decisional factors for SJs. Audio lambda (λ_a), which estimated the processing rate of the auditory cue, showed no main effect of group ($F(1,38)=1.17$, $p = .287$, $\eta^2p= .03$) stimulus ($F(2, 76) =.829$, $p = .44$, $\eta^2p= .021$), or interaction $F(2,76)=.401$, $p=.671$, $\eta^2p =.010$. Visual lambda (λ_v) estimated the visual cue processing rate and showed no main effect of group, $F(1,36)=.133$, $p = .718$, $\eta^2p= .004$ or interaction, $F(2,72)=1.42$, $p=.249$, $\eta^2p =.038$, but a significant main effect of stimulus ($F(2,72)=4.12$, $p< .05$, $\eta^2p= .103$, with FV (M= .056) being smaller than PLD (M= .135), p= .018, and both being comparable to BF (M= .118). Tau (τ) estimated the arrival time differences between auditory and visual cues, with $\tau < 0$ indicating faster auditory and $\tau > 0$ indicating faster visual processing. There was no main effect of group, $F(1,45)=.014$, $p = .908$, $\eta^2p= .001$, or interaction, $F(2,88)=.684$, $p=.507$, $\eta^2p =.015$, but there was a significant main effect of stimulus,

F(1.61,72.62)=21.37, p <.001, η^2p= .322, where τ of BF (M=-39.32) and PLD (M=-50.68) did not differ and indicate that the audio cue was processed faster than the visual cue. However τ of BF and PLD both significantly differed from FV (M= 27.30) where the visual cue was processed faster. Delta (δ) is the resolution parameter determining the ability to discriminate small differences in arrival latency and we found a significant main effect of group, F(1,45)=6.13, p =.017, η^2p= .12, but not a main effect of stimulus, F(2,90)=.229, p = .795, η^2p= .005 or interaction, F(2, 90)= .470, p= .626, η^2p= .010.

2 (ASD, TD) x 3 (BF, PLD, FV) mixed-effects ANOVAs for parameter estimates of TOJ showed no main effects of λ_a, λ_v, τ and δ for group (F(1, 21)=.165, p = 689, η^2p= .008; F(1, 21)=1.15, p = .296, η^2p= .052; F(1, 21)=.788, p = .385, η^2p= .036; F(1, 21)=1.602, p = .219, η^2p= .071, stimulus (F(2,42)=2.65, p=.083, η^2p=.112; F(2,42)=.037, p=.971, η^2p=.003; F(1.01,21.14)=.543, p=.470, η^2p=.025; F(1.01,21.11)=1.30, p=.267, η^2p=.058) or interactions (F(2, 42)= .723, p= .491, η^2p= .033; F(2, 42)= .722, p= .492, η^2p= .033; F(1.01, 21.14)= .505, p= .486, η^2p= .023; F(1.01, 21.11)= 1.18, p= .303, η^2p= .051). TOJ included an additional response bias parameter called Xi, and again no group differences were found, F(1, 21)=.254, p = .619, η^2p= .012, or interaction, F(2, 42)= .700, p= .502, η^2p= .032, but found a marginal main effect of stimulus F(2,42)=3.17, p=.052, η^2p=.131, with Xi for FV (M=.551) being significantly smaller than for PLD (M=.704), p=0.45, but neither differed from BF (M=.685).

3.4.3 Difficulty ratings of Judgements and Stimuli

Wilcoxon signed-rank tests on the difficulty ratings of each task (SJ, TOJ) and stimulus (BF, PLD, FV) combination showed that the participants found TOJs more difficult across all stimulus types (BF: Z=-3.137, p= .02; PLD: Z=-5.796, p< .001; FV: Z=-3.491, p<.001).

Freidman Tests showed that the difficulty of the SJ task differed depending on stimulus: $\chi^2(2) = 28.36$, $p < .001$. Mann-Whitney U tests revealed that the difficulty ratings of the ASD and TD group did not differ in any of the task/stimulus combinations (BFSJ: $U = 312.5$, $p = .614$; BFTOJ: $U = 313$, $p = .633$; PLDSJ: $U = 323$, $p = .752$; PLDTOJ: $U = 267$, $p = .172$; FVSJ: $U = 272.5$, $p = .205$; FVTOJ: $U = 291.5$, $p = .356$).

3.5 Discussion

To investigate the underlying processes of reduced sensitivity to audiovisual asynchrony observed in ASD (e.g., Stevenson et al., 2014, de Boer-Schellekens et al., 2013), the way the ASD and TD group performed on SJ and TOJ tasks was compared on a diverse range of stimulus types. We used Gaussian and cumulative Gaussian curves as well as an ICM (Garcia-Perez & Alcala-Quintana, 2012) to fit the response data to estimate TIW width and PSS in the two participant groups.

The Gaussian fits showed that for SJs, the ASD group's TIW width was wider compared to that of the TD group. However, the TOJs data revealed comparable TIW width of the two groups. In both SJs and TOJs, the TIW width differed across the types of stimuli shown. For SJs, FV had a wider TIW than PLD, which in turn had a wider TIW than BF. For TOJs, FV had a larger TIW than BF, but due to high exclusion rates PLD was not compared. PSS estimated from both SJ and TOJ showed no differences between the two groups, but revealed stimulus differences with SJ's PSS for BF and PLD being larger than that of FV, and TOJ's PSS for BF being larger than for FV.

In contrast to the more traditional Gaussian curve fits, the ICM used in this study gave the flexibility to fit complex asymmetric shapes and thus was able to fit asymmetric TIWs. This is of interest because it has been shown that individuals are better at detecting audio-leading asynchrony in SJ, but video-leading asynchrony in TOJ (Love at al., 2013). Fitting the ICM largely confirmed the results obtained from the Gaussian fits. By fitting the response data of SJ and TOJ together using the ICM, a wider TIW was found in the ASD group than in the TD group supporting the Gaussian fit results. The TIW width also differed across stimuli, with FV and BF having wider TIW than PLD, partly supporting the Gaussian fit results. PSS showed no difference between the groups, but showed differences between the stimuli with PSS for BF and PLD being larger than for FV, replicating the Gaussian fit results.

Fitting the ICM separately to SJ responses revealed a marginally wider TIW in the ASD than the TD group, and showed a wider TIW for FV and BF than for PLD, largely supporting the Gaussian fits. Computing TIW fitting the ICM to TOJ responses revealed no differences between the groups or stimuli (this could be explained by the high participant exclusion rate due to participants being unable to do TOJ on PLDs).

PSS from SJs showed no difference between the groups, but PSS differed across the stimuli, with the PSS of BF and PLD differing form from FV. PSS from TOJ, however, revealed a smaller PSS for ASD than for TD participants. This is a novel finding, which has not been shown by the Gaussian fitting method. PSS also varied across stimuli, with the PSS of BF and PLD being lager than that of FV.

Looking at the estimated parameters describing unisensory and decisional factors for SJs, the processing rate of the auditory cue, audio lambda (λ_a), was comparable across the two groups and all stimulus types. Similarly, visual lambda (λ_v) estimated the processing rate of the visual cue was the same between the groups, but differed across but stimuli. Tau (τ), the arrival time differences between auditory and visual cues, was the same across the two groups, but differed across stimulus types. For BF and PLD, tau indicated that the audio cue was processed faster than the visual cue, whereas, for FV, the visual cue was processed faster. Delta (δ), the resolution parameter determining the ability to discriminate small differences in arrival latency, was larger for the ASD group, indicating that they are less able to discriminate between small differences in arrival latency. This is consistent with a wider TIW in the ASD group. Delta did not change across the range of stimuli used. The parameter of TOJ (audio lambda, visual lambda, tau and delta) were estimated to be the same across the experimental groups and stimuli used. The response bias parameter of TOJ, Xi, found no group differences, but showed a marginal difference across stimulus.

Previous research looked at audiovisual integration in ASD using SJs or TOJs and fitted either linear functions to their response data (de Boer-Schellekens et al., 2013), two different psychometric sigmoid functions were fitted to allow for some asymmetry of the data (Stevenson et al., 2014) or simply compared the ASD group performance of each SOA to that of the controls, without modelling the data (Grossman, Schneps & Tager-Flusberg, 2009). Instead the current current study investigated both SJ and TOJ tasks by also fitting a flexible ICM to the participant's responses, not assuming symmetry of the data.

Fitting Gaussian curves and ICMs revealed a wider and marginally wider, respectively, TIW in ASD, compared to their TD controls when doing SJs. These findings are in line with Stevenson et at. (2014), who found a marginally wider TIW across their types of stimuli used, but their significant group x stimulus interaction

revealed that only the TIW of their complex FV stimulus was wider in ASD and not for simpler non-social stimuli. However, our findings, as well as those of Stevenson et al., (2014) are contrary to Grossman, Schneps and Tager-Flusberg's (2009) findings showing that children and young adults had equal TIW width when performing SJs on FV stimuli. A reason for finding different results to Grossman, Schneps and Tager-Flusberg (2009) could be the different nature of their stimuli, using meaningful phrases with quite big SOA intervals (ranging from 120ms to 500ms).

Our study found no group differences for TIW width in TOJs, a finding that is in opposition with previous results showing that ASD adolescents performing TOJs have a wider TIW than their TD controls (de Boer-Schellekens et al., 2013). Furthermore, the ICM revealed that the PSS in the ASD group is smaller than in the TD group, whereas de Boer-Schellekens et al. (2013) found no group differences for PSS. This differential finding could be explained by the fact that different fitting methods were used. de Boer-Schellekens et al. (2013) used linear fits on their data to estimate the PSS, whereas we used the ICM, which preserved the asymmetry and irregularities of the response data. Therefore, the ICM could be a fitting method allowed for a PSS estimate that is more sensitive at measuring this asynchronous position along the SOAs.

The discrepancies between the finding by Schellekens et al. (2013) and our current results could be explained by the fact that they tested adolescents and a few young adults, whereas we tested adults with ASD (18-40 years of age). Adults with ASD have previously been shown to develop compensatory strategies in tasks that children with ASD are deficient in (McKay et al., 2013). Therefore, we could argue that our adult sample has developed compensatory strategies to do TOJ, which lead to equal behavioural performance between ASD and TD participants. Furthermore, de Boer-Schellekens et al.'s (2013) sample, with 16 participants in each group, was relatively small. It needs to be noted that the equal TIW width between the groups found in our study could also partially be due to the high exclusion rates, in particular for PLD TOJ, as well as the large within group variability within the groups.

The two methods we used to fit the response data of SJs and TOJs showed that for SJs the participants with ASD had a wider TIW than the TD participants, whereas for TOJs TIW width was comparable across the two participant groups. The different

cognitive processes required for SJs and TOJs can help us understand the underlying processes of why temporal audiovisual integration differs in ASD. As proposed in the introduction, the ASD group having a wider TIW in SJs, but not TOJs, suggests that this difference is due to difficulties in combining the audio and the visual cues. SJs require to estimate the temporal correspondence of the audio and visual cue and thus depends on more global level processing (i.e., considering the stimulation as a whole), whereas TOJs could in principle be performed by focusing on only one sensory cue to detect whether it came first or not, thus depending on more local level processing (i.e., considering only the sound). Therefore, audiovisual integration difficulties in ASD are likely to be due to difficulties in processing global information in line with the hypotheses of central coherence deficit and temporal binding deficit in ASD.

The results support previous research showing that audiovisual temporal processing is not just effected in higher order social stimuli, but is also effected in simpler low level stimuli such as beeps and flashes (Foss-Feig et al., 2010; Kwakye et al., 2011; de Boer-Schellekens et al., 2013; Stevenson et al., 2014). In contrast, Stevenson et al. (2014) also found a group x stimulus interaction, only showing a significant group difference in their speech stimulus, but not in their simple or complex non-speech stimuli.

Interestingly, the ICM model explained the TIW width differences between the two groups in SJs by the resolution parameter Delta (δ), which measured the ability to discriminate small differences in arrival latency, and was larger for the ASD group, indicating that they are less able to discriminate between small differences in arrival latency. This is consistent with a wider TIW in the ASD group. The finding that audiovisual temporal processing differences cannot be explained by unisensory processing parameters supports previous findings by Stevenson et al. (2014), showing no group differences in either their audio or visual only TOJ tasks. However, these results are conflicting with studies showing unisensory integration differences in ASD (e.g., Kwakye et al., 2011; Williams at al. (2004).

The current study suggests that adults with ASD are not as good at detecting audiovisual asynchrony and that this difficulty is likely to be due to the less sensitive decisional process and not to unisensory temporal processing differences. These results are encouraging for potential interventions to improve sensory processing in

ASD, especially because it has been shown that the TIW width becomes smaller through training (Powers et al., 2009; Stevenson et al., 2013), and those with the widest TIWs are shown to improve the most after training. Our results would suggest that this training should be done in the multisensory domain, rather than in unisensory domains. Following up on Stevenson et al.'s (2014) results demonstrating the link between decreased sensitivity to audiovisual asynchrony, and the weaker percepts of the McGurk effect, it would be of interest to further explore the link between sensitivity to audiovisual asynchrony and speech perception and comprehension, as well as looking at how training on multisensory TIW width would translate into everyday multisensory speech processing and comprehension. Furthermore, future longitudinal studies could investigate the developmental trajectory of audiovisual temporal processing in ASD. Although behavioural evidence is of great importance, there is a need to understand the neural correlate of these multisensory integration deficits in ASD revealed in this study. Thus far there has been little research investigating the neural underpinnings of the differences in audiovisual integration in ASD. Chapter 4 will investigate audiovisual, audio and visual processing in ASD using fMRI. Moreover, to our knowledge, no fMRI research is published looking at audiovisual temporal processing in ASD during SJ. This will be done in Chapter 5.

Conclusion

This study investigated audiovisual integration in ASD using SJ and TOJ as well as different data fitting methods. More specifically, the analysis of the data included fitting Gaussian functions as well as using an ICM to fit the data (Garcia-Perez & Alcala-Quintana, 2012). Gaussian curve fitting for SJs showed that the ASD group had a wider TIW, but for TOJ no group effect was found. The ICM supported these results and model parameters indicated that the wider TIW for SJs in the ASD group was not due to sensory processing at the unisensory level, but rather due to decreased temporal resolution at a decisional level of combining sensory information. Furthermore, when performing TOJ, the ICM revealed a smaller PSS (closer to physical synchrony) in the ASD group than in the TD group. These behavioural results raise the importance of investigating the neural underpinnings of the differences in audiovisual integration in ASD.

4 An fMRI investigation of the audiovisual, audio and visual processing in Autism Spectrum Disorder

4.1 Abstract

The ability to integrate auditory and visual information is crucial to everyday life. Behavioural results have predominantly shown that individuals with Autism Spectrum Disorder (ASD) have deficits in audiovisual integration. These findings have recently been supported by electroencephalography (EEG) studies (Brandwein et al., 2015). Using functional magnetic resonance imaging (fMRI), we investigated audiovisual, auditory and visual processing in ASD of simple, beep-flash (BF) displays and complex, social face-voice (FV) displays. During a block design experiment, we measured the BOLD signal when 13 adults with ASD and 13 typically developed (TD) age-, sex- and IQ- matched adults were presented with audiovisual, audio and visual information of BF and FV displays. Our analyses revealed that processing of audiovisual as well as unisensory auditory and visual stimulus conditions, in both the BF and FV displays, was associated with reduced activation in ASD. Audiovisual, auditory and visual conditions of FV stimuli revealed reduced activation in ASD in regions of the frontal cortex, while BF stimuli revealed reduced activation in the lingual gyri. In the inferior parietal gyrus we found different sensory conditions of BF to modulate the activation levels differently in ASD than in TD. Conjunction analyses revealed smaller regions of the superior temporal cortex (STC) in ASD to be audiovisual sensitive. Against our predictions, the STC did not reveal any activation differences, per se, between the two groups. However, a superior frontal area was shown to be sensitive to audiovisual face-voice stimuli in the TD group, but not in the ASD group. Overall, this study indicated differences in brain activity for audiovisual, auditory and visual processing of social and non-social stimuli in individuals with ASD compared to TD individuals. These results contrast previous behavioural findings (Chapter 3), suggesting deficient audiovisual integration, yet intact auditory and visual processing in ASD.

4.2 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition described by deficits in social communication and interaction, as well as repetitive patterns of behaviour, interests and activities (American Psychiatric Association, 2013). Differences in sensory processing of ASD, compared to typically developed (TD) individuals, has been reported across the different sensory modalities, such as audition, vision, taste, smell, vestibular and proprioception (Lane et al., 2015; Ludlow et al., 2014; Conner, 2012; Simmons et al., 2009). Moreover, a strong correlation between autistic traits and sensory sensitivities in the general population has also been found (Robertson & Simmons, 2013). The relevance of these sensory differences has recently been recognised and included as diagnostic criteria in the DSM-V. In fact, researchers have started to stress the importance of understanding to what degree sensory sensitivities in individuals with ASD contribute to their social and communication impairments. For example, multisensory processing differences have been proposed to cascade down to cause communication impairments (Brandwein et al., 2015). Similarly, acts of apparent non-compliance, reluctance, lack of interest, as well as aggression, might not be voluntary, and could be secondary to an individual's particular sensory processing and movement profile (Donnellan, Hill & Leary, 2013). This highlights the importance of understanding the neural correlates of these sensory processing differences better, and potentially developing sensory processing interventions.

Interest in multisensory integration in autism has gained increasing recognition as we further appreciate the importance of integrating information from different senses in everyday life. It has been shown that the combination of multiple sensory cues can reduce uncertainty and enhance the ability to make better estimates of the situation (Ernst & Banks, 2002). For example, in a crowded and noisy environment looking at a person's face and lip movements enables us to better understand what the person is telling us.

Audiovisual processing in TD individuals

Using functional magnetic resonance imaging (fMRI), brain regions involved in typical audiovisual integration of simple synchronous displays have been well researched (e.g., Calvert, Campbell & Brammer, 2000; Werner & Noppeney, 2010;

Love, Latinus & Pollick, 2011; Watson et al., 2014). Several neuroimaging studies in adults have begun to identify important brain regions in a network underlying audiovisual simultaneity perception. These include: thalamus (Love, Latinus & Pollick, 2011), hippocampus (Watson et al., 2014), the insula (Calvert, Hansen, Iversen, & Brammer, 2001), inferior parietal lobule (Calvert et al., 2001; Dhamala et al., 2007), superior colliculus (Calvert et al., 2001; Dhamala et al., 2007), posterior superior temporal sulcus (Calvert, Campbell, & Brammer, 2000; Calvert et al., 2001; Dhamala et al., 2007; Beauchamp, Yasar, Frye, & Ro, 2008; Werner & Noppeney, 2010, Steveson & James, 2009; Watson et al., 2014), and unisensory cortices (Noesselt et al., 2007). Moreover, evidence suggests that the Superior Temporal Cortex (STC) has specialist areas for face-voice speech integration (Stevenson et al., 2011; Watson et al., 2014), as well as distinct regions utilised for processing temporal-synchrony (Stevenson et al., 2011). These findings point towards the idea that the STC is a neuronal centre, made up of different regions that underlie a range of low- and high-level multisensory integration processes.

Audiovisual processing in ASD

Research showed that processing multiple sensory signals as a single percept is not as beneficial in ASD as in TD individuals. For instance, people with ASD perceive audiovisual illusions such as the McGurk effect (McGurk & MacDoland, 1976) less often than their TD controls (de Gelder et al., 1991; Irwin et al., 2011; Mongolli et al., 2008), often benefiting less from an additional sensory modality (Smith & Bennetto, 2007) and relying more on one sensory modality (Stevenson et al., 2014).

Electroencephalography (EEG) studies recording high-density brain activity have shown that the neural integration of audiovisual information is atypical in children with ASD (Magnee et al., 2009; Russo et al., 2010; Brandwein et al., 2013, 2015). It has also recently been shown that children with ASD are not as effective at paying attention to a relevant unisensory stream when presented with competing multisensory information (Murphy et al., 2014). Moreover, using fMRI it has been shown that adolescents with ASD use different cortical areas when processing audiovisual emotion stimuli compared to TD adolescents (Doyle-Thomas et al., 2013). More specifically, in this study, brain activation in participants was measured when making emotional judgements of audiovisual displays. Activation patterns revealed that the

ASD group employed parietal and frontal cortices, whereas the TD group recruited frontal and temporal cortices during this task. It was suggested that the absence of integrative emotional networks in ASD might cause the recruitment of the parietofrontal network as a compensatory result. Similarly higher activation patterns were shown in a pilot study by Loveland et al. (2008), who showed that, during emotional congruency tasks, TD participants had more activation compared to the ASD participants in the STC, orbitofrontal cortex, posterior cingulate, parahippocampus and occipital regions (left fusiform gyrus, and bilateral lingual gyrus extending into the left cuneus). However, since both studies (Doyle-Thomas et al., 2013; Loveland et al., 2008) employed emotional stimuli and asked the participants to make emotion judgements, it is likely that these studies also reflect the underlying neural correlates of emotion processing, and not audiovisual processing itself.

Unisensory auditory and visual processing in ASD

Differences in unisensory auditory (Conner, 2012) and visual (Simmons et al., 2009) processing have been frequently reported. Face processing in ASD has been reported to exhibit hypo activations in the face processing network, including regions such as the fusiform face area, occipital face area, pSTS, as well as frontal regions (Pierce, Haist, Sedaghat, & Courchesne, 2004; Pierce & Redcay, 2008; Scherf et al., 2010; Scherf et al., 2015). During object processing tasks, the ASD group revealed hyper-activation in the precuneus (Sherf et al., 2015), while others have found no group differences in processing objects (Humphreys, Hasson, Avidan, Minshew, & Behrmann, 2008). A similar pattern of findings was revealed when looking at auditory processing of voices and non-vocal sounds. Individuals with ASD failed to activate STS voice-selective regions in response to vocal sounds, while they showed typical activation patterns in response to non-vocal sounds (Gervais et al, 2004). These findings suggest abnormal cortical processing of socially relevant auditory and visual information in autism.

The behavioural results in Chapter 3, on audiovisual integration and synchrony perception, made use of the Independent Channels Model (ICM) by Garcia-Perez and Alcala-Quintana (2012) to estimate the unisensory, as well as decisional aspects of synchrony perception tasks. The results point towards no unisensory processing differences between the ASD and the TD group. Instead, the study revealed

underlying decisional deficits in ASD. However, other studies investigating audiovisual processing in ASD have revealed that these results can be attributed to processing deficits at a unisensory level (Williams et al., 2004).

As shown by typical audiovisual processing, the STC appears to be an important neural centre, which includes different regions that underlie a range of low- and high-level multisensory integration processes. Within the STC, the posterior STS has been found to have structural grey matter (GM) differences in ASD (Scheel et al., 2011; Greimel et al., 2013; Doyle-Thomas et al., 2013; Hyde et al., 2010), as well as white matter abnormalities (Barnea-Goraly et al ., 2004) and atypical functional activations (Buckner et al., 2008). These atypicalities of the STS in ASD have been proposed to be the underlying cause of the common aetiology for audiovisual temporal processing deficits observed in ASD and other developmental conditions**.**

This chapter describes a block design experiment which measured the BOLD signal when participants were presented with audiovisual information, audio information and visual information of beep-flash (BF) and face-voice (FV) displays. Whole-brain analyses were run to explore how activation levels were influenced by the different sensory modalities across the experimental groups (ASD and TD). Further conjunction analyses of (AV> A) ∩ (AV> V) were preformed to establish regions sensitive to audiovisual processing separately in the ASD and the TD groups. A conjunction analysis or max-criterion was chosen, as it has previously been shown to be an appropriate criterion to establish areas sensitive to audiovisual information (Kreifelts et al., 2010; Szycik et al., 2008; Love, Pollick, & Latinus, 2011; Watson et al., 2014).

4.3 Methods

4.3. 1 Participants

Thirteen high-functioning adults with Autism Spectrum Disorders (aged between 21 and 41) and 13 age-, sex- and IQ-matched control participants (aged between 21 and 41) took part in this study (Table 4.1). All participants in the ASD group reported to have a diagnosis of having an ASD according to DSM-IV criteria from a qualified clinician. All were native English speakers, had normal or corrected to normal vision and reported no hearing difficulties. The Autism Quotient (AQ), a 50 item autism traits questionnaire by Baron-Cohen, Wheelwright, Skinner, Martin and Clubley

(2001), with the cut off score for Asperger's being 26, supported the diagnoses of the ASD group (M= 37.54, SD = 6.89), and reinforced the assumption that no-one in the TD group had ASD (M=12.31, SD= 4.09). The participants were matched pair-wise on age (t (12)=.82, p=.42) and group-wise on full scale IQ (FSIQ) (t (12)=.51, p=.62), as measured using the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). The experimental procedures were approved by the School of Psychology at the University of Glasgow and also the Greater Glasgow and Clyde National Health Service ethics board.

Group	Age		FSIQ	
	Mean	SD	Mean	SD
ASD	30.54	7.42	119.92	10.13
TD	29.46	5.34	118.23	6.46

Table 4.1 Mean and Standard Deviation (SD) of the ages and Full Scale IQs (FSIQ) of the autism spectrum disorder (ASD) and typically developed (TD) group separately.

4. 3. 2 Stimuli

During the experiment, participants were presented with blocks of synchronous audiovisual (AV) displays, as well as audio only (A) and visual only (V) displays. Two different display types were used: non-social BF displays and complex social FV displays (see Figure 4.1).

Previously recorded synchronous videos used in Chapter 3 and by Love et al. (2013) were used to create the videos for both fMRI experiments. The videos were in QuickTime file format (.mov) and were uploaded into Adobe Premiere Pro CC 2014 (8.0.1) to manipulate their levels of asynchrony. All videos were created to be the same duration, irrespective of their sensory modality (AV, A,V). In order to present the videos using Presentation 14.9 designed by NeuroBehavioral Systems (NBS), the newly created videos were exported in an uncompressed AVI format and then compressed in VirtualDub 1.10.4 to minimise quality loss.

All BF videos were 816 ms long with a frame rate of 60 frames per second. The flash was a white dot (luminance: 85 cd/m^2; visual angle of the diameter: 4.4degrees) on black background, while the beep was a pure tone at 2000 Hz and 84 dB mean

intensity. In the synchronous AV condition, the flash and the beep started at 400 ms. In the V condition, only the flash was presented at 400ms, and in the A condition only the beep was presented at 400ms. During the A presentation, a black screen was presented.

The FV videos were 1920 milliseconds long, with a frame rate of 25 frames per second. The video was of a man saying the word "Tomorrow". The visual angle of the visual speech cue was approximately 12.7 and 18.2 degrees and the mouth region covered about 3.2 by 2.5 degrees of visual angle. This made the mouth region approximately the same size as the flash in the BF videos. Before and after the word was spoken, a still image of the first and the last frame of the video was shown and faded in at 120 ms and faded out at 1840 ms.

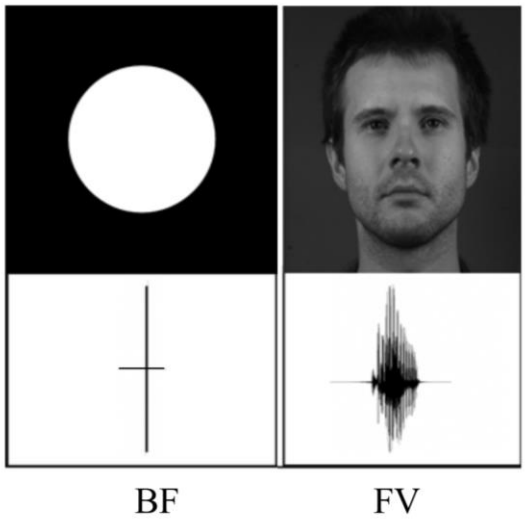


Figure 4. 1 The top panel shows the visual information participants were presented with. The bottom panel shows the auditory waveform for each type of stimulus. The beep-flash (BF) stimulus consisted of a flash of a white dot on a black background and a beep. In the face-voice (FV) stimulus a movie frame is shown and the waveform represents the word "tomorrow". Please note that the images are not to scale, the area of the white flash dot is approximately the same size as the area of the mouth in FV.

4. 3. 3. Design

This experiment aimed to investigate activation differences between the ASD and TD group when perceiving displays of different modalities: AV, A and V. In both the BF and the FV runs, the participants were presented with blocks of three sensory

modalities (AV, A, V), during which the BOLD signal was measured in the fMRI scanner. In the BF run, participants were presented with blocks of BF, beep only and flash only displays. In the FV run, participants were presented with blocks of FV, face only and voice only displays. There was one run of BF and one of FV displays, both containing 21 blocks (seven of each of the 3 sensory modalities). The order of the blocks was pseudo-randomised: each block was always preceded and followed by a block from a different stimulus condition (e.g., a block of A could never be preceded or followed by any other block of A). Two different pseudo-random sequences were created and each of the two sequences were shown to half of the participants in both groups (TD, ASD). The order of the BF and FV runs was counterbalanced to remove any possible order effects. The BF runs were 376 seconds long. Each individual block lasted about 11 seconds (containing 14 repetitions of BF displays), while the FV runs were 368 seconds long, with each block lasting about 11 seconds (containing 7 repetitions of FV displays). In all runs, each block was followed by 4 seconds during which a black screen was shown. Participants were also presented with a black screen for 20 seconds at the start of the each run and 16 seconds at the end of each run.

4. 3. 4. Procedure

Each participant was instructed to pay close attention to the stimuli presented. Participants were shown either the BF run or the FV run first in a counterbalanced order to remove any possible order effects. During both experiments, participants were presented with blocks of audiovisual information, audio information and visual information. Participants were not asked to perform an active task. Together, the 2 experiments took about 14 minutes to complete. All of the MRI and fMRI data collection was performed at the Centre for Cognitive Neuroimaging (CCNi) at the University of Glasgow, UK. Participants were walked through the scanner safety checklist to ensure that they were safe to be scanned. Before entering the scanner, participants were told what stimuli they were going to be presented with and we checked whether they understood the task instructions. All participants provided informed written consent. In the scanner, participants were shown how to use the emergency buzzer in case they felt uncomfortable and wanted to stop the experiment. Participants were made comfortable in the scanner and were given the emergency button. If needed, the participants vision was corrected using the Nordic Neurolabs Visualsystem goggles until participants were able to clearly see the stimuli and

instructions. Once participants were comfortable, they were moved into the scanner and all subsequent communication took place from the control room via an intercom system. The instructions were repeated for each run and participants' comfort was checked. Stimuli were presented using Presentation 14.9 designed by NeuroBehavioral Systems (NBS), via electrostatic earphones (NordicNeuroLab, Norway) at a sound pressure level of 80 dB. In between scans, we checked that participants found the sound pressure level comfortable and loud enough considering the scanner noise. After the study, everyone was reimbursed for their time and transportation

4. 3. 5 Data acquisition parameters

A Siemens 3T Tim Trio MRI scanner was used to acquire sagittal T1 weighted anatomical images and T2 weighted functional images.

4. 3. 5. 1 Functional data

Functional T2 weighted images were acquired covering the whole brain (slices = 32, dimension = 210 x 210 mm, voxel size resolution = 3 x 3 x 3 mm) for each of the 188 and 184 volumes of the BF and FV sub-experiments., using a 32-channel head coil and an echoplanar imaging (EPI) sequence (interleaved, TR = 2 seconds, TE = 30 ms, Flip Angle = 90°) with online motion correction. The first 2 volumes of each functional run comprised 'dummy' gradient and radio frequency pulses, which permitted for steady state magnetisation. During these volumes no stimuli were presented and no fMRI data was collected. Preprocessing and analysis used the motion corrected (moco) series output by the Siemens system.

4. 3. 5. 2 Structural data

At the end of each fMRI session a high-resolution T1-weighted structural image was collected in 192 axial slices and isotropic voxels (resolution: 1 mm x 1mm x 1 mm; dimensions: 256 x 256 mm, TR = 1900 msec, TE = 2.92 msec, time to inversion = 900 msec, FA = 9°). The run time was 10 minutes.

4. 3. 6 fMRI Preprocessing

BrainVoyager QX version 2.8 was used to preprocess and analyse all stages of the fMRI data. The first two functional volumes were excluded to allow for signal

stabilisation. Structural scans were homogeneity corrected and transformed into Talairach space (Talairach & Tournoux 1988) using BrainVoyager QX 2.8 (BrainInnovation, Maastricht, the Netherlands). Functional runs were slice scan time corrected, 3D motion corrected (using trilinear/sinc interpolation) and temporally high-pass filtered at 3 cycles across each run. The functional runs were coregistered to the 1 × 1 × 1 triliniar-interpolated anatomical maps scans and transformed into talairach space. A Gaussian 6mm spatial filter was applied to the 4D volumes in order to improve the signal to noise ratio for group analysis by overcoming differences in intersubjective localisation.

4. 3. 7. 1 Whole-Brain general linear model (GLM)

For both the BF and FV experiments, a second-level, multi-subject, random effects GLM was computed. For both experiments a 2 (group: ASD, TD) x 3 (sensory modality: AV, A, V) mixed-measures ANOVA was run with group as the between-subject factor and sensory modality as the within-subject factor. To account for multiple comparisons, the volume maps were set at a voxel-level uncorrected threshold of $p < 0.001$ and the cluster size threshold estimation (Worsley, Evans, Marrett & Neelin, 1992) was used to control for minimum cluster sizes for each contrast, based on a criterion of $p < 0.05$. A Monte Carlo simulation of 1000 iterations estimated cluster-level false-positive rates. The regions were defined based on those that survived cluster-size threshold estimation, and, where regions covered excessively large areas, effort was made to ensure regions conformed to anatomical boundaries. This method is commonly used to control for multiple comparisons.

4. 3. 7. 2 Conjunction analyses

A series of random-effects conjunction analyses were performed in order to identify regions in which audiovisual integration took place when presented with a simple flash and beep stimulus (BF) and a more complex face-voice stimulus (FV). Each audiovisual condition was contrasted against each of the corresponding unisensory (audio, visual) conditions. Thus, for the BF conjunction analyses run in the ASD and the TD group, separately: beep-flash was contrasted to beep only and beep-flash was contrasted to flash only (i. e., $(BF > B) \cap (BF > F)$; the 'max rule', Beauchamp, 2005; Love, Pollick, & Latinus, 2011). Multisensory audiovisual voxels had to be significantly active in both contrasts. Similarly for the FV conjunction analyses of the

ASD and the TD group: face-voice was contrasted to voice only and face only (FV > F) ∩ (FV > V). These analyses localised regions showing a higher BOLD response to audiovisual stimuli, as compared to both visual only and audio only stimuli. As before, the significance levels were set to $p < 0.001$ and the cluster threshold estimation account for multiple comparisons.

4. 4 Results

4. 4. 1 fMRI activation data

4. 4. 1. 1 Whole-brain GLM of beep-flash displays

For the BF displays, a 2 (group: ASD, TD) x 3 (sensory modality: AV, A, V) REX GLM revealed a significant main effect of group bilaterally in the lingual gyrus, a main effect of sensory modality in bilateral superior temporal gyrus, middle occipital lobule, occipital lobe, and precuneus. In addition, a significant interaction between group and sensory modality was found in the right inferior parietal lobule.

The main effect of group found in the lingual gyri revealed a reduction of activation of the ASD group, compared to the TD group across all sensory modalities (Figure 4. 2, Table 4.2). The interaction between group and sensory modality in right inferior parietal lobule revealed that here the ASD group showed more activation for AV and A than for V, whereas the TD group showed more activation for AV and V than for A (Figure 4.3, Table 4.2). In the result and discussion section, we focus on activation differences between ASD and TD individuals, as well as interaction effects of group performance and stimulus type. Stimulus differences will not be discussed.

Table 4. 2 BF experiment: clusters of activation from a 2 × 3 ANOVA with 'experimental group' as a between-participants factor and 'sensory modality' as a within-participants factor. Legend: BA — Brodmann's area
*activation reached across both hemispheres

Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	F-value	P-value	BA
Group (ASD, TD)						
Lingual gyrus,	Left	-15, -67, 1	128	22.68	0.00008	18
Lingual gyrus	Right	18, -64, -8	90	24.02	0.00005	19
Group (ASD, TD) × Sensory modality (AV, A, V)						
Inferior parietal lobule	Right	63, -22, 25	109	9.96	0.00024	40
Sensory modality (AV, A, V)						
Superior temporal gyrus	Right	39, -22, 7	23123	56.57	0.00001	13
Inferior temporal gyrus	Right	42, -64, 1	9947	46.79	0.00001	37
Occipital lobe	Left*	-12, -67, 7	50568	50.24	0.00001	30
Precuneus	Right*	3, -49, 43	5313	13.48	0.00002	7
Inferior temporal gyrus	Left	-42, -67, 1	7162	38.69	0.00001	37
Superior temporal gyrus	Left	-39, -25, 7	25750	75.87	0.00001	13

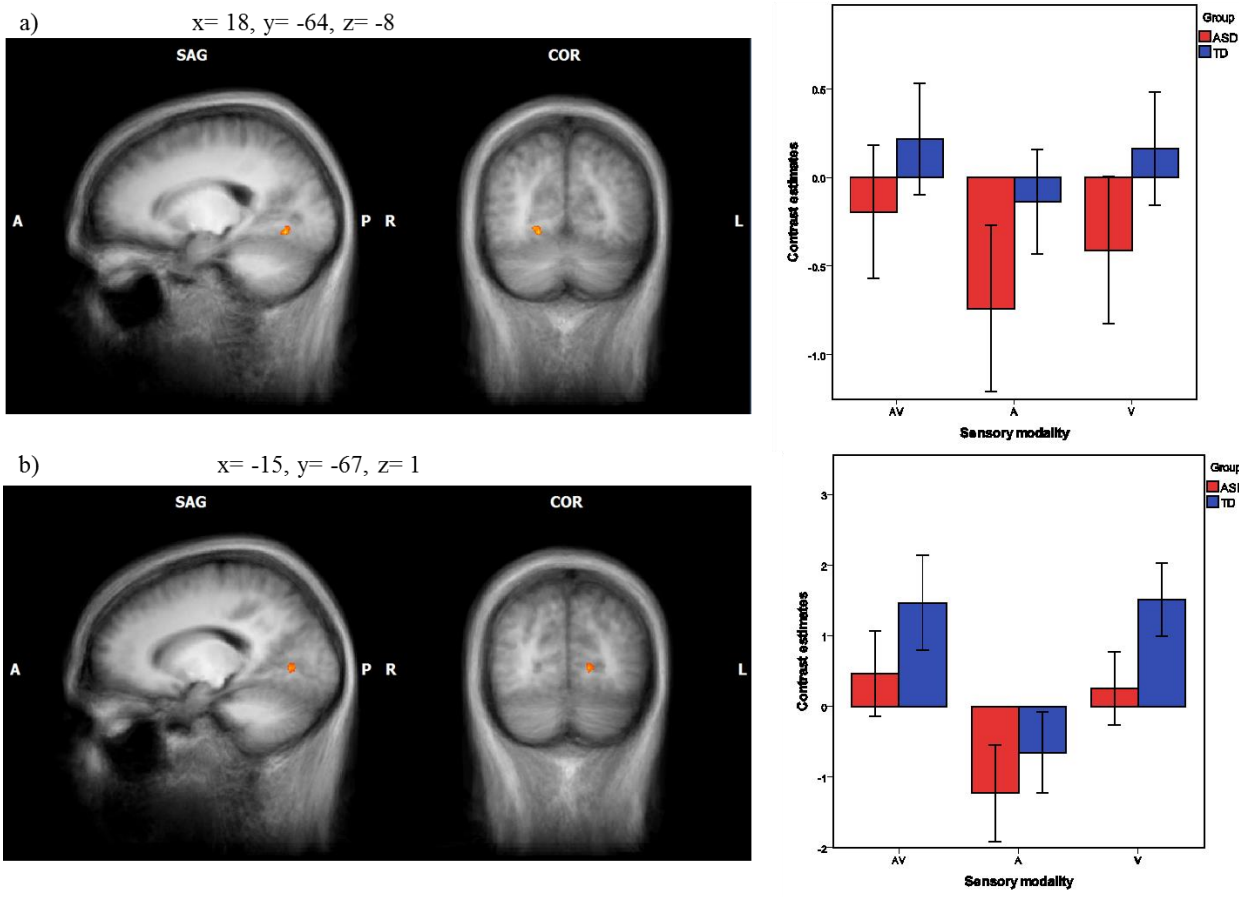


Figure 4. 2 fMRI activation data: Clusters of activation for which the difference between the brain responses to the three beep-flash stimulus conditions (sensory modality: audiovisual, audio and visual) and the two groups of participants (ASD and TD individuals). Coronal and sagittal slices show activation foci at two locations (defined by x y z Talairach coordinates). The average contrast estimates (beta weights) and relative standard errors are shown in histograms for ASD (red) and TD (blue) at each stimulus condition: audiovisual (PSS), audio (A) and visual (V). a) right lingual gyrus b) left lingual gyrus are clusters of activation for which the brain responses differed between two groups of participants when presented with beep-flash stimuli of all three sensory modalities.

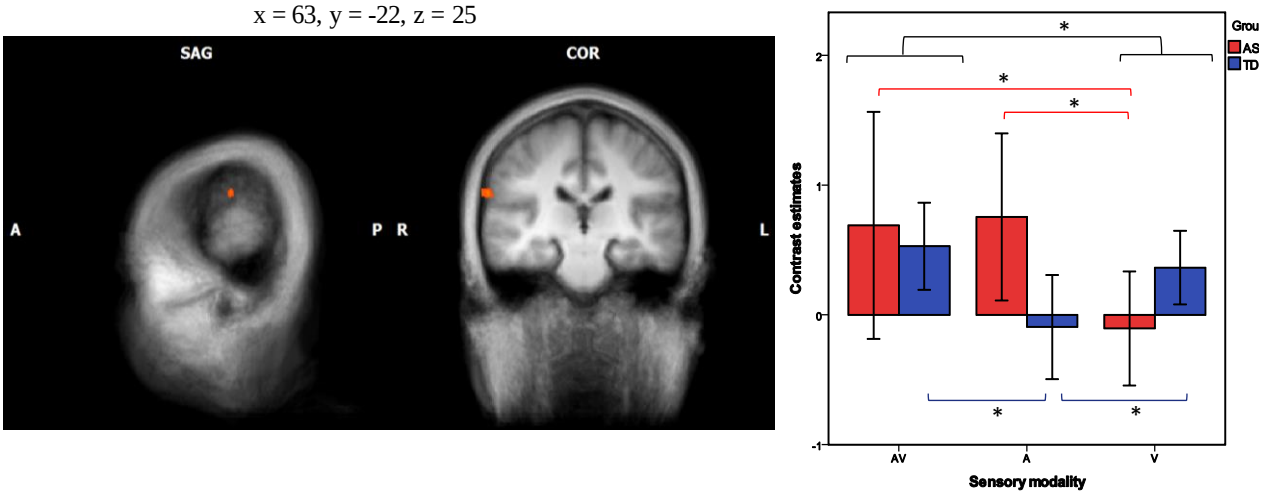


Figure 4. 3 fMRI activation data: in the right inferior parietal lobule we found a cluster of activation for which the brain responses revealed an interaction between the three beep-flash stimulus conditions (sensory modality: audiovisual, audio, visual) and the two groups of participants (ASD and TD individuals). Coronal and sagittal slices show the activation focus at one location (defined by x y z Talairach coordinates). The average contrasts estimates (beta weights) and are shown in the histogram for the ASD (red) and TD (blue) at each stimulus condition: audiovisual, audio, visual. The brackets and * indicate where the pairwise comparisons found significant differences ($p<0.05$) between the conditions. The black brackets indicate significant differences between the conditions across the groups, the red brackets show significant differences between the stimulus conditions in the ASD group and blue brackets indicate significant differences between the stimulus conditions in the TD group.

4. 4. 1. 2 Whole-brain GLM of face-voice displays

For the FV displays, a 2 (group: ASD, TD) x 3 (sensory modality: AV, A, V) REX ANOVA revealed a significant main effect of group in the right inferior frontal gyrus and superior frontal gyrus, a main effect of sensory modality in bilateral superior parietal lobule, occipital lobe, right superior temporal gyrus, caudate, thalamus and culmen, but no significant interaction between group and sensory modality.

The main effect of group in both the inferior and superior frontal gyrus revealed a reduction of activation in the ASD group, compared to the TD group across all sensory modalities (Figure 4. 4, Table 4. 3).

Table 4. 3 Experiment 1 FV clusters of activation from a 2 × 3 ANOVA with 'experimental group' as a between-participants factor and 'sensory modality' as a within-participants factor. Legend: BA — Brodmann's area
*activation reached across both hemispheres

Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	F-value	P-value	BA
Experimental group (ASD, TD)						
Middle frontal gyrus	Right	42, 38, 13	868	27.48	0.00002	46
Superior frontal gyrus	Right	21, 5, 61	443	31.43	0.00001	6
Sensory modality (AV, A, V)						
Superior temporal gyrus	Right	63, -16, 7	23972	95.27	0.00001	22
Cuneus/ Occipital lobe	Right*	12, -101, 4	137300	461.88	0.00001	18
Caudate	Right	21, -40, 16	4974	34.81	0.00001	-
Thalamus	Right	21, -25, 1	3359	115.3	0.00001	-
Superior parietal lobule	Right	21, -46, 61	5491	21.83	0.00001	7
Culmen	Right	6, -37, -23	2013	19.73	0.00001	-
Superior parietal lobule	Left	-18, -49, 58	5608	24.80	0.00001	7

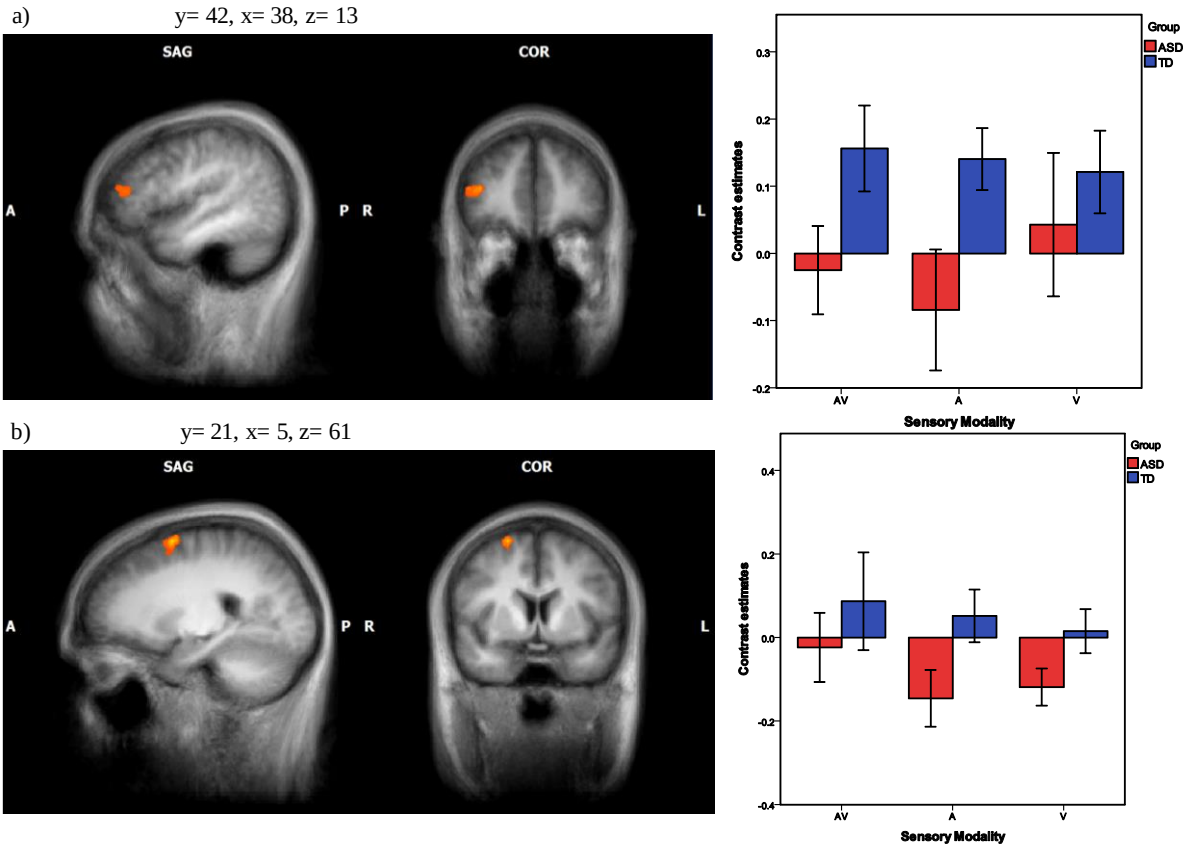


Figure 4. 4 Experiment 1a fMRI activation data: Clusters of activation for which the difference between the brain responses to the three face-voice stimulus conditions (sensory modality: audiovisual, audio and visual) and the two groups of participants (ASD and TD individuals). Coronal and sagittal slices show activation foci at two locations (defined by x y z Talairach coordinates). The average contrast estimates (beta weights) and relative standard errors are shown in histograms for ASD (red) and TD (blue) at each stimulus condition: audiovisual (PSS), audio (A) and visual (V). a) right middle frontal gyrus and b) superior frontal gyrus are clusters of activation for which the brain responses differed between two groups of participants when presented with face-voice stimuli of all three sensory modalities.

4. 4. 2 Conjunction analyses

The conjunction analyses of BF and FV (AV> A) ∩ (AV> V) were performed for both experimental groups (ASD and TD) separately, as well as used to compare them across the experimental groups.

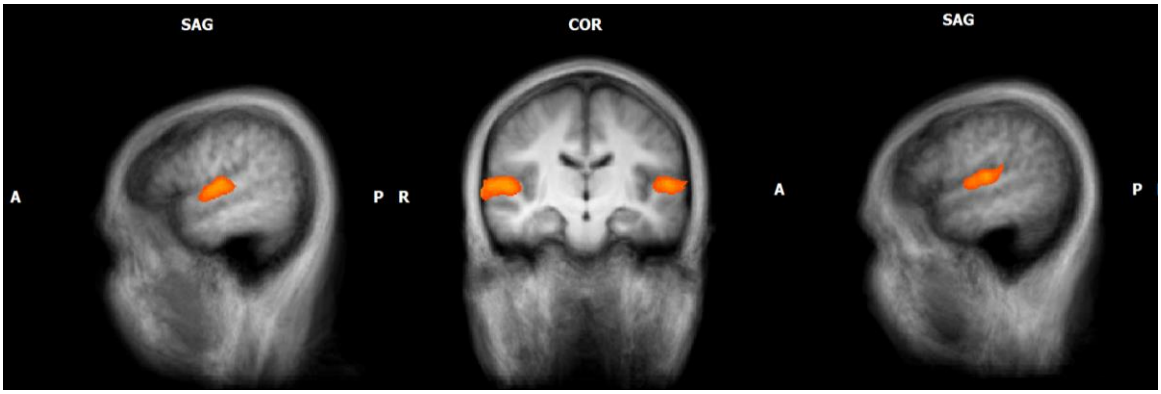
4. 4. 2. 1 Conjunction analyses of beep-flash displays

The RFX conjunction analysis for BF stimuli revealed two clusters of voxels in the right and left superior temporal gyrus in both the ASD and the TD group (Figure 4. 5, Table 4. 4). Comparing the two contrasts between the two groups revealed no activation differences revealing different audiovisual BF sensitivity compared to audio beep and visual flash only conditions.

Table 4. 4 Conjunction analysis of BF stimuli. For the ASD group and the TD group clusters of activation for (AV> A) ∩ (AV> V) are shown. Legend: BA — Brodmann's area

Experimental group	Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	t-value	P-value	BA
ASD							
	Superior temporal gyrus	Right	60, -7, 1	6278	7.06	0.00001	22
	Superior temporal gyrus	Left	-54, -19, 10	4848	6.37	0.00003	41
TD							
	Superior temporal gyrus	Right	63, -19, 10	11493	9.62	0.00001	42
	Superior temporal gyrus	left	-48, -19, 7	11878	11.11	0.00001	22

a) x= 60, y= -7, z = -1 x = -54, y = -10, x = 10



b) x= 63, y =-19, z = 10 x = -48, y = -19, z = 7

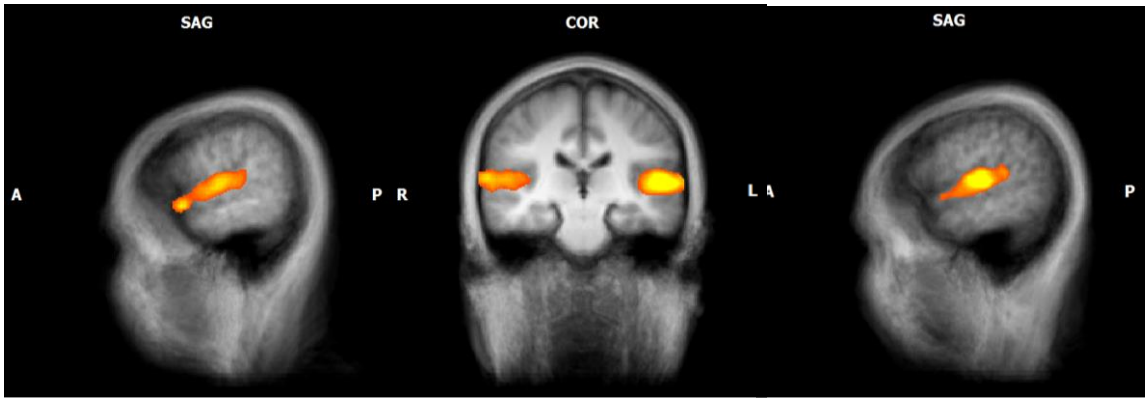


Figure 4. 5 fMRI activation data: Conjunction analysis of (audiovisual > audio) ∩ (audiovisual > visual) to define integrative audiovisual regions separately for a) ASD group and b) TD group for beep-flash stimuli. Coronal and sagittal slices show activation foci at two locations (defined by x y z Talairach coordinates). The clusters of activation sensitive to audiovisual integration are in the bilateral superior temporal sulcus. for both groups of participants (ASD, TD).

4. 4. 2. 2 Conjunction analyses of face-voice displays

Similarly, the RFX conjunction analysis for FV stimuli revealed two clusters of voxels in the right and left superior temporal gyrus in both the ASD and the TD group (Figure 4. 5, Table 4. 5).

Table 4. 5 Conjunction anlalysis of FV stimuli. For the ASD group and the TD group clusters of activation for (AV> A) ∩ (AV> V) are shown. Legend: BA — Brodmann's area

Experimental group	Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	t-value	P-value	BA
ASD							
	Superior temporal gyrus	Right	53, -18, 5	5519	6.66	0.00001	22
	Superior temporal gyrus	Left	-51, -19, 8	3669	5.96	0.00003	41
TD							
	Superior temporal gyrus	Right	58, -20, 1	11957	10.21	0.00001	42
	Superior temporal gyrus	left	-56, -20, 7	12700	11.07	0.00001	41

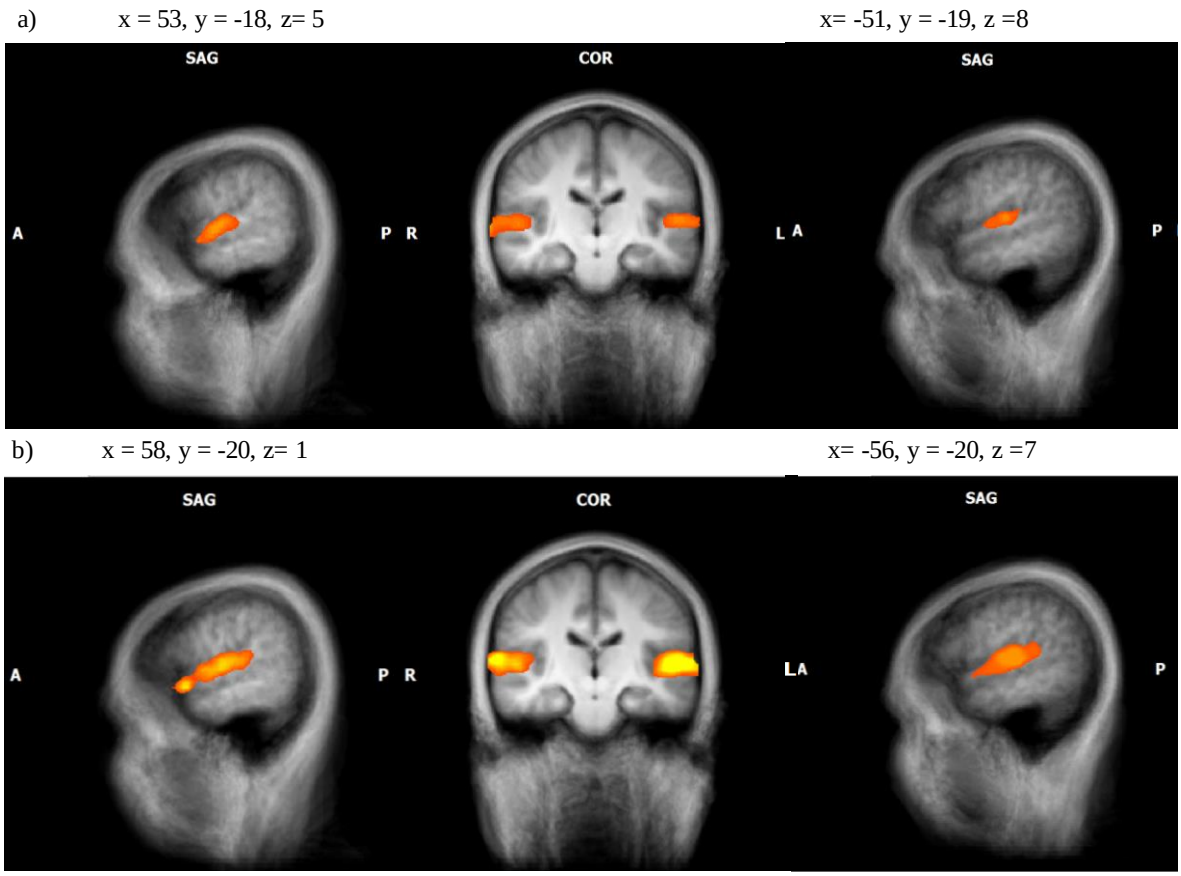


Figure 4.6 fMRI activation data: Conjunction analysis of (audiovisual > audio) ∩ (audiovisual > visual) to define integrative audiovisual regions separately for a) the ASD group and b) the TD group for face-voice stimuli. Coronal and sagittal slices show activation foci at two locations (defined by x y z Talairach coordinates). The clusters of activation sensitive to audiovisual integration are bilaterally in the superior temporal sulcus for both groups of participants (ASD, TD).

The ASD and TD group both showed audiovisual integration sensitive areas in the STS bilaterally when presented with BF and FV stimuli. However, in both the BF and FV contrasts, the two activation maps showed that the ASD group's distributions of the two clusters contained less voxels with a statistically significant conjunction than the TD group. For BF in the ASD group 11, 126 anatomical voxels ($1mm^3$; out of a total of 1,562,139 anatomic voxels) were found to be activated more in response to audiovisual information than audio and visual information alone. In the TD group, more than double the amount of anatomical voxels, 23,371, were found to be more responsive to more audiovisual information. This difference was significant ($p < .01$) according to the Chi-squared test of equality of proportions. Similarly, for FV in the ASD group, we found 9188 anatomical voxels activated more to audiovisual FV stimuli than to audio and visual stimuli alone, whereas, in the TD group, the audiovisual sensitive clusters were larger with 24657 anatomical voxels. This difference was significant ($p < .01$) according to the Chi-squared test of equality of proportions.

4. 4. 2. 3 Between groups conjunction analyses

Comparing the BF and FV conjunction contrasts between the two groups revealed an activation difference for FV in the superior frontal gyrus, showing that in this region the ASD group was less active for audiovisual FV information as defined by (AV> A) ∩ (AV> V) compared to the TD group (Figure 4. 7, Table 4. 6). We did not find group differences for BF.

Table 4. 6 Conjunction analysis of FV stimuli, comparing regions sensitive to audiovisual displays defined by (AV> A) ∩ (AV> V) across the two experimental groups (ASD and TD) Legend: BA — Brodmann's area

Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	t-value	P-value	BA
Superior frontal gyrus	Left	-21, 32, 52	659	-5.08	0.00003	8

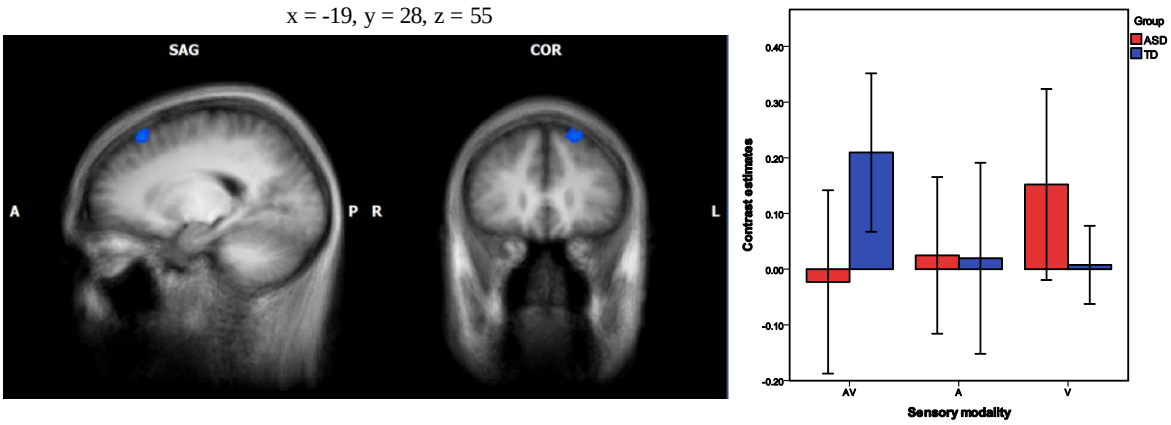


Figure 4. 7 Experiment 1 fMRI activation data: Conjunction analysis of (audiovisual > audio) ∩ (audiovisual > visual) to define integrative audiovisual regions compared across experimental group (ASD, TD) for face-voice stimuli. Coronal and sagittal slices show the activation focus of the location (defined by x y z Talairach coordinates). The cluster of activation sensitive to audiovisual integration is in the left superior frontal gyrus TD individuals but not in ASD.

4. 5 Discussion

In this study we used fMRI to investigate audiovisual, unisensory auditory and visual processing of simple beep-flash, as well as complex and social face-voice displays, in 13 participants with ASD and 13 TD controls. Our results showed cortical activation differences not only in audiovisual conditions, but also in unisensory audio and visual conditions. This suggests that individuals with ASD generally process sensory information differently than TD individuals. Specifically, when presented with simple BF displays of all sensory modalities (AV, A, V), the ASD group exhibited a reduced activation compared to the TD group in the lingual gyri. Additionally, a significant interaction between group and sensory modality was found in right inferior parietal lobule, revealing that the ASD group showed more activation for AV and A than for V, whereas the TD group showed more activation for AV and V than for A. These findings are contradictory to the idea that only the cortical processing of socially relevant auditory and visual information is abnormal in ASD (Gervais et al, 2004; Humphreys, et al., 2008). However, a more generalised sensory processing deficit in ASD is in agreement with our behavioural findings in Chapter 3, which suggested that people with ASD had reduced sensitivity to audiovisual asynchronies across a range of social and non-social displays. For FV displays of all sensory modalities (AV, A, V), we observed reduced activations in the ASD group compared to the TD group in the right middle frontal gyrus and superior frontal gyrus. While it is surprising that we did not reveal any audiovisual specific activation differences, these

findings are consistent with audiovisual and visual face processing research revealing reduced activation in frontal regions in ASD (Pierce et al., 2004; Doyle-Thomas et al., 2013). These findings are in agreement with the results by Williams et al., (2004), which showed that when audio and visual task difficulties were controlled, children with ASD performed no differently on the audiovisual task compared to the controls. However, our results in this study are not in agreement with behavioural findings by Stevenson et al., (2014) and Chapter 3, which suggest intact audio and visual processing in ASD.

The conjunction analyses of BF and FV (AV> A) ∩ (AV> V) for both the ASD and the TD group revealed that the bilateral STG were key to audiovisual integration. Although conjunction analyses for the TD and ASD groups revealed the same regions, for both BF and FV the activated regions in the ASD group were significantly smaller compared to the regions in the TD group. Our findings are consistent with other research showing reduced activation in the STS when presented with voices and as well as audiovisual face-voice displays (Gervais et al, 2004; Doyle-Thomas et al., 2013). Moreover, this finding also suggests an audiovisual specific deficit activation in the STG, which might be revealed more clearly in future studies comprised of larger sample sizes. Comparing audiovisual integration areas between the ASD and TD group revealed that, for audiovisual FV displays, the superior frontal gyrus was less activated in the ASD group than in the TD group. This finding is consistent with previous research showing reduced activation to audiovisual face-voice displays in frontal areas (Doyle-Thomas et al., 2013).

Lingual gyrus

The current study found a reduction of activation of the ASD group compared to the TD group across all sensory modalities of BF stimuli in the lingual gyri. Interestingly, previous research has found the left lingual gurus to be activated less in ASD when performing visuospatial and linguistic reasoning tasks (Sahyoun et al., 2009). However, others have found increased activation of the lingual gyri in ASD during word categorisation tasks (Gaffrey et al., 2007). Structural differences in the lingual gyrus in ASD have also been reported. More specifically, grey matter (GM) and cortical thickness (CT) increases have been found in both lingual gyri in ASD (Ecker et al., 2010; Hyde et al., 2010), whereas, more recently, Ecker et al. (2013) found CT increases in regions of the right lingual gyrus in ASD. In typical participants, research

has also shown that the lingual gyri and lateral occipital cortices are more sensitive to audiovisual stimuli, compared to unisensory stimuli (Vander Wyk et al. 2010; Calvert et al. 2001; Stevenson & James 2009). Furthermore, Petrini et al. (2011) used fMRI to examine brain activity of people watching audiovisual point-light drumming, and showed that when there was a natural covariation between sound intensity and velocity of the drumming strike, the lingual gyrus was more activated compared to displays in which this natural covariation was eliminated.

Inferior parietal lobule

For BF stimuli, we found that in ASD the right inferior parietal lobule was more activated during AV and A stimulation than in V stimulation, while in the TD group, it was more activated for AV and V than for A. Thus the AV stimuli activated this area comparably across groups, but the unisensory conditions lead to different activation patterns. Although the inferior parietal lobule has previously been shown to have atypical activation patterns in ASD visual processing (Huble, et al., 2003), these findings contradict the finding by Doyle-Thomas et al. (2013), showing that during audiovisual emotional matching tasks people with ASD relied more on areas of the parietal lobe including the middle parietal lobule and precuneus. Li, Xue, Ellmore, Frye and Wong (2014) used diffusion tensor imaging (DTI) to show stronger local connectivity in ASD in inferior parietal regions including the BA 40. The right inferior parietal lobule is also an area commonly associated with grey matter and cortical thickness differences in ASD (Ecker et al., 2010; Hyde et al, 2010; Wallace et al., 2010; Hadjikhani et al., 2006; Doyle-Thomas et al., 2013), which can be linked to atypical activation.

The importance of the inferior parietal lobule in the integration of audiovisual information has been also been demonstrated (Calvert et al., 2000, 2001; Dhamala et al., 2007). Although, for speech perception, the left inferior parietal lobule has often been shown to be involved in audiovisual speech perception (Miller & D'Esposito, 2005; van Wassenhove, Nusbaum & Small, 2007; Szycik, Tausche & Münte, 2008), the right inferior parietal lobule might also play a specific role in both unimodal and multimodal event order judgments (Snyder & Chatterjee, 2004; Battelli, Pascual-Leone & Cavanagh, 2007), and has been suggested to contribute to the perception of synchrony between events across sensory modalities.

Premotor BA 6

The ASD group showed reduced activation in Brodmann area 6 in the superior frontal gyrus when observing FV stimuli of different modalities. Interestingly, the area lies within the premotor cortex, an area which has previously been associated with audiovisual perception. Neuroimaging studies have shown that areas like the premotor cortex are not only involved in speech production, but also help speech perception (Meister, Wilson, Deblieck, Wu & Iacoboni, 2007). It has also been suggested to facilitate speech perception by mapping unimodal and multimodal sensory features onto articulatory speech gestures (Callan, Jones, & Callan, 2014). Similarly, the superior precentral cortex has been found to be involved in audiovisual sentence processing (Capek et al., 2004). Interestingly, speech perception in ASD has been shown to be atypical. For example, in ASD speech perception has been shown to be less influenced by a talking face than in TD peers (de Gelder et al., 1991; Irwin et al., 2011; Mongolli et al., 2008). The premotor cortex has also been activated by execution, as well as observation of execution of action (a mirror neuron system property) (Callen et al., 2004; Mashal, Solodkin, Dick, Chen & Small, 2012). People with ASD have been frequently shown to have an executive functioning deficit (e.g., Liss, et al. 2001). Intriguingly, in ASD, increases of CT in the superior temporal gyrus were found (Waiter et al., 2004, McAlonan et al., 2002, Ecker et al, 2013), whereas reductions of GM volume in ASD were found by McAlonan et al. (2002).

Dorsolateral prefrontal cortex BA 8 and BA 46

Comparing the FV conjunction contrasts between the two groups revealed that for an area in the superior frontal gyrus, BA 8, the ASD group was less active during audiovisual FV displays as defined by (AV> A) ∩ (AV> V), compared to the TD group. Area BA 46 in the middle frontal gyrus revealed a reduction of activation in the ASD group, compared to the TD group across all sensory modalities of FV displays. These results are in line with Doyle-Thomas et al.'s (2013) results, which also revealed reduced activation in ASD in the middle frontal gyrus during audiovisual tasks of emotion matching. However, Loveland et al., (2008) in their pilot study, revealed a higher activation in the right middle frontal gyrus in their ASD participants compared to their controls. The BA 8 and BA 46 are both part of the dorsolateral

prefrontal cortex (DLPFC). In distinction to the activation pattern we found in BA 46 (reduced activation during stimuli of all modalities), a study recently showed that people with ASD have greater activation in the DLPFC when attending to faces and houses (Herrington, Riley, Grupe & Schultz, 2015). In BA 8, we found a similar trend of increased activation to our face stimulus in the ASD group. Moreover, more extensive connectivity in ASD between the thalamus and the middle frontal regions has been found (Mizuno, Villalobos, Davies, Dahl & Müller, 2006). Furthermore, evidence that DLPFC is activated in typical audiovisual processing in sentence processing and temporal order judgements (Capek et al., 2004; Adhikari, Goshorn, Lamichhane & Dhamala, 2013) supports the discovery of BA 8 being an audiovisual integration area in the TD group. Evidence from structural studies, revealing GM volume and CT increases and decreases in ASD in the right inferior frontal cortex, also support the notion of atypical cortical activation in those areas (Ecker et al., 2010; Hyde et al, 2010; Hadjikhani et al., 2006; Doyle-Thomas et al., 2013; Ecker et al, 2013).

Conclusion

The results of the current study reveal that audiovisual and unisensory auditory and visual processing of both social face-voice and simple beep-flash stimuli are associated with reduced activation in ASD. Audiovisual, auditory and visual conditions of human face-voice stimuli revealed reduced activation in ASD participants compared to TD participants in regions of the frontal cortex, while beep-flash stimuli revealed reduced activation in the lingual gyri. The inferior parietal gyrus revealed that its activation was modulated differently by the different sensory stimulus conditions of visual-flash stimuli in ASD and TD participants. Specifically, we found increased activation in audiovisual and auditory conditions compared to the visual condition in individuals with ASD, while TD controls showed increased activation in audiovisual and visual conditions compared to the auditory condition. Although smaller regions of the STC were found in ASD to be sensitive to audiovisual stimuli as computed by conjunction analyses, against our predictions, we did not find any activation differences, per se, of the STC between the two groups. However, a superior frontal area was shown to be sensitive to audiovisual face-voice stimuli in the TD group, but not in the ASD group. Overall, this study has indicated that brain

activity, prompted by audiovisual, auditory and visual processing of social and non-social stimuli, is different in people with ASD compared to TD.

These results are in contrast to previous behavioural findings (Chapter 3), which suggested deficient audiovisual integration, while auditory and visual processing is intact. The current results reveal the need for further investigations to explain the relationship between our results and those found in Chapter 3. Chapter 5 will look more specifically at the neural correlates of Synchrony Judgements in ASD and TD participants.

5 fMRI investigation of audiovisual temporal processing in Autism Spectrum Disorder using Synchrony Judgements

5.1 Abstract

The integration of information from different senses is important in everyday life. In Autism Spectrum Disorder (ASD), behavioural results have shown deficits in audiovisual integration. This view also been supported by electroencephalography (EEG) studies (Brandwein et al., 2015). Using functional magnetic resonance imaging (fMRI), we investigated audiovisual temporal processing in ASD. In 13 adult males with ASD and 13 age-, sex-, and IQ-matched typically developed (TD) controls, we investigated temporal asynchrony of audio and visual information in simple beep-flash (BF) displays, as well as complex and social face-voice (FV) displays. Blood oxygenation level dependent (BOLD) signals were measured while the ASD and TD participants were asked to make synchrony judgements (SJ) on audiovisual displays of different levels of asynchrony: the participants' point of subjective simultaneity (PSS), audio leading visual information (audio first), visual leading audio information (visual first). Whereas no effect of group was found with BF displays, increased putamen activation was observed in ASD participants compared to TD participants when making SJs on FV displays. Investigating SJ on audiovisual displays in the bilateral superior temporal gyrus (STG), an area involved in audiovisual integration (see Chapter 4), we found no group differences or interaction between group and levels of audiovisual asynchrony. The investigation of different levels of asynchrony revealed a complex pattern of results, indicating a network of areas more involved in processing PSS than audio first and visual first, as well as areas responding differently to audio first compared to video first. These activation differences between audio first and video first stimuli in different brain areas are constant with the view that audio leading and visual leading stimuli are processed differently.

5.2 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterised by the DSM-V as deficits in social communication and interaction, as well as repetitive patterns of behaviour, interests and activities (American Psychiatric Association, 2013). Differences in sensory processing in ASD compared to typically developed

(TD) individuals has been reported across the different sensory modalities (Lane et al., 2015, Ludlow et al., 2014, Simmons et al., 2009). Recently, Robertson and Simmons (2013) demonstrated a strong correlation between autistic traits in the general population and sensory sensitivities. The relevance of these sensory differences has recently been recognised and included as diagnostic criteria in the DSM-V. In fact, researchers have started to stress the importance of understanding to what degree sensory anomalies in individuals with ASD contribute to their social and communication impairments. For example, multisensory processing differences have been proposed to cascade down to causing communication impairments (Brandwein et al., 2015) or acts of apparent non-compliance, reluctance, lack of interest and aggression (Donnellan, Hill & Leary, 2013). This highlights the importance of better understanding the neural correlates of these sensory processing differences and potentially developing sensory processing interventions.

The ability to integrate auditory and visual information is crucial to everyday life. The combination of multiple sensory cues reduces uncertainty and enables us to make better estimates of situations (Ernst & Banks, 2002). For example, in a crowded and noisy environment, looking at a person's face and lip movements enables us to better understand what a person is telling us. Moreover, to interact appropriately with the environment, the multisensory integration process needs a degree of specificity to combine only the information that belongs together and keep other information apart. To make a judgement of whether audio and visual information belongs together, the temporal correspondence of the two incoming cues need to be considered. When looking at the temporal aspect of audiovisual integration, it has been shown that individuals integrate incoming audio and visual information, and perceive them as one unitary event, even if they are hundred milliseconds or more apart; this is called the temporal integration window (TIW; as seen in Chapter 3; Hairston et al., 2005; Love et al., 2013; Petrini et al., 2009a,b; Stevenson & Wallace, 2013; Stevenson et al., 2014; van Wassenhove et al., 2007; van Eijk et al., 2008). Usually the participants are shown audiovisual stimulus pairs with varying stimulus onset asynchronies (SOAs) between the audio and visual information. As shown and discussed in Chapter 3, the TIW can be measured by Synchrony Judgements (SJ) and Temporal Order Judgements (TOJ), which are likely to tap into different perceptual mechanisms (Love, et al., 2013; Love et al., in preparation). The width of the TIW varies from

participant to participant (Stevenson, Zemtsov, & Wallace, 2012), as well as across people's life span. The TIW has been shown to be wider in childhood, becomes narrower in late adolescence (Hillock et al., 2011; Hillock-Dunn and Wallace, 2012a), and becomes wider again in late adulthood (Chan, Pianta & McKendrick, 2014a).

Temporal audiovisual processing in TD individuals

In Chapter 4 we showed the importance of the superior temporal gyrus (STG) in audiovisual integration, however, when temporal processing aspects are introduced, the role of the STG is less clear (Miller & D'Esposito, 2005; Stevenson et al., 2010; Stevenson, VanDerKlok, Pisoni, & James, 2011; Stevenson, Mullin, Wallace & Steeves, 2013; Love, 2011). These studies have specifically looked at audiovisual synchrony judgements (SJ) which ask participants to judge whether audiovisual stimuli, presented at a range of different levels of asynchrony, are in synch or out of synch. A network of regions responding more to synchronous than asynchronous speech, including right mSTG, bilateral superior colliculus, fusiform gyrus, lateral occipital cortex, and extrastriate visual cortex, has been found (Stevenson et al., 2010). Similarly, Love (2011) examined participants' brain activation during synchrony judgment tasks on synchronous and asynchronous audiovisual speech displays. Similarly to Stevenson et al., 2010, he showed an asynchrony network and a synchrony network, but only for perceptually (audio preceding visual information by about 90 ms) defined synchrony and not when it was physically (SOA = 0) defined synchronous speech. He also distinguished two regions of the superior temporal cortex (STC): a middle region of STC, responding to synchronous speech, and a posterior region, responding to asynchronous speech.

Furthermore, Love et al. (in preparation) compared the neural mechanism underlying SJs with temporal order judgement tasks (TOJs), which act as another popular task used to assess people's sensitivity to asynchrony. Using audiovisual drumming displays at different levels of asynchrony (audio first, video first) and synchrony [physically synchronous, participant's point of subjective simultaneity (PSS)], their results showed that the two judgements use different brain areas. The middle occipital cortex was found to show sustained activation during SJ and deactivated during TOJ, whereas transient activation was greater in TOJ than in SJ in regions of the left middle

occipital, middle frontal, precuneus and medial superior frontal lobe. This can be taken as evidence that the SJ and TOJ measure different aspects of audiovisual synchrony perception. A recent repetitive transcranial magnetic stimulation (rTMS) study by Stevenson et al. (2013) revealed that stimulation prior to making SJs on beep-flash stimuli of the multisensory region (STS) caused an overall widening of the TIW (increased tolerance for visual-leading stimuli). Furthermore, stimulation of auditory (Hechl's gyrus) and visual (striate cortex) regions caused a broadening within the audio first stimuli and video first stimuli, respectively. The broadening of the TIW, to the more ecologically valid visual first stimuli with STS disruption, advocates that audiovisual temporal processing in STS reflects learned environmental information.

Audiovisual processing in ASD

From behavioural research we know that for individuals with ASD, processing multiple sensory signals as a single percept is not obtained as frequently as for TD individuals. For instance, people with ASD perceive audiovisual illusions such as the McGurk effect (McGurk & MacDoland, 1976) less often than their TD controls (de Gelder et al., 1991; Irwin et al., 2011; Mongolli et al., 2008); often benefiting less from an additional sensory modality (Smith & Bennetto, 2007) and relying more on one sensory modality (Stevenson et al., 2014).

Electroencephalography (EEG) studies recording high-density brain activity have shown that the neural integration of audiovisual information is deficient in children with ASD (Magnee et al., 2009; Russo et al., 2010; Brandwein et al., 2013, 2015). Moreover, using fMRI, it has been shown that adolescents with ASD use different cortical areas when processing audiovisual emotion stimuli compared to TD adolescents (Doyle-Thomas et al., 2013). The activation patterns revealed that the ASD group employed more parietal and frontal cortices, whereas the TD group recruited frontal and temporal cortices during this task. It was suggested that the absence of integrative emotional networks in ASD might cause the recruitment of the parietofrontal network as a compensatory result. Similarly higher activation patterns were shown in a pilot study which asked participants to make emotional congruency judgements (Loveland et al., 2008). Moreover, Chapter 4 revealed people with ASD exhibit reduced activation when presented with audiovisual, unisensory auditory and visual simple beep-flash, and social face-voice stimuli. Audiovisual, auditory and

visual conditions of face-voice stimuli revealed reduced activation in ASD in regions of the frontal cortex, while beep-flash stimuli revealed reduced activation in the lingual gyri. Activation levels in the inferior parietal gyrus revealed an interaction between the multisensory and sensory stimulations and experimental group. While the conjunction analyses highlighted the STG to be significantly involved in audiovisual integration in both groups, the ASD group revealed smaller regions of the STG to be audiovisual sensitive compared to the controls. Interestingly, a superior frontal area was shown to be sensitive to audiovisual face-voice stimuli in the TD group, but not in the ASD group. Overall, this study suggests that the processing of audiovisual, auditory and visually presented social and non-social stimuli are different in people with ASD compared to TD people.

Temporal audiovisual processing in ASD

Thus far the literature on multisensory processing in ASD has concentrated on behavioural responses to demonstrate audiovisual temporal processing differences. Commonly, the TIW is shown to be wider in ASD (Chapter 3, Stevenson et al., 2014; de Boer-Schellekens, Eussen & JeanVroomen, 2013). The sensitivity to asynchrony has been measured using video clips of simple beeps and flashes, complex audiovisual human actions and audiovisual speech (Chapter 3; de Boer-Schellekens, Eussen & JeanVroomen, 2013; Grossman, Schneps & Tager-Flusberg, 2009; Stevenson et at., 2014), as well as using different tasks: SJs (Chapter 3; Grossman, Schneps & Tager-Flusberg, 2009; Stevenson et at., 2014) or TOJs (Chapter 3; de Boer-Schellekens, Eussen & JeanVroomen, 2013).

Although behavioural evidence is of great importance, it is crucial to obtain a deeper understanding of the processes underlying the differences in audiovisual integration in ASD. Thus far there has been little research investigating the neural underpinnings of the differences in audiovisual integration in ASD. While Chapter 4 highlights that even basic fMRI research investigating audiovisual, audio and visual processing in ASD is lacking, investigating audiovisual temporal processing during SJ tasks would enable us to reveal the neural correlates underlying the observed behavioural differences. Moreover, behavioural similarities do not always mean that the underlying brain processes are the same, since compensatory mechanisms might have

been developed (McKay et al., 2013). Therefore it is important to investigate what neural substrates underlie these audiovisual processing differences.
Looking at the behavioural results in Chapter 3 and the fMRI research on audiovisual integration and synchrony perception, we can postulate how a widened TIW might be reflected in the underlying neural correlates that serve the integration of audiovisual information. As we have shown in Chapter 4, the STG is clearly an important neural centre of multisensory integration processes. In ASD, structural differences within the STG and posterior STS are a common finding within the grey matter (Scheel et al., 2011; Greimel et al., 2013; Doyle-Thomas et al., 2013; Hyde et al., 2010) and white matter (Barnea-Goraly et al ., 2004). Additionally, functional activation differences of the STG in ASD have been commonly observed (Gusnard & Raichle, 2001; Buckner et al., 2008). These atypicalities of the STG in ASD have been proposed to be the underlying cause of the common aetiology for audiovisual temporal processing deficits observed in ASD and other developmental conditions (Stevenson et al., 2014; Wallace & Stevenson, 2014).

To our knowledge, no study to date has investigated audiovisual temporal processing in ASD using neuroimaging techniques. Therefore we used fMRI to investigate the underlying neural correlates of audiovisual temporal integration in ASD when making SJs on simple beep-flash (BF) and more complex and social face-voice (FV) displays. Thus the current experiment measured BOLD signals while participants were asked to make SJ on audiovisual BF and FV displays which were: perceptually synchronous to the individual participant (PSS; previously collected in Chapter 3), asynchronous with audio leading visual information (audio first; AF) and visual leading the auditory information (video first; VF). A whole-brain analysis was run to establish whether ASD and TD individuals recruit the same areas when processing audiovisual displays of different levels of asynchrony. The regions established through the conjunction analysis in Chapter 4 were also used to create masks of regions sensitive to audiovisual integration to see whether we would find any group differences, specifically in those areas, when processing audiovisual displays of different asynchronies. Moreover, the masks are of the STG, and therefore will allow us to investigate the predictions from other research that atypicalities of the STG in ASD could be an underlying cause of audiovisual temporal processing deficits observed in ASD.

5.3 Methods

5.3. 1 Participants

The participants in this study had also taken part in Chapter 4. Thirteen high-functioning adult males with ASD (aged between 21 and 41) and 13 age-, sex- and IQ-matched control participants (aged between 21 and 41) took part in this fMRI study (Table 5. 1). All participants in the ASD group reported to have a diagnosis of having an ASD according to DSM-IV criteria from a qualified clinician. All were native English speakers, had normal or corrected to normal vision and reported no hearing difficulties. The Autism Quotient (AQ), a 50 item autism traits questionnaire by Baron-Cohen, Wheelwright et al. (2001), supported the diagnoses of the ASD group (M= 37.54, SD = 6.89) and reinforced the assumption that no-one in the TD group had ASD (M=12.31, SD= 4.09). The participants were matched pair-wise on age (t (12) =.82, p=.42) and group-wise on full scale IQ (FSIQ) (t (12) =.51, p=.62) as measured using the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). The experimental procedures were approved by the School of Psychology at the University of Glasgow and also the Greater Glasgow and Clyde National Health Service ethics board.

Group	Age		FSIQ	
	Mean	SD	Mean	SD
ASD	30.54	7.42	119.92	10.13
TD	29.46	5.34	118.23	6.46

Table 5. 1 Mean and Standard Deviation (SD) of the ages and Full Scale IQs (FSIQ) of the autism spectrum disorder (ASD) and typically developed (TD) group separately.

5. 3. 2 Stimuli

Participants were presented with AV displays showing the participants' subjectively perceived level of synchrony (PSS), asynchrony with audio first (AF) in which audio information precedes visual information, and visual first (VF) in which visual information precedes audio information. Similar to the experiment in Chapter 4, this experiment also contained non-social BF displays and complex social FV displays (Figure 5. 1). The timings for the PSS displays were obtained from the previous behavioural experiment described in Chapter 3 (see Table 5. 2). We used the individual PSS estimates from the Independent Channels Model (ICM) fit, as this fitting procedure takes asynchronies and variabilities of the response data (Garcia-

Perez & Alcala-Quintana, 2012). The PSS obtained from BF and FV revealed no group differences, t(24)=0.56, p= .58 and t(24) = 0.84, p=.41, respectively.

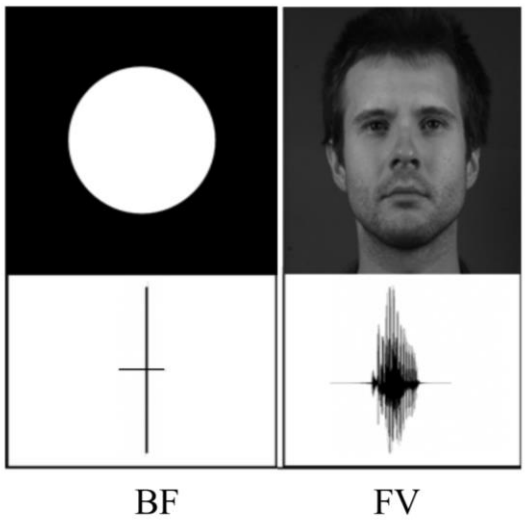


Figure 5. 1 The top panel shows the visual information participants were presented with. The bottom panel shows the auditory waveform for each type of stimulus. The beep-flash (BF) stimulus consisted of a flash of a white dot on a black background and a beep. In the face-voice (FV) stimulus a movie frame is shown and the waveform represents the word "tomorrow". Please note that the images are not to scale, the area of the white flash dot is approximately the same size as the area of the mouth in FV.

ASD participants	**PSS (ms)**		**TD participants**	**PSS (ms)**	
	BF	**FV**		**BF**	**FV**
1	2.1	-89.79	**1**	9.62	-6.94
2	108.12	82.47	**2**	27.37	42.74
3	22.69	18	**3**	25.36	11.38
4	47.84	44.85	**4**	23.95	48.93
5	-29.29	-51.47	**5**	19.12	-50.46
6	154.44	13.99	**6**	25.37	50.26
7	-29.79	12.31	**7**	74.77	-68.04
8	51.48	46.8	**8**	53.14	9.01
9	78.59	-81.7	**9**	43.19	53.14
10	-5.72	43.44	**10**	-0.83	9.73
11	44.82	-25.41	**11**	40.43	-19.59
12	61.92	-15.99	**12**	57.66	23.76
13	63.27	-55.52	**13**	56.38	38.19
Mean	43.88	-4.46	**Mean**	35.04	10.93
SD	53.05	53.75	**SD**	21.46	38.68

Table 5. 2 shows each participant's perceived level of synchrony when presented with BF and FB audiovisual stimuli. A positive value indicated that the participant's PSS was when audio information was leading visual information, whereas a positive PSS indicated that the participant's PSS was when the visual information was leading the audio information. ASD = Autism Spectrum Disorder, TD = typically developed, PSS = point of subjective synchrony (PSS) as measured by the Independent Channels model in Chapter 3, BF= beep-flash, FV = face-voice, SD = standard deviation

All videos were created to be the same duration, irrespective of their degree of asynchrony (PSS, AF, VF). For BF, all videos were 816 ms long with a frame rate of 60 frames per second. The flash was a white dot (luminance: 85 cd/m^2; visual angle of the diameter: 4.4 degrees) on a black background, while the beep was a pure tone at 2000 Hz and 84 dB mean intensity. In the synchronous condition, the flash and the beep started at 400 ms, while in the asynchronous conditions and the participants' individual PSS, the flash always started at 400 ms and the beep was shifted to be presented either before (for audio leading) or after the flash (for video leading). In the asynchronous videos, the beep was presented 333ms before or after the flash. The PSS videos were individually created for each participant. See Table 5.2 for more details.

The FV videos were 1920 milliseconds long, with a frame rate of 25 frames per second. The video was of a man saying the word "Tomorrow". The visual angle of the visual speech cue was approximately 12.7 and 18.2 degrees and the mouth region covered about 3.2 by 2.5 degrees of the visual angle. This made the mouth region approximately the same size as the flash in the BF videos. The voice was shifted to come either before or after the lips moved. Before and after the word was spoken, a still image of the first and the last frame of the video was used to fill the gap that the shifting of the audio stream created. These still images faded in at 120 ms and faded out at 1840 ms. In the asynchronous videos, the audio stream was presented 400 ms before or after the video stream. The videos had a black background that matched the screen background to minimise the predictability of the cues. In neither of the movies was the sound clipped at the beginning or the end.

5. 3. 3 Design

Two event-related fMRI experiments aimed to examine activation differences while making SJs of AV displays at different levels of asynchrony: participants' PSS, AF and VF. The sequence of the stimuli presented in the event-related fMRI study was optimised for detecting signals between event types using the Genetic Algorithm (GA) developed by Wager and Nichols (2003). The one-back counterbalancing of the optimal sequences was checked and shown to be well counterbalanced. The GA has been shown to optimise contrast efficiency of event-related designs compared to randomised, fully counterbalanced and m-sequence (maximum length shift-register

sequence) designs (Wager & Nichols, 2003; Kao, Mandal, Lazar & Stufken, 2009). It is also less strenuous than going through all possible permutations of designs to find the ones with maximum efficiency. The GA for each run was set up to generate 100000 iterations of 1000 designs, each with a maximum running time set to 12 hours. Temporal jitter of the inter stimulus intervals (ISI) was also introduced to the design in order to further maximise the efficiency of the design.

We ran the GA four times and used the four most efficient sequences for the BF and FV experiments. Each participant was shown the same two sequences (one for each run) for both experiments. The two runs of each experiment were always shown together, however the order of presentation was counterbalanced. The order in which the BF and the FV experiments were presented was also counterbalanced. Furthermore, half of the participants of each group (ASD, TD) were presented with the sequences one and two, whereas the other half of the participants were presented with sequences three and four.

Each run contained 22 stimulus displays for each of the 3 conditions (PSS, AF and VF). The lengths of the ISIs were on average 5.3 seconds (ranging from 5-15 seconds) for the BF experiment, and 4.3 seconds (ranging from 4-14 seconds) for the FV experiment. During the ISIs, participants were presented with a black background. The stimulus length was 816 ms for the BF displays and 1920 ms for the FV displays. In total, each BF run was 444 seconds long, with 20 seconds rest before and 16 seconds rest after each run. The FV runs were 448 seconds long, with 20 seconds rest at the beginning and 16 seconds rest at the end.

5. 3. 4 Procedure

For BF and FV experiments, participants were presented with two runs of audiovisual displays. The order of the experiments was counterbalanced, and so were the runs within each experiment. Each experiment contained audiovisual displays of three different levels of asynchrony (PSS, AF, VF). The participants were instructed to make SJs and thus were asked to judge whether the audio and visual information in the displays were in synch or out of synch. This was a forced choice task and participants responded by pressing one of two buttons. The buttons corresponding to the “in synch” and “out of synch” responses were counterbalanced. Overall this study took about 35 minutes to complete. In the same way as in Chapter 4, the participants

walked through the scanner safety checklist to ensure that they were safe to be scanned. Before entering the scanner we explained the task to the participants and allowed them to ask any questions. All participants provided informed written consent. If needed, the participants vision was corrected using the Nordic Neurolabs Visualsystem goggles until participants were able to clearly see the stimuli and instructions. Stimuli were presented using Presentation 14.9 designed by NeuroBehavioral Systems (NBS), via electrostatic earphones (NordicNeuroLab, Norway) at a sound pressure level of 80 dB. In between scans we checked that participants found the sound pressure level comfortable and loud enough considering the scanner noise. After the study, everyone was reimbursed for their time and transportation.

5. 3. 5 Data acquisition parameters

A Siemens 3T Tim Trio MRI scanner was used to acquire sagittal T1 weighted anatomical images and T2 weighted functional images.

5. 3. 5. 1 Functional data

Functional T2 weighted images were acquired (interleaved, TR= 2 seconds, TE = 30ms, Flip angle = 90°). We collected 32 slices for each of 222 and 224 volumes for the BF and FV sub-experiment, respectively, at a resolution of 3mm x 3mm x 3mm voxel size resolution and dimensions 210 x 210 mm per image with online motion correction. The first 2 volumes of each functional run comprised 'dummy' gradient and radio frequency pulses, which permitted for steady state magnetisation. During these volumes no stimuli were presented and no fMRI data was collected. The data sets used for the analysis were the motion corrected (moco) series output by the Siemens system.

5. 3. 5. 2 Structural data

The high-resolution T1-weighted structural images collected were the same as in Chapter 4 and comprised of 192 axial slices and isotropic voxels (resolution: 1 mm x 1mm x 1 mm; dimensions: 256 x 256 mm, TR = 1900 msec, TE = 2.92 msec, time to inversion = 900 msec, FA = 9°). The run time was 10 minutes.

5. 3. 6 fMRI Preprocessing

BrainVoyager QX version 2.8 was used to preprocess and analyse all stages of the fMRI data. The first two functional volumes were excluded to allow for signal stabilisation. The preprocessed (homogeneity corrected and in Talairach space) structural scans from Chapter 4 were used. Functional runs were slice scan time corrected, 3D motion corrected (using trilinear/sinc interpolation) and temporally high-pass filtered at 4 cycles across each run. The functional runs were coregistered to the 1 × 1 × 1 triliniar-interpolated anatomical maps scans and transformed into Talairach space. A Gaussian 6mm spatial filter was applied to the 4D volumes in order to improve the signal to noise ratio for group analysis by overcoming differences in intersubjective localisation.

5. 3. 7 Analyses

A random-effects general linear model (RFX GLM) was used to compute first-level statistics on the z-normalised BOLD signal for each individual. Using the GLM, parameter estimates for each condition were calculated for each voxel within the brain. The first level analysis results were then entered into the second-level random-effects analyses to account for variability between subjects and to allow generalisations at a population level.

5. 3. 7. 1 Whole-Brain GLM

For both BF and FV experiments, a 2 (group: ASD, TD) × 3 (level of asynchrony: PSS, AF, VF) RFX GLM was run, with group as the between-subjects factor and level of asynchrony as the within-subject factor. Cluster size threshold estimation was used to control for multiple comparisons.

5. 3. 7. 2 Audiovisual synchrony analysis restricted to audiovisual integration regions

We performed 2 (group: ASD, TD) × 3 (level of asynchrony: PSS, AF, VF) RFX GLM on the beta values from the first order statistical analysis, restricted to only those voxels that were found to be significantly more active in the audiovisual conditions than the audio and visual only conditions, as revealed by the conjunction analyses in Chapter 4. These voxels were used to create masks to restrict analysis to only regions sensitive to audiovisual integration. Four separate masks were created from the results of the conjunction analyses: one for each group (ASD and TD) as well as for the different display types (BF and FV). The audiovisual sensitive regions were defined

by the max-criterion, as it has previously been shown to be an appropriate criterion to establish regions sensitive to audiovisual information (Kreifelts et al., 2010; Szycik et al., 2008; Love, Pollick, & Latinus, 2011; Watson et al., 2014). Applying the mask greatly reduced the number of voxels that were entered into the GLM compared to the unconstrincted whole-brain analysis. This provided sufficient power to run a REX GLM and be able to apply the results to the wider population.

5. 4 Results

5. 4. 1. Whole-brain GLM of BF displays

For BF displays, a 2 (group: ASD, TD) x 3 (level of asynchrony: PSS, AF, VF) RFX GLM revealed a significant main effect of level of asynchrony in supramarginal gyrus (bilaterally), left medial frontal gyrus, left precentral gyrus and the right middle temporal gyrus, but no main of group or interaction between group and level of asynchrony (Figure 5. 2, Table 5. 3). The left middle frontal gyrus revealed reduced activation when perceiving PSS displays compared to AF and VF (Figure 5.2 a). Similarly, we found that the PSS led to reduced activation compared VF in the precentral gyrus (Figure 5. 2 b). The putamen responded most to the AF condition, compared to PSS and VF (Figure 5. 2 c). Lastly, the bilateral supramarginal gyri showed reduced activation in the PSS condition compared to AF (Figure 5.2 d and e).

Table 5. 3 Experiment 2a BF clusters of activation from a 2 × 3 ANOVA with 'experimental group' as a between-participants factor and 'level of asynchrony' as a within-participants factor. Legend: BA — Brodmann's area

Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	F-value	P-value	BA
Level of asynchrony (PSS, AF, VF)						
Supramarginal gyrus	Right	48, -46, 31	540	12.13	0.00005	40
Putamen	Right	27, -10, 4	178	10.10	0.00022	-
Precentral gyrus	Left	-36, -1, 34	520	11.77	0.00007	6
Middle frontal gyrus	Left	-48, 47, 10	278	15.87	0.00001	10
Supramarginal gyrus	Left	-54, -55, 31	302	12.37	0.00005	40

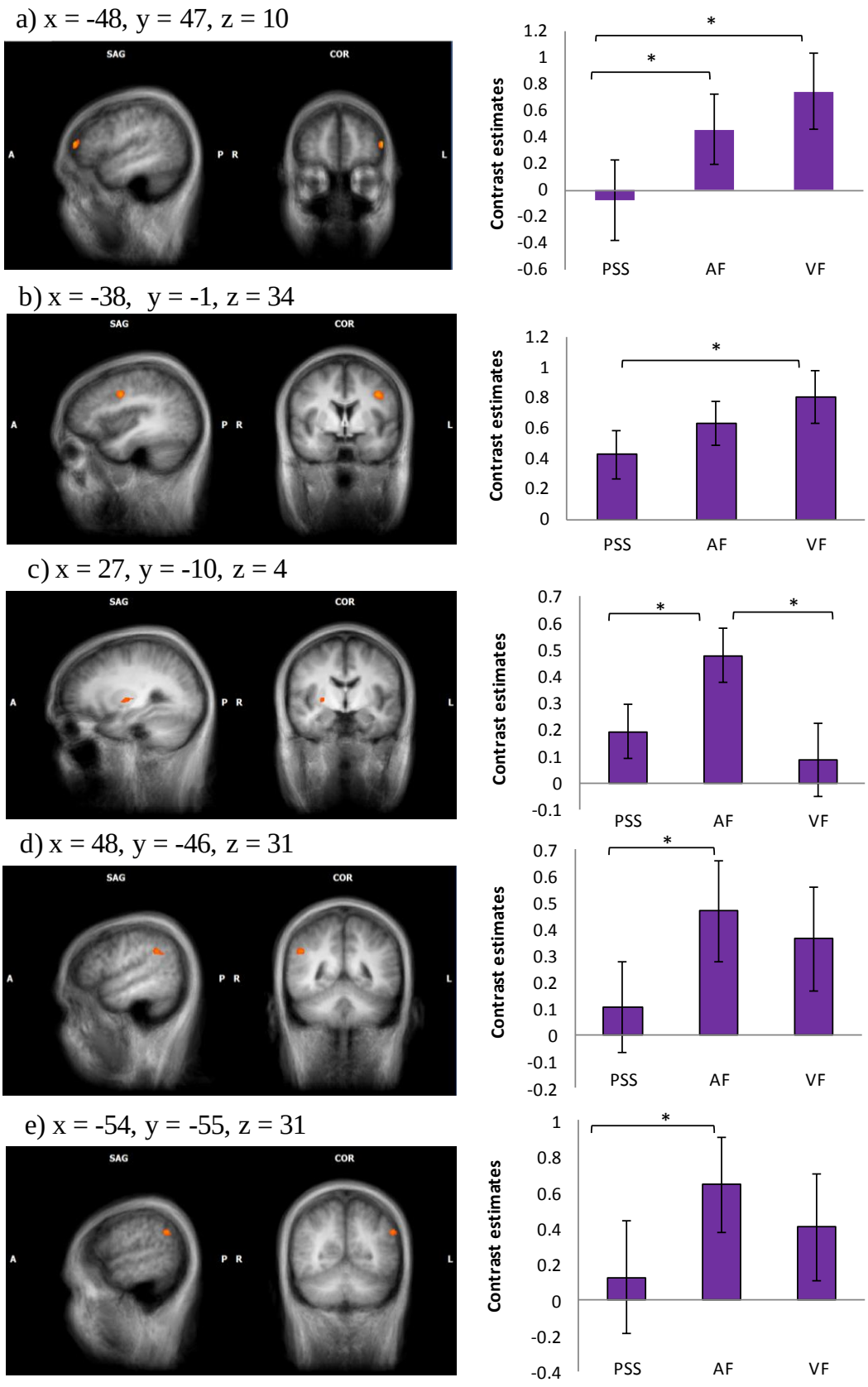


Figure 5.2 Beep-flash experiment fMRI activation data: Clusters of activation for which the brain responses, collapsed across groups (ASD, TD), differed across the three beep-flash stimulus conditions: point of subjective simultaneity, audio first and visual first. Sagittal and coronal slices show activation foci at locations defined by x y z Talairach coordinates. The average contrast estimates (beta weights) and relative standard errors are shown in the histograms. The brackets and * indicate where the pairwise comparisons found significant differences ($p<0.05$) between the conditions. Point of subjective simultaneity (PSS), audio first (AF), visual first (VF). a) middle frontal gyrus, b) precentral gyrus, c) putamen, d) right supramarginal gyrus, e) left supramarginal gyrus

5. 4. 2 Whole-brain GLM of FV displays

For FV displays, a 2 (group: ASD, TD) x 3 (level of asynchrony: PSS, AF, VF) RFX GLM revealed a significant main of group in the right putamen (Figure 5.3, Table 5.4). A significant main effect of level of asynchrony was found in the bilateral precentral, gyrus, cingulate gyrus, right transverse temporal gyrus, angular gyrus, inferior frontal gyrus, superior frontal, insular and left caudate, declive, middle temporal gyrus (Figure 5. 4, Table 5.4). However, we did not find an interaction between group and level of asynchrony.

The main effect of group in the putamen revealed an increased activation in the ASD group compared to the TD group, across all levels of asynchrony (Figure 5.3).

The main effect of asynchrony in the right inferior frontal gyrus (Figure 5.4 a) revealed that conditions PSS and AF elicited stronger activation than VF. The right superior frontal gyrus (Figure 5.4 b) responded more to the PSS condition than to AF. The precentral gyri (Figure 5.4 c and d) responded more to conditions PSS and VF than to AF. The right superior temporal gyrus (Figure 5.4 e) only revealed a difference in activation between AF and VF, with a higher activation for VF. In the left, declive (Figure 5.4 f) activation was higher for PSS compared to AF and VF. The right insular revealed a stronger activation to VF and to AF (Figure 5.4 g). The cingulate gyri l) and the right precuneus (Figure 5.4 h, i and l) elicited stronger activation to PSS and VF than for AF. Increased activation to the AF and VF conditions, compared to PSS, were found in the left caudate (Figure 5.4 j), while the right angular gyrus revealed increased activation to the PSS condition (Figure 5.4 k). Lastly, the left middle temporal gyrus revealed higher activation to PSS than to AF (Figure 5.4 m).

Table 5. 4 Experiment 2b FV clusters of activation from a 2 × 3 ANOVA with 'experimental group' as a between-participants factor and 'level of asynchrony' as a within-participants factor. Legend: BA — Brodmann's area

Anatomical region	Hemisphere	Talairach coordinate of peak voxel (x, y, z)	Number of voxels	F-value	P-value	BA
Group (ASD, TD)						
Putamen	Right	27, -1, 10	181	21.81	0.00010	-
Level of asynchrony (PSS, AF, VF)						
Superior temporal gyrus	Right	57, -22, 10	239	9.79	0.00027	41
Precentral gyrus	Right	51, -10, 40	1476	14.69	0.00001	4
Angular gyrus	Right	46, -67, 31	668	11.74	0.00007	39
Inferior frontal gyrus	Right	42, 47, 1	924	12.34	0.00005	10
Insular	Right	39, -31, 13	178	11.94	0.00006	10
Superior frontal gyrus	Right	18, 26, 55	507	10.72	0.00014	6
Cingulate gyrus	Right	6, -43, 31	861	12.66	0.00004	31
Precuneus	Right	3, -61, 25	152	10.05	0.00023	31
Cingulate gyrus	Left	-3, -28, 34	359	10.61	0.00015	31
Caudate	Left	-12, -25, 22	245	11.49	0.00008	-
Declive	Left	-33, -70, -20	153	10.05	0.00023	-
Middle temporal gyrus	Left	-45, -61, 28	279	9.78	0.00027	39
Precentral gyrus	Left	-51, -13, 43	1569	14.93	0.00001	3

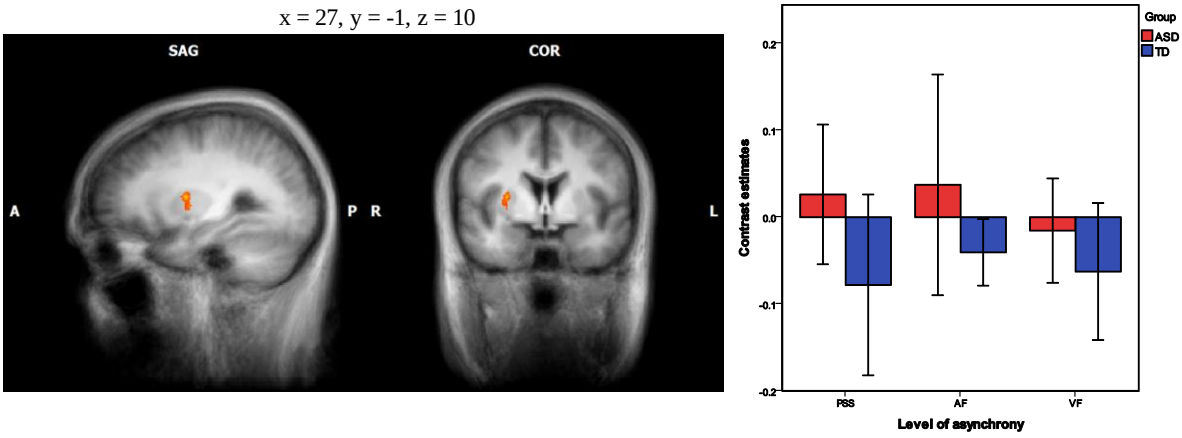


Figure 5. 3 Face-voice experiment fMRI activation data: Cluster of activation in the right putamen for which the difference between the brain responses to the three stimulus conditions (point of subjective simultaneity, audio first and visual first) varied across the two groups of participants (individuals with ASD and TD individuals). Sagittal and coronal slices show an activation focus at a location defined by x y z Talairach coordinates. The average contrast estimates (beta weights) and relative standard errors are shown in the histograms. The ASD group is depicted in red and TD group in blue at each stimulus condition: point of subjective simultaneity (PSS), audio first (AF) and visual first (VF).

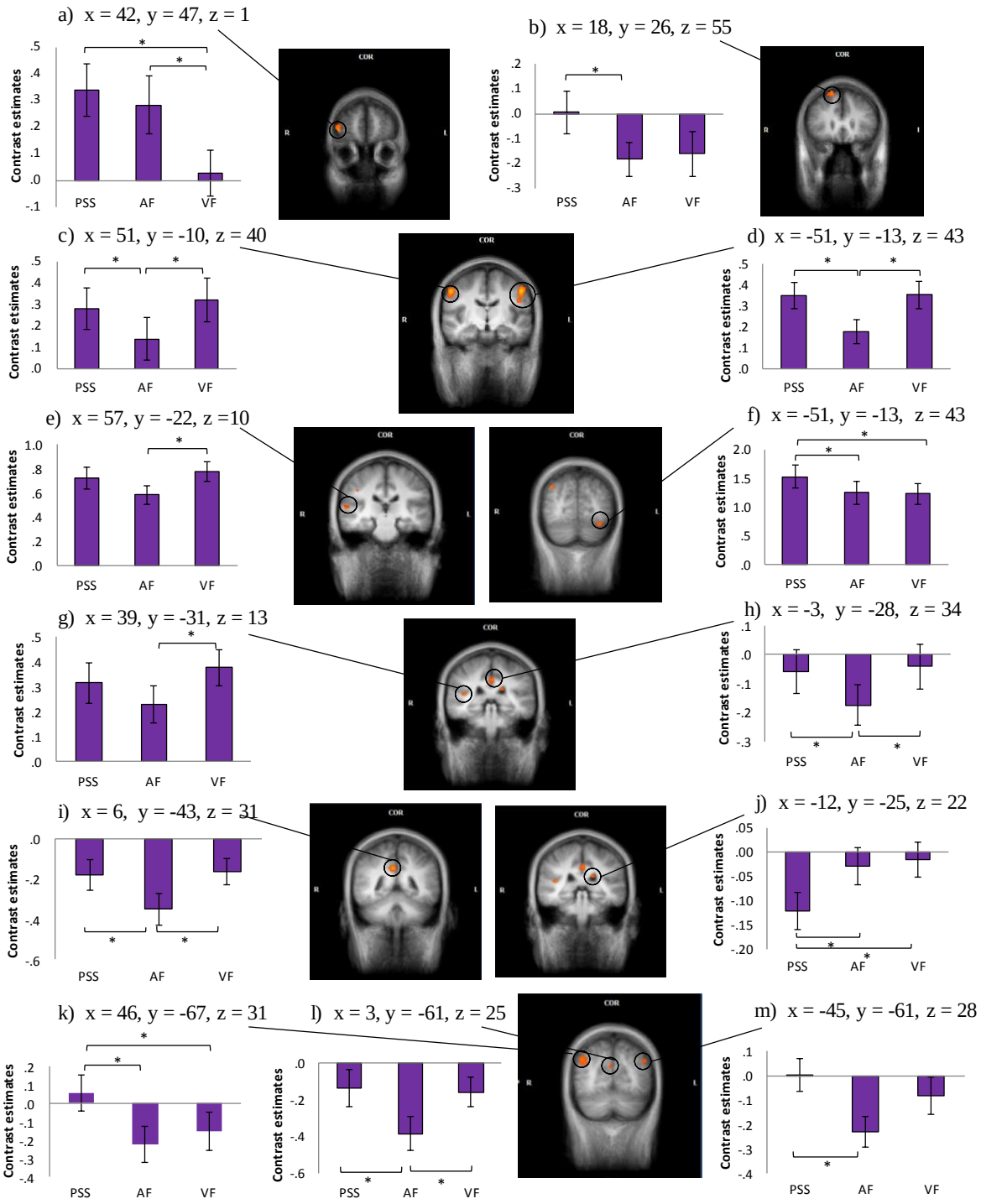


Figure 5. 4 Face-voice experiment fMRI activation data: Clusters of activation for which the brain responses, collapsed across groups (ASD, TD), differed across the three face-voice stimulus conditions: point of subjective simultaneity, audio first and visual first. Sagittal and coronal slices show activation foci at locations defined by x y z Talairach coordinates. The average contrast estimates (beta weights) and relative standard errors are shown in the histograms. The brackets and * indicate where the pairwise comparisons found significant differences ($p<0.05$) between the conditions. Point of subjective simultaneity (PSS), audio first (AF), visual first (VF). a) right inferior frontal gyrus, b) right superior frontal gyrus, c) right precentral, d) left precentral gyrus), e) right superior temporal gyrus, f) left declive, g) right insular, h) left cingulate gyrus, i) right cingulate gyrus, j) left caudate, k) right angular gyrus, l) right precuneus, m) middle temporal gyrus.

5. 4. 3 Audiovisual synchrony analyses restricted to Conjunction Regions

In Chapter 4 audiovisual integration regions were defined (by (AV> A) ∩ (AV> V)) for both experimental groups (ASD and TD) and for both stimulus types (BF and FV). These regions were created into masks to restrict the analyses investigating how activations to the three levels of asynchrony (PSS, AF, VF) differed between the experimental groups (ASD and TD). For BF, this restricted analysis revealed no significant main effect of group, level of asynchrony and no significant interaction. Similarly, for FV, no significant main effect or interaction was found, but the main effect of level of asynchrony revealed an area in the STG, the Heschl gyrus, which had previously been found by the whole-brain analysis.

5. 5 Discussion

In the present study, we investigated audiovisual temporal processing in ASD by asking ASD and TD participants to make SJs when presented with audiovisual BF and FV stimuli at different levels of asynchrony (the participants subjective point of simultaneity, audio first, video first). For BF displays, we did not find any difference in activation patterns between the two groups. However, we revealed that the levels of asynchrony modulated activation in the supramarginal gyrus (bilaterally), left medial frontal gyrus, left precentral gyrus and the right middle temporal gyrus. Similarly, for FV displays, the activation patterns were similar across the two groups, except for an area within the putamen, which revealed increased activation in the ASD group compared to the TD group. However, levels of asynchrony modulated activation of the bilateral precentral, gyrus, cingulate gyrus, right STG, angular gyrus, middle frontal gyrus, superior frontal, insular and left caudate, declive and middle temporal gyrus. Analyses restricted to only STG regions, which were shown to be sensitive to audiovisual integration in Chapter 4, showed no group differences across the levels of asynchrony.

While the activation patterns to SJ on BF displays were comparable across the ASD and TD participants, increased activation of the ASD group during SJ of FV displays was found in the right putamen. This is a very interesting result and can be linked to other literature. Von Saldern and Noppeney (2013) have demonstrated the importance of the putamen for integrating auditory and visual motion information when

performing motion discrimination tasks. Additionally, Watson et al. (2015) pointed out the putamen's heteromodal processing functions. In ASD, abnormal activity has been demonstrated in the striatum, of which the lateral parts correspond to the putamen, during social processing (Delmonte et al., 2012) and reward processing tasks (Delmonte et al., 2012). Moreover, DTI research reports decreased white matter connectivity between the striatum and prefrontal cortex in ASD, compared with TD individuals (Langen et al., 2012). The putamen has been found to have increased as well as decreased GM volume in ASD (Sato et al, 2014; Ecker et al., 2010). Sato et al. (2014) hypothesised that the increased volume of the putamen found in high-functioning adults with ASD might reflect structural or histological abnormalities of the putamen, and therefore could be the underlying cause of symptoms such as repetitive and stereotyped behaviours and impaired social interactions. Altogether, this data suggests that the abnormal structure and function of the putamen in ASD could be the underlying cause audiovisual integration deficits, as well as the diagnostic characteristics of ASD. Therefore, it would be of particular interest to investigate activation in the putamen relates to severity if diagnostic deficits in ASD.

Moreover, the putamen is known to be sensitive to temporal structure of sensory signals, especially auditory signals (Grahn & Rowe, 2009; 2013). Studies on experts have shown that people who are good at a task show less activation than novices in task specific brain regions (e.g., Petrini et al., 2011). Thus, these two results taken together suggest that extra activation of the putamen in ASD indicates that people with ASD are working harder to reflect the temporal structure of the displays. This explanation would be consistent with the results showing that individuals with ASD are less sensitive at detecting asynchrony between audio and visual information than TD individuals (Chapter 3).

Moreover, the lack of group difference of SJ on BF displays could suggest that audiovisual asynchrony detection of audiovisual displays might be more deficient in speech than non-speech. This is merely a speculation, as the two display types (BF and FV) are fundamentally different, and we cannot draw any solid conclusions from these findings. However, these findings would not be consistent with our previous behavioural SJ results, showing a marginally wider TIW in ASD across a range of different audiovisual displays, including BF and FV. Other behavioural results,

however, support this speculation, as they suggest speech-specific deficit in audiovisual temporal processing in ASD (Stevenson et al., 2014). A possible explanation of the differential activation patterns across display types is that the perceptual system is more tolerant to asynchrony as the complexity of the audiovisual displays increase (Love et al., 2013), and therefore SJs are harder to make when displays are more complex. Furthermore, the speech-specific activation differences in audiovisual SJ tasks could also underlie extended speech processing in ASD (Cardy et al., 2005).

A complex pattern of activation was revealed when studying the levels of asynchrony when participants made SJs. We propose that our findings shed light onto the asymmetry of the TIW. The asynchrony between the left and right sides of the TIW has commonly been observed in behavioural data (Miller & D'Esposito 2005; van Wassenhove, et al., 2007; Vroomen & Keetels, 2010; Stevenson & Wallace, 2013). The asymmetry is driven by the right side (containing conditions in which visual stimuli precede audio stimuli) being wider than the left side. Thus participants are much more likely to perceive visual first trials as synchronous when compared with audio first. This asymmetry has been argued to have ecological validity, since, in natural surroundings, visual stimulus energy from an audiovisual event will always reach the retina prior to auditory energy reaching the cochlea (Stevenson & Wallace, 2013). Therefore, the TIW might reflect the natural temporal statistics of stimuli within our environment. Additionally, there are considerable timing differences of the transduction processes and neural conduction of incoming audio and visual information (Lamb & Pugh, 1992; Lennie, 1981). As such, the PSS is also often where visual information is leading audio information in SJ tasks (Roach, Heron, Whitaker, & McGraw, 2011).

While a few fMRI studies have investigated activation differences between synchronous and asynchronous audiovisual displays (Stevenson et al, 2010; Lewis & Noppeney, 2010; Love, 2011), to our knowledge, none have compared the activation patterns elicited by perceivable AV displays compared to perceived VA. Interestingly, for BF, none of the audiovisual temporal asynchrony sensitive areas (middle frontal gyrus, precentral gyrus, putamen, bilateral supramarginal gyri) were found to elicit more activation during the PSS condition than the asynchronous conditions. This is a

contradictory finding to the results by Lewis and Noppeney (2010), which reported greater activation to their physical synchronous condition, compared to their asynchronous conditions of non-speech displays. A possible explanation for the discrepancies in these findings are the different neural networks involved in processing physical synchronous stimuli, compared to perceptually defined synchronous stimuli (Love, 2011). Moreover, for BF, the putamen responded most to the AF condition, compared to PSS and VF, a further finding that might underlie the asynchronous TIW.

In our FV results, we find an even more complex pattern of activations across the different levels of asynchrony. Firstly, in a majority of regions, the PSS condition elicited a stronger response than AF (left middle temporal gyrus, right superior frontal gyrus, precentral gyri, the cingulate gyri & right precuneus), VF (right inferior frontal gyrus) or both conditions (left declive & right angular gyrus). Only the left caudate revealed decreased activation to PSS compared to AF and VF conditions. Furthermore, differential activation between AF and VF conditions revealed that only the right inferior frontal gyrus responded more to AF than VF conditions, while the precentral gyri, cingulate gyri, right precuneus, Heschl gyrus in the right STG and right insula responded more to VF than to AF conditions. These findings might be the underlying reason of the observed asynchrony of the TIW. Moreover, the findings of higher activation to VF than to AF conditions is in agreement with study by Perrodin, Kayser, Logothetis and Petkov (2015), which recently showed that natural asynchronous (visual leading) dynamic face-voice stimuli regulate network oscillations and neuronal excitability in the voice-sensitive cortex of macaques, located in the anterior part of the temporal gyrus. Although Love (2011) looked at a range of different asynchronous displays and made different contrasts, some of our results are in agreement with his. Moreover, he defined the posterior right STC as a neural correlate. This reflects the common finding that people are better detecting asynchrony in audio-leading stimuli, and is a result that our findings could be in agreement with. However, further investigation is needed to fully explore the neural correlates of the observed asymmetry in AF and VF conditions.

Conclusion

This study measured BOLD signals while ASD and TD participants made SJ on audiovisual displays of different levels of asynchrony: PSS, AF and VF. While the activations related to SJ on BF displays were comparable across the ASD and TD participants, SJ on FV displays revealed increased putamen activation in ASD participants. This finding supports research suggesting that audiovisual asynchrony detection of audiovisual displays might be less effective in speech than non-speech. However, since the display types (BF and FV) are fundamentally different, this result remains somewhat speculative. Furthermore, the increased putamen activation in ASD for SJ on FV displays is potentially significant given its fundamental role in sensitivity to temporal structure of sensory signals (Grahn & Rowe, 2009; 2013). Studies on experts show that people who are good at a task show less activation than novices (e.g., Petrini et al., 2011). We therefore suggest that extra activation of the putamen reflects that people with ASD try harder to reflect the temporal structure of the displays. This explanation is also consistent with the wider TIW found in ASD shown in Chapter 3. Additionally, the investigation of different levels of asynchrony revealed a complex pattern of results, indicating a network of areas more involved in processing PSS than AF and FV, as well as areas responding differently to VF compared to AF. These activation differences between audio first and video first in different brain areas are in agreement with the view that AF and VF are processed differently.

6. General Discussion

Conclusions from experimental results

There are four key conclusions to be drawn from this Ph.D. thesis:

1. Autistic traits correlate with cortical thickness (CT) measurements in areas associated with atypical activation in autism spectrum disorder (ASD) (Chapter 2).
2. Compared to typically developed (TD) individuals, individuals with ASD have a marginally wider temporal integration window (TIW) when making synchrony judgements (SJs), while the TIW estimated from temporal order judgements (TOJs) is of similar width in both groups. A model-based approach indicates that this widening of the TIW in SJs is due to decreased temporal resolution at a decisional level in ASD (Chapter 3).
3. Audiovisual, audio and visual processing of simple beep-flash (BF) and more complex face-voice (FV) displays mainly revealed a reduction of activation in brain areas in the ASD group compared to the TD group (Chapter 4).
4. SJ of audiovisual FV displays are underpinned by greater activation in the putamen in the ASD group compared to the TD group (Chapter 5).
5. SJ of audiovisual BF and FV displays reveal a complex pattern of activations providing evidence for a potential neural basis of the commonly reported asymmetry of the TIW, which reflects the enhanced ability to detect asynchrony during audio leading asynchronous displays (Chapter 5).

The results provide new insights into autistic traits and ASD and the underlying behavioural, as well as functional and structural brain abnormalities. While Chapter 2 showed how the structure of CT is associated with autistic traits in the general population, Chapters 3, 4 and 5 helped further our understanding of the basis of audiovisual integration differences in ASD, utilising psychophysical as well as functional magnetic resonance imaging (fMRI) methods.

6.1 Conclusion 1 from Chapter 2

The objective of Chapter 2 was to investigate the relationship between cortical thickness (CT) and autistic traits as measured by the Autism Spectrum Quotient questionnaire (AQ). Traits of ASD, such as social communication and interaction

deficits, as well as repetitive patterns of behaviour, interests and activities, are thought to be present in the typical population, and the AQ was developed to assess the prevalence of these autistic traits in the general population. Von dem Hagen et al (2011) showed that AQ is associated with white matter (WM) and grey matter (GM) volume using voxel-based-morphometry (VBM), but found no GM differences in areas associated with social cognition. However, research shows that VBM potentially conflates information about morphology, size and position (Ashburner & Friston, 2001), while CTA is less susceptible to positional variance and provides a more direct index of cortical morphology (Kim et al., 2005; Jiao et al., 2010). Therefore, the current study made use of the semi-automatic, surface-based CTA tools in Brainvoyager to further investigate the relationship between CT and AQ in the same sample previously investigated by von dem Hagen et al. (2011).

A whole-brain analysis was employed, which revealed positive correlations between CT and AQ in the left temporo-occipital junction, left posterior cingulate, right precentral gyrus and bilateral precentral sulcus, areas previously associated with structural and functional differences in ASD. Our findings were supported by previous research showing that these areas are often associated with functions impaired in ASD, such as social processing, attention switching or motor skills. Additionally, these areas have been related to structural and functional activation abnormalities in ASD. Our findings provide further evidence that the autistic traits (measured by the AQ) and CT are correlated in the general population.

Moreover, the discrepancy between our results and those by von dem Hagen et al. (2011) provides further evidence that results of CT measures and GM volume measures are not necessarily comparable. This is a commonly found observation, and research investigated the cause for the heterogeneity of cortical morphology estimates. Hazlett et al. (2011) examined GM volume, CT and surface area (SA) in ASD and suggested that increased GM volume might be associated with increased SA rather than CT. Moreover, Raznahan et al. (2010), in a cross-sectional study in ASD, reported altered neurodevelopmental trajectories for GM volume and CT, but not SA. These results were supported in a recent study by Ecker et al. (2013) which investigated GM volume, SA, and CT, as well as their relationship in a large sample of men with ASD and well matched typically developed controls. These results suggest that GM volume measurements are derived from measurements of SA and CT, which

are measurements associated with different developmental pathways. These pathways are likely to be controlled by different underlying neurobiological mechanisms. Therefore, CT is a more direct measure of cortical morphology than GM volume. Our results also suggest that CT measurements might be more sensitive to differences in cortical morphology than GM volume measurements. Similar conclusions have been drawn from studies measuring CT as well as GM volume in the same population (Hyde et al., 2007).

6.2 Conclusion 2 from Chapter 3

The main aim of Chapter 3 was to examine audiovisual temporal integration in ASD using different stimulus types, tasks and data fitting methods. The ability to integrate auditory and visual information is crucial to everyday life, but results in the literature are mixed regarding how individuals with ASD integrate audiovisual information. To investigate this question, we examined the TIW, which indicates how precisely sight and sound need to be temporally aligned so that a unitary audiovisual event can be perceived. A total of 26 adult males with ASD and 26 age-and IQ-matched TD males were presented with BF, point-light drummer (PLD), and FV displays at 11 values of stimulus onset asynchrony (SOA), as well as synchrony, while participants were making SJs and TOJs.

Analysis of the data included fitting Gaussian functions as well as fitting an Independent Channels Model (ICM) (Garcia-Perez & Alcala-Quintana, 2012; Garcia-Perez & Alcala-Quintana, 2013). The ICM was used to fit the response data from SJ and TOJ in a more flexible manner than the Gaussian function fits, allowing for individual asymmetries and irregularities in the data. Gaussian curve fitting for SJs showed that the ASD group had a wider TIW, but no group effect was found for TOJ. Possible differences in cognitive processes required for SJs and TOJs can help us understand the underlying processes of why temporal audiovisual integration differs in ASD. The finding that the ASD group had a wider TIW in SJs, but not TOJs, suggests that this difference is due to difficulties in combining the audio and the visual cues. SJs require one to estimate the temporal correspondence of the audio and visual information, and thus depend on more global levels of processing (i.e., considering the stimulation as a whole), whereas TOJs could, in principle, be performed by focusing on only one sensory cue to detect whether it came first or not, thus depending on more

local level processing (i.e., considering only the sound). Therefore, audiovisual integration difficulties in ASD are likely to be due to difficulties in processing global information. This is in line with the hypotheses of a central coherence deficit or temporal binding deficit in ASD (Brock et al., 2002).

The ICM supported these results and model parameters indicated that the wider TIW for SJs in the ASD group was not due to unisensory processing, but rather due to decreased temporal resolution at a decisional level of combining the sensory information. The results of the wider TIW for SJ is largely in agreement with Stevenson et al. (2014), who showed a wider TIW for SJ in ASD, but no unisensory processing differences in ASD. While our results showed a wider TIW across all stimulus types, and thus suggest generalised deficit, Stevenson et al. (2014) did not find wider TIWs for non-speech displays, suggesting a speech-specific audiovisual temporal processing deficit. However, other research supports our findings by showing audiovisual temporal integration differences in ASD when presented with simple beep and flash stimuli (e.g., Foss-Feig et al., 2010; Kwakye et al., 201).

Furthermore, when modelling TOJ, the ICM revealed a smaller Point of Subjective Simultaneity (PSS; closer to physical synchrony) in the ASD group than in the TD group. This result is in disagreement with the findings by de Boer-Schellekens et al., (2013), which revealed no PSS differences between the ASD and TD group. This discrepancy between the findings could be due to the different fitting methods being employed to estimate PSS. Their null finding could also be explained by the PSS being highly variable across participants and due to their small sample size.

These results are encouraging for potential interventions to improve sensory processing in ASD, especially because it has been shown that the TIW width becomes smaller through training (Powers et al., 2009; Stevenson et al., 2013). It would also be of importance to further explore the link between sensitivity to audiovisual asynchrony and speech perception and comprehension, as well as looking at how training on multisensory TIW width would translate into everyday multisensory speech processing and comprehension. Moreover, our behavioural results motivate the use of fMRI to aid us in understanding the underlying differences in audiovisual integration in ASD. Thus far there has been little research investigating the neural underpinnings of the differences in audiovisual integration in ASD.

6.3 Conclusion 3 from Chapter 4

Our behavioural results in Chapter 3 revealed that individuals with ASD have deficits in audiovisual integration. Subsequently, our aim of Chapter 4 was to investigate whether these audiovisual integration deficits in ASD would be reflected in neural activation patterns. The existence of such differences have recently been supported by electroencephalography (EEG) studies (Brandwein et al., 2015). Using fMRI we investigated audiovisual, auditory and visual processing in ASD of simple BF displays and complex, social FV displays. During a block design experiment, we measured the BOLD signal while 13 adults with ASD and 13 typically developed (TD) age-, sex- and IQ- matched adults were presented with audiovisual, audio and visual information of BF and FV displays.

The results revealed that audiovisual, unisensory auditory and visual processing of both social FV and simple BF stimuli are mostly associated with reduced activation in ASD. Audiovisual, auditory and visual conditions of human FV stimuli revealed reduced activation in ASD participants, compared to TD participants in regions of the frontal cortex. This finding is generally constant with results by Doyle-Thomas et al. (2013), while BF stimuli revealed reduced levels of activation in the lingual gyri. An interaction between group and sensory modality condition of BF stimuli revealed that the activation of the inferior parietal gyrus was differentially modulated by the different sensory modalities in ASD and TD participants. In detail, we found increased activation in audiovisual and auditory conditions compared to the visual condition in individuals with ASD, while TD controls showed increased activation in audiovisual and visual conditions compared to the auditory condition. Taking the results of FV and BF together, we show that not only cortical processing of socially relevant audiovisual, auditory and visual information is abnormal in ASD, but that sensory processing defects could be more generalised, including simple non-social information. This finding is in agreement with our behavioural findings in Chapter 3.

Furthermore, the conjunction analyses testing for regions sensitive to audiovisual integration discovered the STC in both groups and for both display types. The importance of the STC in audiovisual integration is supported by other studies in the literature (e.g., Stevenson & James, 2009; Watson et al., 2014). Interestingly, the activated regions in ASD were smaller than those in TD individuals, which is

consistent with the literature showing reduced activation in STG during audiovisual emotional FV matching tasks (Doyle-Thomas et al., 2013), as well as structural abnormalities (Ecker et al., 2010; Hyde et al, 2010). However, against our predictions, we did not find any activation differences, per se, of the STC between the two groups. Instead, a superior frontal area was shown to be sensitive to audiovisual FV stimuli in the TD group, but not in the ASD group. This finding is in agreement with previous studies looking at audiovisual emotional matching tasks (Doyle-Thomas et al., 2013; Loveland et al., 2008). Overall, this study indicated that audiovisual, auditory and visual processing of social and non-social stimuli led to different activation patterns in individuals with ASD compared to TD individuals. These results are mostly in support of the recent EEG findings showing neural markers of auditory processing and multisensory integration to be correlated with severity of autistic symptoms (Brandwein et al., 2015). However, correlation was found between the neural markers of auditory and multisensory processing against clinical measures of visual and auditory sensitivities. Overall, this data supports the idea that abnormal multisensory and unisensory processing contributes to autism symptoms. In future experiments it would be interesting to investigate the relationship between our activation levels to audiovisual, audio and visual stimulation and the severity of autistic symptoms. Furthermore, it would be a good idea to control for clinical measures of auditory and visual sensitivities, or investigate whether our activation results could be related to sensory sensitives.

Moreover, compensatory processing mechanisms in adults with ASD have been previously found. For example, McKay et al., 2013 revealed that adults with ASD used different brain networks when given biological motion tasks, while no behavioural differences were observed. Thus, finding abnormal activation patterns to audiovisual, audio, visual stimuli in this chapter, but only finding deficits in audiovisual integration while auditory and visual processing was intact in our behavioural chapter (Chapter 3), suggests the possibility that our adults with ASD have developed compensatory strategies for audio and visual processing that are only revealed by abnormal levels of activation in ASD. Therefore, it is important to investigate neural substrates of audio and visual perception in ASD further.

6.4 Conclusion 4 and 5 from Chapter 5

The aim of this study was to further investigate audiovisual temporal processing in ASD. This was based on the behavioural findings in Chapter 3 showing that individuals with ASD are less sensitive to audiovisual asynchronies when making SJ. Using functional magnetic resonance imaging (fMRI), we investigated audiovisual temporal processing in ASD. In 13 adult males with ASD and 13 age-, sex-, and IQ-matched typically developed (TD) controls, we investigated temporal asynchrony of audio and visual information in simple BF displays, as well as complex and social FV displays. The study measured BOLD signals while the ASD and TD participants made SJ on the aforementioned audiovisual displays of different levels of asynchrony: the participants' PSS, audio leading visual information (audio first), visual leading audio information (visual first).

While no activation differences between the groups were found in SJ on BF displays, SJ on FV displays revealed increased putamen activation in ASD participants compared to TD participants. This finding supports research suggesting that deficits of audiovisual asynchrony detection of non- speech stimuli might be less affected in ASD than for speech stimuli (Stevenson et al., 2014). However, this interpretation is not in line with our results from Chapters 3 and 4. In these Chapters we report atypical performance and activation levels in individuals with ASD across non-social and social conditions, which provide evidence of a more generalised sensory processing deficit. Importantly, it needs to be mentioned that our display types (BF and FV) not only differ in their social content, but are fundamentally different in their visual, auditory and temporal characteristics, which are all aspects that could influence the results. Therefore, this result remains somewhat speculative, and should be further researched using highly controlled stimuli.

Furthermore, the increased activation levels of the ASD group in the putamen during SJs of FV displays is of potential significance given its fundamental role in sensory processing and the detection of temporal beat structure (Grahn & Rowe, 2009; 2013). Studies on experts have shown that people who have extensive experience with a task show less activation than novices (Petrini, et al., 2011). Taken together these findings suggest that extra activation of the putamen reflects that individuals with ASD recruit more resources to determine the temporal structure of the displays. This explanation is

also consistent with the wider TIW found in ASD (Chapter 3). Therefore, we conjecture that the higher activation found with ASD in the putamen is a reflection of the audiovisual temporal integration deficit observed in SJ.

The second aim of this experiment was to investigate the neural correlates of the often observed asymmetry of the TIW, which shows that people are better at detecting audiovisual asynchrony in audio first conditions (e.g., van Wassenhove et al., 2007; Stevenson & Wallace, 2013). Therefore, we investigated activation patterns that were elicited in response to the different levels of asynchrony (PSS, audio first, video first). We revealed a complex pattern of results indicating a network of areas more involved in the perception of PSS than audio-first and visual-first displays, as well as areas responding differently to audio-first compared to video-first. The activation differences between audio-first and video-first conditions are in agreement with the view that audio-first and visual-first are processed differently. This is new fMRI evidence for a potential neural basis of the well-defined behavioural result of the asymmetry of the TIW. We did not measure the asymmetry of our behavioural data in Chapter 3, and therefore cannot conclude on these results. However, this could be further investigated by measuring the activation responses while making SJ across a wider range of SOAs, similar to the study by Love (2011).

6.5 Linking results of all chapters together

The results provide new insights into autistic traits and ASD, and the underlying behavioural as well as functional and structural brain abnormalities. Chapter 2 showed how the structure of CT is associated with autistic traits in the general population. Chapters 3, 4 and 5 helped us further our understanding of the basis of audiovisual integration differences in ASD, utilising psychophysical as well as functional magnetic resonance imaging (fMRI) methods. Chapter 3 showed that the TIW is wider in ASD when making SJ, but not when making TOJ, and that this widening was due to a decreased temporal resolution at a decisional level of combining the sensory information, and not due to deficits in sensory processing. Overall, Chapter 4 showed evidence that audiovisual, auditory and visual processing of social and non-social stimuli led to different activation patterns in individuals with ASD, compared to TD individuals. The finding that both social and non-social displays led to different activation patterns in ASD was in agreement with our behavioural findings in Chapter

3, which showed that individuals with ASD were less sensitive to asynchrony across all stimulus types (including social and non-social). However, audiovisual, visual and audio processing elicited different activations in ASD, which differed from the results of the Chapter 3, in which the ICM predicted no unisensory processing differences in ASD. Chapter 5 looked more specifically at the neural correlates of SJs in ASD and TD participants. While making audiovisual SJs on FV stimuli increased activation levels in the putamen were found in individuals with ASD compared to the TD controls.. Since novice versus expert studies indicate that increased activation can potentially reflect inefficient performance (Petrini et al., 2011), these findings can be taken in agreement with the results of Chapters 3 and 4. Chapter 2 provided evidence of anatomical differences associated with the autistic traits and motivated a closer consideration of anatomical brain differences. However, Chapters 3, 4 and 5 were not linked to these anatomical abnormalities because the areas in which we found CT to correlate with autistic traits did not correspond to areas associated with abnormal activation patterns in ASD elicited by our sensory stimulation.

6.*6 Limitations of the research*

It must be acknowledged that there are some limitations to the conclusions of the current thesis. One limitation of the research presented in this Ph.D. thesis is that we were not able to confirm the ASD diagnoses of our participants, other than obtaining their AQ scores. The gold-standard method used is a combination of the Autism Diagnostic Interview- Revised (ADI-R) and the Autism Diagnostic Observation Schedule (ADOS) (see Simmons et al., 2009, for a brief description). However, these diagnostic techniques for adults are currently being debated (Matson & Neal, 2009). Performing these diagnostic assessments would have greatly increased the recruitment difficulties, as well as putting constraints on other resources, such as limited funding available. However, participants have provided us with diagnostic information. Moreover, the ASD group's mean AQ score (M= 36.64, SD = 8.80) was well above the cut off score (26/50) for Asperger's, and therefore supporting the diagnoses of individuals in the ASD. Furthermore, parents or life partners were asked to complete a demographics questionnaire indicating the diagnoses received by their child or partner.

Secondly, we were unable to collect IQ scores for the participants in Chapter 2. Interestingly, Hoekstra et al. (2010) has previous discovered a modest negative correlation between autistic traits and IQ (r= –.27), however this was mainly driven by communication problems and it was emphasised that autistic traits and IQ are mostly unrelated. Furthermore, IQ has been found to correlate with CT (Narr et al., 2007; Choi et al., 2008). However, the participants were recruited through the University of Cambridge subject pool, and thus were likely to mainly be students with average or above average IQ. Furthermore, our CT did not control for age and gender, which again are characteristicsthat are linked to CT differences (Zielinski et al., 2014; Sowell et al. (2007). Moreover, the sex differences are also found in mean AQ (Baron-Cohen et al., 2014; Ruzich et al., 2015). Therefore, future experiments studying the association of AQ scores and structural differences in the general population CT should control for age, IQ and sex differences in order to better isolate this specific relationship between AQ and CT. Heterogeneity in ASD samples in general is an important discussion point and is likely to be the underlying factor of frequently reported inconstant results in ASD research. Therefore, Chapters 3 and 4 only investigated homogenous groups with ASD and TD participants being matched on age, IQ and gender. In fact, we only included men in our sample.

6.7 Is AQ, the best measure of autistic traits?

The total AQ has been shown to have good test-retest reliability as well as good internal consistency (Baron-Cohen et al., 2001). Moreover, the AQ has been reported to have suitably high sensitivity and specificity in individuals referred for diagnosis (Woodbury-Smith, Robinson, Wheelwright, Baron-Cohen, 2005). Therefore, the AQ is a sensitive measure of autistic traits in the general population. Consistent with genetic evidence showing that ASD is heritable, AQ scores have been shown to be heritable within families (Hoekstra, Bartels, Verweij & Boomsma, 2007). Additionally, the AQ scores predict performance on tasks commonly associated with superior performance in individuals with ASD (Almeida et al., 2010), while also predicting performance on tasks that are impaired in ASD (Baron-Cohen et al., 2001; Wyer, et al., 2012). However, it needs to be mentioned that the AQ is not the only measure of autistic traits. For example, the Broad Autism Phenotype Questionnaire (BAPQ) was developed by Hurley, Losh, Parlier, Reznick and Piven (2007), while the adult Social Responsiveness Scale (SRS) was originally developed by Constantino and Todd

(2005). A study by Brooke, Hopwood, Wainer and Donnellan (2011) compared these three self-report measures of autistic traits and showed that the BAPQ and SRS clearly demonstrated sex differences and had better internal consistency than the AQ. Moreover, in this study, the BABQ was the only measure to show normal distributions of its total score as well as sub-scores. Generally the SRS and BAPQ were shown to have better criterion variability.

Furthermore, Gregory and Plaisted-Grant (2013) investigated whether the similarity in performance by high-AQ individuals and people diagnosed with ASD actually reflects the same underlying perceptual processes. The authors administered two visual search tasks to a large sample of TD individuals, as well as assessed individuals using the AQ. The results suggested that using AQ scores as a substitution for ASD requires unverified assumptions about high-AQ scoring individuals and their relationship to individuals with an ASD. Furthermore, research has not fully explained the endophenotypes related to ASD, and thus the AQ can only function as an approximation of these. However, when no individuals with ASD are available, the AQ enables researchers to study healthy individuals who have been scored for AQ instead. However, it is important to bear in mind that this might come at a scientific cost.

6.8 Implications of findings of audiovisual processing differences

The multisensory integration differences that we reveal in Chapters 3, 4 and 5 could have cascading effects in the early development of social communication skills. For example, early language learning in TD children involves integrating incoming audio (speech sounds) and visual (lip movements) information (Teinonen et al., 2008). The benefit that people typically get from such multisensory inputs during speech perception has been shown to be considerably impaired in children with ASD (Foxe et al. 2013; Stevenson et al. 2014). Similarly, social communication could also be impacted by impaired multisensory integration through the misinterpretation of non-linguistic social cues such as facial expressions, and changes in prosody of the speech signals are needed to interpret a speaker's emotion and intention (Ethofer et al., 2006). Furthermore, our behavioural results in Chapter 3, looking at the audiovisual temporal processing, could suggest that individuals with ASD rely more on integrating redundant sensory information, as their integration system seems less precise at

detecting asynchrony between incoming audio and visual information. This could lead to falsely integrating information together that does not belong together. In social situations this can lead to misinterpretation of social cues. Additionally, Brandwein et al. (2015) suggested that these deficits of precisely integrating audiovisual information could underlie existing deficits observed in ASD, such as the feeling of 'sensory overload'. This is in agreement with Molholm et al. (2004), who suggested that the integration of multisensory information is crucial to group together the information that comes in through the separate sensory systems. Moreover, the feeling of 'sensory overload' can lead to withdrawal and defensive behaviours (Brandwein et al., 2015). Similarly, Donnellan, Hill & Leary (2013) proposed that acts of apparent non-compliance, reluctance, lack of interest as well as aggression might not be voluntary, and could be secondary to an individual's particular sensory processing differences.

6.9 Future experiments

I believe the work described in this thesis provides a springboard for furtherresearch. For example, it would be interesting to run similar experiments as those in Chapters 3, 4 and 5, and measure ASD symptom severity across participants, correlating symptom severity with measures of TIW and PSS. To my current knowledge, there are no published studies looking at the relationship between the severity of symptoms in ASD or autistic traits in the general population, and performance on SJ and TOJ tasks. Moreover, it would be interesting to regress symptom severity against functional activations elicited through multisensory and unisensory information, as well as through making SJ on multisensory displays. This would give us a better understanding of how multisensory deficits in ASD are related to actual symptoms in ASD.

As mentioned previously, there has been very little fMRI research investigating audiovisual integrating in ASD. Our results in Chapter 4 suggest that atypical activation levels in response to audiovisual and unisensory stimulation are not unique to social speech stimuli, but are also present in BF displays. Conversely, Chapter 5 revealed only activation differences in our speech displays. This result might mislead people into thinking that this is evidence for a speech specific audiovisual integration deficit. However, as mentioned before, these displays had some fundamental differences that could potentially be the cause for not finding group differences in our BF condition. Therefore, it would be interesting to investigate the specificity of the

audiovisual integration deficit in ASD using highly controlled stimuli, such as ensuring that displays have the same luminance, display size, similar complexity, similar length and similar sound features. Furthermore, our behaviour results from the ICM revealed that the TOJ resulted in PSS differences between the ASD and the TD group. Therefore it would be interesting to explore this difference further and investigate whether neural correlates reflect those differences in PSS.

Wallace and Stevenson (2014) propose that using approaches from perceptual plasticity (Powers et al., 2009, Powers et al., 2012, Stevenson et al., 2013) to provide training in multisensory perception could be utilised to improve unisensory and multisensory temporal acuity. Successful training can narrow the width of individuals TIW (Powers et al., 2009; Stevenson et al., 2013). Similarly, Petrini et al., (2011) showed that the TIW is narrower in people with musical expertise (such as professional drummers), and that this performance difference is clearly reflected in activation levels of the brain of audiovisual temporal perception. Similarly, simple training on audiovisual temporal integration tasks, like the SJ, have been shown to be translated into the neural correlates of audiovisual temporal processing (Powers et al., 2012). The most promising result of these studies is that individuals with the widest TIW are the ones that benefit the most from training (i.e., showed the most significant changes of TIW width) (Powers et al., 2009; Stevenson et al., 2013). Therefore, it is likely that people with ASD, or other populations with audiovisual temporal processing difficulties such as dyslexia and Schizophrenia, would benefit from training. Moreover, the implications of such training could be researched to see whether the training translates into more general changes in multisensory integration, beyond the task that they are trained in. In the near future, it would be fascinating to see how long these training effects last. Furthermore, research would need to be extended to show whether such training would lead to improvements of real life functions, such as social skills, communication as well as hypo and hyper-sensory processing.

6.10 General Conclusion

Overall, this thesis aids our understanding of how individuals with ASD process audiovisual information, as well as how cortical structure is related to autistic traits in the general population. To achieve this understanding, structural and functional

magnetic resonance imaging (fMRI), as well as psychophysical techniques, were employed. Our results showed evidence of cortical thickness differences associated with the autistic traits.

We showed that individuals with ASD are less sensitive at detecting asynchronies between sight and sound when making synchrony judgements. Further fMRI analyses revealed that audiovisual, audio and visual processing of simple non-social and social displays elicit mainly a reduction of activation in brain areas in individuals with ASD compared to the TD individuals. Moreover, synchrony judgements of audiovisual social displays were underpinned by greater activation in the putamen in individuals with ASD compared to TD individuals. Lastly, we found that synchrony judgements of audiovisual displays revealed a complex pattern of activations, providing evidence for a potential neural basis of the commonly reported asymmetry of the temporal integration window.

References

Adhikari, B. M., Goshorn, E. S., Lamichhane, B., & Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain connectivity*, *3*(5), 536-545.

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior research methods*, *45*(4), 972-998.

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition (text revision).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Ameis, S. H., Fan, J., Rockel, C., Voineskos, A.N., Lobaugh, N. J., Soorya, L., Wang, A.T., Hollander, E., Anagnostou, E., 2011. Impaired structural connectivity of socioemotional circuits in autism spectrum disorders: a diffusion tensor imaging study. *PLoS One*,*6*, e28044.

Anagnostou, E., & Taylor, M. J. (2011). Review of neuroimaging in autism spectrum disorders: what have we learned and where we go from here. *Mol Autism*, *2*(4).

Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of vision*, *6*(3), 6.

Ashburner, J. & Friston, K. J. (2000). Voxel-based morphometry-the methods. *Neuroimage, 11(6),* 805-821.

Ashburner, J. & Friston, K. J. (2001). Why voxel-based morphometry should be used. *Neuroimage, 14,* 1238-1243.

Assaf, Y., Pasternak, O., 2008. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. *Journal of Molecular Neuroscience 34*, 51–61.

Axial Diffusivity in the ILF and IFOF is Related to Autism Symptom Severity. *Journal of vision*, *15*(12), 639-639.

Aylward, E. H., Minshew, N. J., Field, K., Sparks, B. F. & Singh, N. (2002). Effects of age on brain volume and head circumference in autism. *Neurology, 59,* 175-183.

Aylward, E. H., Minshew, N. J., Field, K., Sparks, B. F., & Singh, N. (2002). Effects of age on brain volume and head circumference in autism. *Neurology*,*59*(2), 175-183.

Bargary, G., & Mitchell, K. J. (2008). Synaesthesia and cortical connectivity.*Trends in neurosciences*, *31*(7), 335-342.

Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., & Reiss, A. L. (2004). White matter structure in autism: preliminary evidence from diffusion tensor imaging. *Biological psychiatry*, *55*(3), 323-326.

Barnea-Goraly, N., Lotspeich, L.J., Reiss, A.L., 2010. Similar white matter aberrations in children with autism and their unaffected siblings: a diffusion tensor imaging study using tract-based spatial statistics. Archives of General Psychiatry 67, 1052–1060.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes?. *British Journal of Developmental Psychology*, *13*(4), 379-398.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*(1), 5-17.

Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The 'when'pathway of the right parietal lobe. *Trends in cognitive sciences*, *11*(5), 204-210.

Battelli, L., Pascual-Leone, A., Cavanagh, P. (2007). The 'when' pathway of the right parietal lobe. *Trends in Cognitive Science, 11,* 204–210.

Beauchamp, M. S. (2005). Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics*, *3*(2), 93-113.

Beauchamp, M. S., Yasar, N. E., Frye, R. E., & Ro, T. (2008). Touch, sound and vision in human superior temporal sulcus. *Neuroimage*, *41*(3), 1011-1020.

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.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309-312.

Bhat A. N., Galloway J. C., Landa R. J. (2012). Relation between early motor delay and later communication delay in infants at risk for autism. Infant Behav. Dev. 35, 838–846.

Billeci, L., Calderoni, S., Tosetti, M., Catani, M., & Muratori, F., (2012). White matter connectivity in children with autism spectrum disorders: a tract-based spatial statistics study. *BMC Neurology 12*, 148.

Binder, M. (2015). Neural correlates of audiovisual temporal processing–Comparison of temporal order and simultaneity judgments. *Neuroscience, 300,* 432-447.

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology*, *19*(6), 503-508.

Boenke, L. T., Deliano, M., & Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental brain research*, *198*(2-3), 233-244.

Brandwein A. B., Foxe J. J., Butler J. S., Frey H. P., Bates J. C., Shulman L. H., et al. (2015). Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *Journal of Autism Developmental Disorder, 45,* 230–244.

Brandwein A. B., Foxe J. J., Butler J. S., Russo N. N., Altschuler T. S., Gomes H., et al. (2013). The development of multisensory integration in high-functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cerebral Cortex, 23,* 1329–1341.

Brandwein A. B., Foxe J. J., Butler J. S., Russo N. N., Altschuler T. S., Gomes H., et al. (2013).The development of multisensory integration in high-functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. Cerebral Cortex, 23, 1329–1341.

Brandwein, A. B., Foxe, J. J., Butler, J. S., Frey, H. P., Bates, J. C., Shulman, L. H., & Molholm, S. (2015). Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *Journal of autism and developmental disorders*, *45*(1), 230-244.

Brang, D., Williams, L. E., & Ramachandran, V. S. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex*, *48*(5), 630-637.

Brang, D., Williams, L. E., & Ramachandran, V. S. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex*, *48*(5), 630-637.

Brian J., Bryson S. E., Garon N., Roberts W., Smith I. M., Szatmari P., et al. (2008). Clinical assessment of autism in high-risk 18-month-olds. Autism 12, 433–456

Brock, J., Brown, C., Boucher, J., & Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Developmental Psychopathology, 14,* 209-224.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences, 1124*(1), 1-38.

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

Callan, D. E., Jones, J. A., Munhall, K. G., Kroos, C., Callan, A. M., & Vatikiotis-Bateson, E. (2002). Mirror neuron system activity and audiovisual speech perception. In *Proceeding of the 8th International Conference on Functional Mapping of the Human Brain on CD-Rom in NeuroImage* (Vol. 16, No. 2).

Callan, D., Callan, A., & Jones, J. A. (2014). Speech motor brain regions are differentially recruited during perception of native and foreign-accented phonemes for first and second language listeners. *Frontiers in neuroscience*, *8*, 275.

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-

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, *14*(2), 427-438.

Capek, C. M., Bavelier, D., Corina, D., Newman, A. J., Jezzard, P., & Neville, H. J. (2004). The cortical organization of audio-visual sentence comprehension: an fMRI study at 4 Tesla. *Cognitive brain research*, *20*(2), 111-119.

Cardy, J. E. O., Flagg, E. J., Roberts, W., Brian, J., & Roberts, T. P. (2005). Magnetoencephalography identifies rapid temporal processing deficit in autism and language impairment. *Neuroreport*, *16*(4), 329-332.

Chan, Y. M., Pianta, M. J., & McKendrick, A. M. (2014). Reduced audiovisual recalibration in the elderly. *Frontiers in aging neuroscience*, *6*.

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *Jama*, *290*(3), 337-344.

Courchesne, E., Karns, C.M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D. Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology, 57,* 245-254.

Courchesne, E., Pierce, K., Schumann, C.M., Redcay, E., Buckwalter, J.A., Kennedy, D.P., ... Morgan, J. (2007). Mapping early brain development in autism. *Neuron, 56,* 399-413.

Dager, S. R. (2002). Brain structural abnormalities in young children with autism spectrum disorder. *Neurology, 59,* 184-192.

Dale, A. M., Fischl, B. & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage, 9,* 179-194.

Dawson, G., Munson, J., Webb, S. J., Nalty, T., Abbott, R., & Toth, K. (2007). Rate of head growth decelerates and symptoms worsen in the second year of life in autism. *Biological psychiatry*, *61*(4), 458-464.

de Boer-Schellekens, L., Eussen, M., & Vroomen, J. (2013). Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Frontiers in integrative neuroscience*, *7*.

de Boer-Schellekens, L., Stekelenburg, J. J., Maes, J. P., Van Gool, A. R., & Vroomen, J. (2014). Sound improves diminished visual temporal sensitivity in schizophrenia. *Acta psychologica*, *147*, 136-142.

de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., & Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia research*, *72*(2), 195-203.

de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia Research, 72(2–3),* 195–203.

De Jong, J. J., Hodiamont, P. P. G., Van den Stock, J., & De Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophrenia research*, *107*(2), 286-293.

Delmonte, S., Balsters, J. H., McGrath, J., Fitzgerald, J., Brennan, S., Fagan, A. J., & Gallagher, L. (2012). Social and monetary reward processing in autism spectrum disorders. *Mol. Autism*, *3*(7).

Dhamala, M., Assisi, C. G., Jirsa, V. K., Steinberg, F. L., & Kelso, J. S. (2007). Multisensory integration for timing engages different brain networks.*Neuroimage, 34*(2), 764-773.

Dhamala, M., Assisi, C.G., Jirsa, V.K., Steinberg, F.L., Kelso, J. A. (2007). Multisensory integration for timing engages different brain networks. *Neuroimage, 34,* 764–773.

Diwadkar, V. A., Goradia, D., Hosanagar, A., Mermon, D., Montrose, D. M., Birmaher, B., ... & Zajac-Benitez, C. (2011). Working memory and attention deficits

in adolescent offspring of schizophrenia or bipolar patients: comparing vulnerability markers. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(5), 1349-1354.

Diwadkar, V. A., & Keshavan, M. S. (2002). Newer techniques in magnetic resonance imaging and their potential for neuropsychiatric research. *Journal of psychosomatic research*, *53*(2), 677-685.

Doyle-Thomas, K. A., Goldberg, J., Szatmari, P., & Hall, G. B. (2013). Neurofunctional underpinnings of audiovisual emotion processing in teens with autism spectrum disorders. *Frontiers in psychiatry*, *4*.

Ecker, C., Ginestet, C., Feng, Y., Johnston, P., Lombardo, M. V., Lai, M. C., ... & MRC AIMS Consortium. (2013). Brain surface anatomy in adults with autism: the relationship between surface area, cortical thickness, and autistic symptoms. *JAMA psychiatry*, *70*(1), 59-70.

Ecker, C., Marquand, A., Mourão-Miranda, J., Johnston, P., Daly, E. M., Brammer, M. J., ... & Murphy, D. G. (2010). Describing the brain in autism in five dimensions—magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. *The Journal of Neuroscience*, *30*(32), 10612-10623.

Erlebacher, A., & Sekuler, R. (1971). Response frequency equalization: A bias model for psychophysics. *Perception & Psychophysics*, *9*, 315-320.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433.

Ethofer, T., Van De Ville, D., Scherer, K., & Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Current Biology*, *19*(12), 1028-1033.

Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences 97*: 11050–11055.

Fischl, B., Sereno, M. I. & Dale, A. M. (1999a). Cortical surface-based analysis. II: Inflation, flatterning and a surface-based coordinate system. *NeuroImage, 9,* 195-207.

Fischl, B., Sereno, M. I. , Tootel, R. B. H. & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping, 8*, 272-284.

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,* 381-389.

Foucher, J. R., Lacambre, M., Pham, B. T., Giersch, A., & Elliott, M. A. (2007). Low time resolution in schizophrenia: lengthened windows of simultaneity for visual, auditory and bimodal stimuli. *Schizophrenia Research,97*(1), 118-127.

Foxe, J. J., Molholm, S., Del Bene, V. A., Frey, H. P., Russo, N. N., Blanco, D., ... & Ross, L. A. (2015). Severe multisensory speech integration deficits in high-functioning school-aged children with autism spectrum disorder (ASD) and their resolution during early adolescence. *Cerebral Cortex*, *25*(2), 298-312.

Freitag C. M., Kleser C., Schneider M., von Gontard A. (2007). Quantitative assessment of neuromotor function in adolescents with high functioning autism and Asperger Syndrome. J. Autism Dev. Disord. 37, 948–959.

Freitag, C. M., Luders, E., Hulst, H. E., Narr, K. L., Thompso, P. M., Toga, A. W., … Konrad, C. (2009). Total brain volume and corpus callosum size in medication-naive adolescents and young adults with autism spectrum disorder. *Biological Psychiatry, 66,* 316-319.

Frith, U. (1989). Autism: Explaining the enigma.

Frith, U. (1991). *Autism and Asperger syndrome*. Cambridge University Press.

Froyen, D. J., Bonte, M. L., van Atteveldt, N., & Blomert, L. (2009). The long road to automation: neurocognitive development of letter–speech sound processing. *Journal of Cognitive Neuroscience*, *21*(3), 567-580.

Froyen, D., Van Atteveldt, N., Bonte, M., & Blomert, L. (2008). Cross-modal enhancement of the MMN to speech-sounds indicates early and automatic integration of letters and speech-sounds. *Neuroscience letters*, *430*(1), 23-28.

Fujisaki, W., & Nishida, S. Y. (2009). Audio–tactile superiority over visuo–tactile and audio–visual combinations in the temporal resolution of synchrony perception. *Experimental brain research*, *198*(2-3), 245-259.

Fusar-Poli, P., Borgwardt, S., Crescini, A., Deste, G., Kempton, M. J., Lawrie, S., ... & Sacchetti, E. (2011). Neuroanatomy of vulnerability to psychosis: a voxel-based meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(5), 1175-1185.

Gaffrey, M. S., Kleinhans, N. M., Haist, F., Akshoomoff, N., Campbell, A., Courchesne, E., & Müller, R. A. (2007). A typical participation of visual cortex during word processing in autism: an fMRI study of semantic decision.*Neuropsychologia*, *45*(8), 1672-1684.

García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. *Psychonomic bulletin & review*, *19*(5), 820-846.

Gebauer, L., Foster, N. E., Vuust, P., & Hyde, K. L. (2015). Is there a bit of autism in all of us? Autism spectrum traits are related to cortical thickness differences in both autism and typical development. *Research in Autism Spectrum Disorders*, *13*, 8-14.

Gelder, B. D., Vroomen, J., & Van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology*, *3*(1), 69-86.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., ... & Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature neuroscience*, *7*(8), 801-802.

Geurts, H. M., Ridderinkhof, K. R., & Scholte, H. S. (2013). The relationship between grey-matter and ASD and ADHD traits in typical adults. *Journal of autism and developmental disorders*, *43*(7), 1630-1641.

Ghosh, S., Basu, A., Kumaran, S. S., & Khushu. S. (2010). Functional mapping of language networks in the normal brain using a word-association task. *Indian Journal of Radiology and Imaging. 20(3),* 182–187.

Giampietro, V. P., et al. (2003). Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses. *Human Brain Mapping, 20,* 29–40.

Gibbard, C. R., Ren, J., Seunarine, K. K., Clayden, J. D., Skuse, D. H., & Clark, C. A. (2013). White matter microstructure correlates with autism trait severity in a combined clinical–control sample of high-functioning adults.*NeuroImage: Clinical, 3,* 106-114.

Glozman, J. M. (2014). Selective Involvement of Lingual Gyrus in Working Memory and Perception of Different Types of Visual Stimuli. *Journal of the International Neuropsychological Society*, *20,* 43.

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

Goebel, R., Esposito, F. & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with Brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping, 27,* 392-401.

Goebel, R., Hasson, U., Harel, M., Levy, I. & Malach, R. (2004). Statistical analyses across aligned cortical hemispheres reveal high-resolution population maps of human visual cortex. *NeuroImage Supplement* (Human Brain Mapping, Budapest).

Goebel, R., Staedtler, E., Munk, M.H.J. & Muckli, L. (2002). Cortex-based alignment using functional and structural constraints. *NeuroImage Supplement.*

Gomot, M., Belmonte, M.K., Bullmore, E.T., Bernard, F.A., Baron-Cohen, S. (2008). Brain hyper-reactivity to auditory novel targets in children with high-functioning autism. *Brain, 131,* 2479–2488.

Gottlieb J. From thought to action: the parietal cortex as a bridge between perception, action, and cognition. Neuron. 2007;53(1):9–16.

Grahn, J. A., & Rowe, J. B. (2009). Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *Journal of Neuroscience*, *29*(23), 7540–7548.

Grahn, J. A., & Rowe, J. B. (2013). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex*, *23*(4), 913–921. doi:10.1093/cercor/bhs083

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature neuroscience*, *8*(2), 145-146.

Greimel, E., Nehrkorn, B., Schulte-Rüther, M., Fink, G. R., Nickl-Jockschat, T., Herpertz-Dahlmann, B., ... & Eickhoff, S. B. (2013). Changes in grey matter development in autism spectrum disorder. *Brain Structure and Function*, *218*(4), 929-942.

Grèzes, J., Wicker, B., Berthoz, S., de Gelder, B. (2009). A failure to grasp the affective meaning of actions in autism spectrum disorder subjects. *Neuropsychologia, 47,* 1816–1825.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in cognitive sciences*, *5*(1), 36-41.

Grossman, R. B., Schneps, M. S., & Tager-Flusberg, H. (2009). Slipped lips: onset asynchrony detection of auditory-visual language in autism. *The Journal of Child Psychology and Psychiatry, 50,* 491-497.

Hadjikhani, N., Joseph, R. M., Snyder, J. & Tager-Flusberg, H. (2006). Anatomical differences in the mirror neuron system and social cognition network in autism. *Cerebral Cortex, 16,* 1276-1282.

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.

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*, 441–449.

Haist, F., Adamo, M., Westerfield, M., Courchesne, E., Townsend, J., 2005. The functional neuroanatomy of spatial attention in autism spectrum disorder. *Developmental Neuropsychology, 27 (3),* 425–458.

Hallett M., Lebiedowska M. K., Thomas S. L., Stanhope S. J., Denckla M. B., Rumsey J. (1993). Locomotion of autistic adults. Arch. Neurol. 50, 1304–1308

Hardan, A., Minshew, N., Diwadkar, V., Yorbik, O., Sahni, S., & Keshavan, M. (2002). A voxel based morphometry study of grey matter in autism. In*International Meeting for Autism Research (IMFAR), Orlando, Florida*.

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.

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.

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.

Henry, M. K. (1998). Structured, sequential, multisensory teaching: The Orton legacy. *Annals of Dyslexia*, *48*(1), 1-26.

Heron, J., Roach, N. W., Hanson, J. V., McGraw, P. V., & Whitaker, D. (2012). Audiovisual time perception is spatially specific. *Experimental brain research*,*218*(3), 477-485.

Herrington, J. D., Riley, M. E., Grupe, D. W., & Schultz, R. T. (2015). Successful Face Recognition is Associated with Increased Prefrontal Cortex Activation in Autism Spectrum Disorder. *Journal of autism and developmental disorders*, *45*(4), 902-910.

Hillock, A. R., Powers, A. R., & Wallace, M. T. (2011). Binding of sights and sounds: age-related changes in multisensory temporal processing.*Neuropsychologia*, *49*(3), 461-467.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature neuroscience*, *3*(1), 80-84.

Hubl, D., Bölte, S., Feineis–Matthews, S., Lanfermann, H., Federspiel, A., Strik, W., et al., 2003. Functional imbalance of visual pathways indicates alternative face processing strategies in autism, *Neurology*, *61,* 1232–1237.

Humphreys, K., Hasson, U., Avidan, G., Minshew, N., & Behrmann, M. (2008). Cortical patterns of category-selective activation for faces, places and objects in adults with autism. *Autism Research*, *1*(1), 52-63.

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. *The Journal of Neuroscience*, *27*(47), 13028-13032.

Hyde, K. L., Samson, F., Evans, A.C., Mottron L. (2010). Neuroanatomical differences in brain areas implicated in perceptual and other core features of autism revealed by cortical thickness analysis and voxel-based morphometry. *Human brain mapping. 31 (4),* 556-566.

Iidaka, T., Miyakoshi, M., Harada, T., & Nakai, T. (2012). White matter connectivity between superior temporal sulcus and amygdala is associated with autistic trait in healthy humans. *Neuroscience letters*, *510*(2), 154-158.

Irwin, J. R. (2006). Audiovisual speech integration in children with autism spectrum disorders. *Journal of the Acoustic Society of America*, *120,* 3348-3348.

Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z. & Herskovits E. H. (2010). Predictive models of autism spectrum disorder based on brain regional cortical thickness. *NeuroImage, 50,* 589-599.

Jones, J., Jarick, M. (2006). Multisensory integration of speech signals: The relationship between space and time. *Experimental Brain Research, 174,* 588–594.

Jones, S. E., Buchbinder, B. R., Aharon, I. (2000). Three-dimensional mapping of cortical thickness using Laplace's Equation. *Human brain mapping. 11(1),* 12-32

Just, M.A., Cherkassky, V.L., Keller, T.A., Kana, R.K., Minshew, N.J. (2006). Functional and anatomical cortical underconnectivity in autism: evidence from an fMRI Study of an executive function task and corpus callosum morphometry. *Cerebral Cortex,* bhl006.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217–250.

Kannurpatti, S. S., Motes, M. A., Rypma, B., & Biswal, B. B. (2010). Neural and vascular variability and the fMRI-BOLD response in normal aging. *Magnetic resonance imaging*, *28*(4), 466-476.

Keehn, B., Brenner, L., Palmer, E., Lincoln, A.J., Muller, R.-A., 2008. Functional brain organization for visual search in ASD. *Journal of the International Neuropsychological Society*, *14*, 990–1003.

Kim. J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., ... Evans, A. C. (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage, 27*, 210-221.

Kirkovski, M., Enticott, P. G., Maller, J. J., Rossell, S. L., & Fitzgerald, P. B. (2015). Diffusion tensor imaging reveals no white matter impairments among adults with autism spectrum disorder. *Psychiatry Research: Neuroimaging*.

Knaus, T.A., Silver, A.M., Lindgren, K.A., Hadjikhani, N., Tager-Flusberg, H., (2008). fMRI activation during a language task in adolescents with ASD. Journal of the International *Neuropsychological Society, 14*, 967–979.

Koolschijn, P. C. M., Geurts, H. M., van der Leij, A. R., & Scholte, H. S. (2015). Are Autistic Traits in the General Population Related to Global and Regional Brain Differences?. *Journal of autism and developmental disorders,45*(9), 2779-2791.

Kosaka, H., Omori, M., Munesue, T., Ishitobi, M., Matsumura, Y., Takahashi, T., ... & Morita, T. (2010). Smaller insula and inferior frontal volumes in young adults with pervasive developmental disorders. *Neuroimage*, *50*(4), 1357-1363.

Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., & Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI-activation patterns during the perception of social signals from voice and face. *Human brain mapping*, *31*(7), 979-991.

Kubler A., Dixon V., Garavan H. (2006). Automaticity and reestablishment of executive control-an fMRI study. *Journal of Cognitive Neuroscience, 18,* 1331–1342

Lamb, T. D., & Pugh, E. N. (1992). A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. *The Journal of physiology*, *449*(1), 719-758. Lane, A. E., Young, R. L., Baker, A. E., & Angley, M. T. (2010). Sensory processing subtypes in autism: Association with adaptive behavior. *Journal of autism and developmental disorders*, *40*(1), 112-122.

Langen, M., Durston, S., Staal, W. G., Palmen, S. J., & van Engeland, H. (2007). Caudate nucleus is enlarged in high-functioning medication-naive subjects with autism. *Biological psychiatry*, *62*(3), 262-266.

Langen, M., Leemans, A., Johnston, P., Ecker, C., Daly, E., Murphy, C. M., ... & AIMS Consortium. (2012). Fronto-striatal circuitry and inhibitory control in autism: findings from diffusion tensor imaging tractography. *Cortex*, *48*(2), 183-193.

Latinus, M., Crabbe, F., & Belin, P. (2011). Learning-induced changes in the cerebral processing of voice identity. *Cerebral Cortex*, *21*(12), 2820-2828.

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.

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., Chabriat, H., (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging 13,* 534–546.

Leech R, Sharp DJ (July 2013). "The role of the posterior cingulate cortex in cognition and disease". *Brain* 137 (Pt 1): 12–32.

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism Developmental Disorders, 37,* 894-910.

Lennie, P. (1981). The physiological basis of variations in visual latency. *Vision research, 21*(6), 815-824.

Lewis, R., Noppeney, U. (2010). Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *Journal of Neuroscience, 30,* 12329–39.

Li, H., Xue, Z., Ellmore, T. M., Frye, R. E., and Wong, S. T. (2012). Network-based analysis reveals stronger diffusion-based connectivity and different correlations with language skills in brains of children with high functioning autism spectrum disorders. Human Brain Mapping, 35 (2), 396–413.

Liss, M., Fein, D., Allen, D., Dunn, M., Feinstein, C., Morris, R., ... & Rapin, I. (2001). Executive functioning in high-functioning children with autism.*Journal of Child Psychology and Psychiatry*, *42*(2), 261-270.

Love, S. A., (2011). An fMRI and preychophysical investiagtion of temporal factor of audiovisual integration, PhD Thesis

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments.*PloS one*, *8*(1), e54798.

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.

Love, S. A., Petrini, k., Pernet, C., Latinus, M., & Pollick., F. (in preparation). Different neural correlates for SJs and TOJs.

Love, S. A., Pollick, F. E., & Petrini, K. (2012). Effects of experience, training and expertise on multisensory perception: investigating the link between brain and

behavior. In *Cognitive behavioural systems* (pp. 304-320). Springer Berlin Heidelberg.

Loveland, K. A., Steinberg, J. L., Pearson, D. A., Mansour, R., & Reddoch, S. (2008). Judgments of auditory-visual affective congruence in adolescents with and without autism: a pilot study of a new task using fMRI. *Perceptual and motor skills*, *107*(2), 557-575.

Magnée, M., de Gelder, B., van England, H., Kemner, C. (2008). Audiovisual speech integration in pervasive developmental disorder: evidence from event-related potentials. *Journal of Child Psychology and Psychiatry*, *49,* 995-1000.

Maier, J. X., Di Luca, M., & Noppeney, U. (2011). Audiovisual asynchrony detection in human speech. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 245.

Martin, B., Giersch, A., Huron, C., & van Wassenhove, V. (2013). Temporal event structure and timing in schizophrenia: preserved binding in a longer "now". *Neuropsychologia*, *51*(2), 358-371.

Mashal, N., Solodkin, A., Dick, A. S., Chen, E. E., & Small, S. L. (2012). A network model of observation and imitation of speech. *Frontiers in psychology*,*3*.

Massaro, D. W. (1998). Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. MIT Press, Cambridge, MA

McAlonan, G. M., Daly, E., Kumari, V., Critchley, H. D., van Amelsvoort, T., Suckling, J., ... Murphy, D. G. (2002). Brain anatomy and sensorimotor gating in Asperger's syndrome. *Brain, 127,* 1594 -1606.

McCarthy, G., Puce, A., Belger, A., Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9,* 431–44.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

McKay, L. S., Simmons, D. R., McAleer, P., Marjoram, D., Piggot, J., & Pollick, F. E. (2012). Do distinct atypical cortical networks process biological motion

information in adults with Autism Spectrum Disorders?. *Neuroimage*,*59*(2), 1524-1533.

Leech, R., Braga, D. J. Sharp (2013). Echoes of the brain within the posterior cingulate cortex. *The Journal of Neuroscience* 32 (1): 215–222.

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*,*17*(19), 1692-1696.

Miller, L. M., D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *Journal of Neuroscience, 25,* 5884–5893.

Mizuno, A., Villalobos, M. E., Davies, M. M., Dahl, B. C., & Müller, R. A. (2006). Partially enhanced thalamocortical functional connectivity in autism. *Brain research*, *1104*(1), 160-174.

Mongillo, E., Irwin, J., Whalen, D., Klaiman, C., Carter, A., & Schultz, R. (2008). Audiovisual processing in children with and without autism spectrum disorders. *Journal of Autism Developmental Disorder, 38,* 1349-1358.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research*, *17*(1), 154-163.

Mossbridge, J. A., Fitzgerald, M. B., O'Connor, E. S., & Wright, B. A. (2006). Perceptual-learning evidence for separate processing of asynchrony and order tasks. *The Journal of neuroscience, 26(49),* 12708-12716.

Müller, R.-A. (2007). A typical participation of visual cortex during word processing in autism: an fMRI study of semantic decision. *Neuropsychologia, 45*, 1672–1684.

Narr, K. L., Woods, R. P., Lin, J., Kim, J., Phillips, O. R., Del'Homme, M., Levitt, J. G. (2009). Widespread cortical thinning is a robust anatomical marker for attention deficit/hyperactivity disorder ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry. 48 (10),* 1014.

Narr, K. L, Woods, R. P., Thompson, P. M., Szeszko, P., Robinson, D., Dimtcheva, T., Bilder, R. M. (2007). Relationships between IQ and regional cortical gray matter thickness in healthy adults. *Cerebral Cortex. 17 (9)*, 2163-2171.

Neufeld, J., Sinke, C., Zedler, M., Dillo, W., Emrich, H. M., Bleich, S., & Szycik, G. R. (2012). Disinhibited feedback as a cause of synesthesia: evidence from a functional connectivity study on auditory-visual synesthetes. *Neuropsychologia*, *50*(7), 1471-1477.

Neufeld, J., Sinke, C., Zedler, M., Dillo, W., Emrich, H. M., Bleich, S., & Szycik, G. R. (2012). Disinhibited feedback as a cause of synesthesia: evidence from a functional connectivity study on auditory-visual synesthetes. *Neuropsychologia*, *50*(7), 1471-1477.

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: a review. *Neuroscience & Biobehavioral Reviews*, *36*(2), 836-854.

Olson, I. R., Gatenby, J. G., Gore, J. C. (2002). A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Cognitive Brain Research, 14*, 129–138

Osipowicz, K., Bosenbark, D. D., & Patrick, K. E. (2015). Cortical Changes Across the Autism Lifespan. *Autism Research*, *8*(4), 379-385.

Pearl, D., Yodashkin-Porat D, Katz N, Valevski A, Aizenberg D, Sigler M, Kikinzon L, et al. (2009). Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Comprehensive Psychiatry*, *50(2)*, 186–192.

Pelphrey, K.A. (2010). Cortical integration of audio-visual speech and non-speech stimuli. *Brain and Cognition, 74*, 97–106.

Perkins, T.J., Stokes, M.A., McGillivray, J.A., Mussap, A.J., Cox, I.A., Maller, J.J., Bittar, R. G., 2014. Increased left hemisphere impairment in High-functioning autism: a tract based spatial statistics study. *Psychiatry Research: Neuroimaging 224*, 119–123.

Perrodin, C., Kayser, C., Logothetis, N. K., & Petkov, C. I. (2015). Natural asynchronies in audiovisual communication signals regulate neuronal multisensory interactions in voice-sensitive cortex. *Proceedings of the National Academy of Sciences*, *112(1),* 273-278.

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F., Puce, A. & Pollick, F. E. (2009a). Multisensory integration of drumming actions: musical expertise affects perceived audiovisual asynchrony, *Experimental brain research, 198,* 339–52.

Petrini, K., Holt, S. P. and Pollick, F. E. (2010a). Expertise with multisensory events eliminates the effect of biological motion rotation on audiovisual synchrony perception, *Journal of Vision, 10(5),* 1–14.

Petrini, K., Holt, S. P., & Pollick, F. (2010b). Expertise with multisensory events eliminates the effect of biological motion rotation on audiovisual synchrony perception. *Journal of Vision*, *10(5),* 2.

Petrini, K., Pollick, F.E., Dahl, S., McAleer, P., McKay, L., Rocchesso, D., Waadeland, C. H., Love, S., Avanzini, F., Puce, A. (2011). Action expertise reduces brain activity for audiovisual matching actions: an fMRI study with expert drummers. *Neuroimage, 56,* 1480–1492.

Petrini, K., Russell, M. and Pollick, F. E. (2009b). When knowing can replace seeing

Pierce, K., & Redcay, E. (2008). Fusiform function in children with an autism spectrum disorder is a matter of "who". *Biological psychiatry*, *64*(7), 552-560.

Pierce, K., Haist, F., Sedaghat, F., & Courchesne, E. (2004). The brain response to personally familiar faces in autism: findings of fusiform activity and beyond. *Brain*, *127*(12), 2703-2716.

Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). Portraits or people? Distinct representations of face identity in the human visual cortex. *Journal of Cognitive Neuroscience*, *17*(7), 1043-1057.

Powers, A. R., Hillock, A.R., Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *Journal of Neuroscience, 29,* 12265-12274.

Ranganath, C., Johnson, M.,K., D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia, 41,* 378–389.

Redcay, E. & Courchesne, E. (2005). When is the brain enlarged in autism? A meta-analysis of all brain size reports. *Biological Psychiatry, 58,* 1-9.

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. *Neuroscience and Biobehavioural Review 32,* 123-142.

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.

Redcay, E., Courchesne, E., (2008). Deviant functional magnetic resonance imaging patterns of brain activity to speech in 2–3-year-old children with autism spectrum disorder. *Biological Psychiatry, 64,* 589–598.

Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental Disorders*, *43*(4), 775-784.

Roine, U., Roine, T., Salmi, J., Wendt, N. V., Leppämäki, S., Rintahaka, P., ... & Sams, M. (2013). Increased coherence of white matter fiber tract organization in adults with Asperger syndrome: a diffusion tensor imaging study. *Autism Research*, *6*(6), 642-650.

Rojas, D. C., Peterson, E., Winterrowd , E., Reite, M. L., Rogers, S. J., & Tregellas, J. R. (2006). Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. *BMC Psychiatry,* 6:56.doi:10.1186/1471-244X-6-56.

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., van der Kouwe, A., Jenkins, B. G., Dale, A. M., Fischl, B., (2002). Regional and

progressive thinning of the cortical ribbon in Huntington's disease. *Neurology, 58,* 695-701.

Ross, L. A., Del Bene1, V. A., Molholm1, S., Frey, H-P., & Foxe, J. J. (2015). Sex differences in multisensory speech processing in both typically developing children and those on the autism spectrum. *Frontiers in Neuroscience, 9,* 185 1-13.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia Research, 97,* 173–183.

Russo, N., Foxe, J. J., Brandwein, A. B., Altschuler, T., Gomes, H., & Molholm, S. (2010). Multisensory processing in children with autism: high-density electrical mapping of auditory–somatosensory integration. *Autism Research, 3*(5), 253-267.

Sahyoun, C., Soulières, I., Belliveau, J., Mottron, L., Mody, M., (2009). Cognitive differences in pictorial reasoning between high-functioning autism and Asperger's syndrome. *Journal of Autism and Developmental Disorders, 39,* 1014–1023.

Saito, Y., Suga, M., Tochigi, M., Abe, O., Yahata, N., Kawakubo, Y., ... & Yamasue, H. (2013). Neural correlate of autistic-like traits and a common allele in the oxytocin receptor gene. *Social cognitive and affective neuroscience*, nst136.

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., Morris, J. C., Dale, A. M., Fischl, B., (2004). Thinning of the cerebral cortex in aging. *Cerebral Cortex, 14,* 721-730.

Sato, K., Kaji, R., Matsumoto, S., & Goto, S. (2007). Cell type-specific neuronal loss in the putamen of patients with multiple system atrophy. *Movement disorders*, *22*(5), 738-742.

Scheel, C., Rotarska-Jagiela, A., Schilbach, L., Lehnhardt, F. G., Krug, B., Vogeley, K., & Tepest, R. (2011). Imaging derived cortical thickness reduction in high-functioning autism: key regions and temporal slope. *Neuroimage*, *58*(2), 391-400.

Scherf, K. S., Legault, J., Turney, I., Elbich, D., Minshew, N., & Behrmann, M. (2015). Axial Diffusivity in the ILF and IFOF is Related to Autism Symptom Severity. *Journal of vision*, *15*(12), 639-639.

Scherf, K. S., Luna, B., Minshew, N., & Behrmann, M. (2010). Location, location, location: alterations in the functional topography of face-but not object-or place-related cortex in adolescents with autism. *Frontiers in Human Neuroscience*, *4*.

Schmitz, N., Rubia, K., Daly, E., Smith, A., Williams, S., Murphy, D. (2006). Neural correlates of executive function in autistic spectrum disorders. *Biological Psychiatry, 54*(1), 7–16.

Sekuler, R., & Erlebacher, A. (1971). The two illusions of Müller-Lyer: Confusion theory reexamined. *The American journal of psychology*, 477-486.

Setti A., Finnigan S., Sobolewski R., McLaren L., Robertson I. H., Reilly R. B., et al. (2011b).Audiovisual temporal discrimination is less efficient with aging: an event-related potential study. *Neuroreport, 22,* 554–558.

Setti, A., Burke, K. E., Kenny, R. A., & Newell, F. N. (2011a). Is inefficient multisensory processing associated with falls in older people?. *Experimental brain research*, *209*(3), 375-384.

Shams, L., Kamitani, Y. & Shimojo, S. (2002). Visual illusion induced by sound. *Cogntive. Brain Research, 14,* 147–152.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408,*788-788.

Shukla, D. K., Keehn, B., Smylie, D. M., & Müller, R. A. (2011). Microstructural abnormalities of short-distance white matter tracts in autism spectrum disorder.*Neuropsychologia*, *49*(5), 1378-1382.

Shukla, D.K., Keehn, B., Muller, R.A., (2011). Tract-specific analyses of diffusion tensor imaging show widespread white matter compromise in autism spectrum disorder. *Journal of Child Psychology and Psychiatry 52,* 286–295.

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P. & Pollick, F. E. (2009). Vision in Autism Spectrum Disorders. *Vision Research, 49 (22),* 2705-2739.

Singh-Curry V, Husain M. (2009).The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy. *Neuropsychologia.* 47(6), 1434–1448.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex, 17,* 2387–2399.

Smith, E., Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry, 48,* 813-821.

Snyder, J. J., & Chatterjee, A. (2004). Spatial-temporal anisometries following right parietal damage. *Neuropsychologia*, *42*(12), 1703-1708.

Snyder, J. J., Chatterjee, A. (2004). Spatial-temporal anisometries following right parietal damage. *Neuropsychologia, 42,* 1703–1708.

Sowell, E. R., Peterson, B. S., Kan, E., Woods, R. P., Yoshii, J., Bansal, R., ... & Toga, A. W. (2007). Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. *Cerebral cortex,17*(7), 1550-1560.

Stanfield, A. C., McIntosh, A. M., Spencer, M. D., Philip, R., Gaur, S., & Lawrie, S. M. (2008). Towards a neuroanatomy of autism: a systematic review and meta-analysis of structural magnetic resonance imaging studies. *European Psychiatry*, *23*(4), 289-299.

Stanfield, A. C., McIntosh, A., M., Spencer, M. D., Philip, R., Gaur, S. & Lawrie, S. M. (2008). Towards a neuroanatomy of autism: A systematic review and meta-analysis of structural magnetic resonance imaging studies. *European Psychiatry, 23,* 289-299.

Stekelenburg, J. J., Maes, J. P., Van Gool, A. R., Sitskoorn, M., & Vroomen, J. (2013). Deficient multisensory integration in schizophrenia: An event-related potential study. *Schizophrenia research*, *147*(2), 253-261.

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.

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research, 227(2),* 249–261.

Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *Neuroimage, 49,* 3308–18.

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.

Stevenson, R. A., Kim, S., & James, T. W. (2009). An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Experimental Brain Research, 198(2–3),* 183–194.

Stevenson, R. A., Mullin, C. R., Wallace, M. T., & Steeves, J. K. (2013). fMRI-guided TMS of the superior temporal sulcus impairs multisensory temporal processing. *Multisensory Research, 26,* 208-208.

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.

Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., & James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *Neuroimage, 55*(3), 1339-1345.

Szelag, E., Kowalska, J., Galkowski, T., & Pöppel, E. (2004). Temporal processing deficits in high-functioning children with autism. *British Journal of psychology*, *95*(3), 269-282.

Szycik, G. R., Münte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H. M., & Dietrich, D. E. (2009). Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophrenia research*, *110*(1), 111-118.

Szycik, G. R., Munte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H. M., Dietrich, D. E. (2009). Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophrenia Research, 110,* 111–118.

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.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging.

Toal, F., Daly, E. M., Page, L., Deeley, Q., Hallahan, B., Bloemen, O., ... Murphy, D. G. (2010). Clinical and anatomical heterogeneity in autistic spectrum disorder: a structural MRI study. *Psychological Medicine, 40,* 1171-1181.

Toga, A. W., Thompson, P. M., & Sowell, E. R. (2006). Mapping brain maturation. *Focus*.

van der Smagt, M. J., van Engeland, H., & Kemner, C. (2007). Can you see what is not there? Low-level auditory-visual integration in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37,* 2014-2019.

Vander Wyk, B. C., Ramsay, G. J., Hudac, C. M., Jones, W., Lin, D., Klin, A., ... & Pelphrey, K. A. (2010). Cortical integration of audio–visual speech and non-speech stimuli. *Brain and cognition*, *74*(2), 97-106.

Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & Van De Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Attention, Perception and Psychophysics, 70,* 955-968.

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory–visual speech perception. *Neuropsychologia, 45(3),* 598–607.

Vander Wyk, B. C., Ramsay, G. J., Hudac, C. M., Jones, W., Lin, D., Klin, A., ... & Pelphrey, K. A. (2010). Cortical integration of audio–visual speech and non-speech stimuli. *Brain and cognition*, *74*(2), 97-106.

Vatakis, A., & Spence, C. (2006a). Audiovisual synchrony perception for music, speech, and object actions. *Brain research*, *1111*(1), 134-142.

Vatakis, A., & Spence, C. (2006b). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neuroscience letters*, *393*(1), 40-44.

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & Psychophysics*, *69*(5), 744-756.

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.

von Saldern, S., & Noppeney, U. (2013). Sensory and striatal areas integrate auditory and visual signals into behavioural benefits during motion discrimination. *Journal of Neuroscience. 33*, 8841–8849.

Waiter, G. D., Williams, J. H., Murray, A. D., Gilchrist, A., Perrett, D. I. & Whiten, A. (2004): A voxel-based investigation of brain structure in male adolescents with autistic spectrum disorder. *Neuroimage, 22*, 619-625.

Walker, L., Gozzi, M., Lenroot, R., Thurm, A., Behseta, B., Swedo, S., & Pierpaoli, C. (2012). Diffusion tensor imaging in young children with autism: biological effects and potential confounds. *Biological psychiatry*, *72*(12), 1043-1051.

Wallace, G. L., Dankner, N., Kenworthy, L., Giedd, J. N., & Martin, A. (2010). Age-related temporal and parietal cortical thinning in autism spectrum disorders. *Brain. 133 (12)*, 3745-3754.

Watson, R., Latinus, M., Charest, I., Crabbe, F., & Belin, P. (2014). People-selectivity, audiovisual integration and heteromodality in the superior temporal sulcus. *Cortex, 50,* 125-136.

Wechsler, D. (1999). WASI manual. *San Antonio: Psychological Corporation*.

Werner, S., & Noppeney, U. (2010). Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *The Journal of Neuroscience, 30*(7), 2662-2675.657.

Whittingham, K. M., McDonald, J. S., & Clifford, C. W. (2014). Synesthetes show normal sound-induced flash fission and fusion illusions. *Vision research,105,* 1-9.

Williams, J. H. G., Waiter, G. D., Gilchrist, S., Perret, D. I., Murray, A. D. & Whiten, A. (2006). Neural mechanisms of imitation and 'mirror neuron' functioning in autistic spectrum disorder. *Neuropsychologia 44 (4),* 610-621.

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.

Williams, L. E., Light, G. A., Braff, D. L., & Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia, 48*(10), 3128-3136.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research, 152*(2), 198-210.

Zielinski, B.A., Prigge M.B., Nielsen, J.A., Froehlich, A.L., Abildskov T.J., Anderson J.S., et al. (2014). Longitudinal changes in cortical thickness in autism and typical development. *Brain, 137(6),* 1799–1812.