Visual-tactile integration

and individual differences in speech perception

by

Katie Bicevskis

B.A. (Visual), Australian National University, 2001

Graduate Diploma in Arts, University of Melbourne, 2010

Postgraduate Diploma in Arts, University of Melbourne, 2012

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Master of Arts

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Linguistics)

The University of British Columbia

(Vancouver)

September 2015

© Katie Bicevskis, 2015

Abstract

Integration of speech information is evident in audio-visual (McGurk & MacDonald, 1976) and audio-tactile (Gick & Derrick, 2009) combinations and an asymmetric window of multimodal integration exists which is consistent with the relative speeds of the various signals (Munhall et al., 1996; Gick et al., 2010). It is presently unclear whether integration is possible if the audio speech signal is removed. The current thesis utilizes synchronous and asynchronous visual and aero-tactile speech stimuli to investigate potential integration effects of this modality combination and explores the shape of the potential window of visual-tactile integration. Results demonstrate that the aero-tactile stimulus significantly affects categorization of speech segments so that individuals are more likely to perceive a voiceless aspirated stop when they experience a combination of visual-tactile integration which reflects relative speeds of light and speech airflow is also evident. These results add to our knowledge of multimodal speech integration and support notions that speech is perceived as a holistic, modality neutral event.

Children with Autism Spectrum Disorder (ASD) have exhibited differential multimodal integration behaviour (Gelder et al., 1991; Mongillo, et al., 2008; Irwin et al., 2011; Stevenson, et al., 2014) and differences in temporal acuity (Stevenson, et al., 2014) as compared to typically developing children, however it is unclear whether these differential findings are specific to this clinical population or can be considered part of a continuum of multimodal integration behaviour which includes typically developed adults. The current thesis examines individual differences in visual-tactile integration based on temporal acuity and behavioural traits associated with ASD, in a typically developed adult population. Results show that temporal acuity and behavioural traits associated with ASD, especially the trait of imagination, significantly influence the range of asynchronous stimuli over which visual-tactile integration occurs and also affect individuals' abilities to differentiate visually similar speech stimuli. These results reveal a relationship between visual-tactile integration rates, traits associated with ASD and temporal acuity and suggest that the differential behaviour observed in child ASD populations forms part of a continuum which extends to typically developed adults.

Preface

This Master's thesis is an original, unpublished work by the author, K. Bicevskis. The research was conducted with the approval of the UBC Behavioural Research Ethics Board, as part of the research project entitled "Processing Complex Speech Motor Tasks", certificate number B04-0337, principal investigator Bryan Gick.

Table of Contents

Abstractii
Prefaceiii
Table of Contentsiv
List of Tables vi
List of Figures
Acknowledgmentsviii
Dedicationix
1. Introduction
1.1 Multimodal Speech Perception
1.2 Individual Differences in Multimodal Perception
1.3 Outline of Current Study
2. Visual-Tactile Integration
2.1 Overview
2.1.1 Hypothesis and Predictions
2.2 Methods
2.2.1 Participants
2.2.2 Procedure
2.2.2.1 Visual-Tactile Integration Task
2.2.3 Stimuli
2.2.3.1 Visual Stimuli
2.2.3.2 Tactile Stimuli
2.2.3.3 Coordination of Stimuli 17

2.3 Analysis and Results	
2.3.1 Visual-only Condition	
2.3.2 Synchronous Condition	
2.3.3 Asynchronous Conditions and Asymmetry	
2.4 Discussion	
3. Individual Differences in Integration Based on Behavioural Attributes and Tempora	d Binding
Windows	
3.1 Overview	
3.1.1 Hypothesis and Predictions	
3.2 Methods	
3.2.1 Simultaneity Judgment Task	
3.2.2 Autism Spectrum Quotient Questionnaire	
3.3 Analysis and Results	
3.3.1 Creation of Temporal Binding Windows	
3.3.2 Autism Spectrum Quotient Score as a Predictor of Integration Rates	
3.3.2.1 Synchronous Condition	
3.3.2.2 Asynchronous Conditions	
3.3.3 Temporal Binding Window Width as a Predictor of Integration Rates	
3.3.3.1 Synchronous Condition	
3.3.3.2 Asynchronous Conditions	
3.3.4 Relationship Between Autism Spectrum Quotient Score and Temporal Bind	ling
Window Width	
3.4 Discussion	
4. Conclusions	
References	
Appendix A: Autism Spectrum Quotient Questionnaire	

List of Tables

Table 1: Results of mixed model with structure: Response ~ Visual stimulus * Puff condition + (1 + (Visual stimulus * Puff condition)/Participant) + (1/Token)......22Table 2: Results of mixed model with structure: Response ~ Visual stimulus * SOA + (1 + Visual $stimulus * SOA/Participant) + (1/Token) \dots 24$ Table 3: Results of mixed model with structure: Response ~ Visual stimulus *Poly(SOA, degree Table 4: Results of mixed model with structure: Response ~ (Visual Stimulus * SOA) + Table 5: Results of mixed model with structure: Response ~ Visual stimulus * Puff condition * Social Skill + $(1 + (Visual stimulus + Puff condition + Social Skill)/Participant) + (1 \setminus Token)$ 39 Table 6: Results of mixed model with structure: Response ~ Visual stimulus * Puff condition * Social Skill + (1 + (Visual stimulus * Puff condition * Social Skill)/Participant) + (1/Token)40 Table 7: Results of mixed model with structure: Response ~ Visual stimulus * SOA * AQ + (1 + 1)Table 8: Results of mixed model with structure: Response ~ Visual stimulus * SOA * Imagination + (1 + (Visual stimulus + SOA + Imagination)/Participant) + (1/Token)......44Table 9: Results of mixed model with structure: Response ~ Visual stimulus * SOA * TBW + (1+ (Visual stimulus + SOA + TBW)/Participant) + (1/Token)......49Table 10: Results of mixed model with structure: Response ~ Visual stimulus * SOA * RTBW + $(1 + (Visual stimulus + SOA + RTBW)/Participant) + (1/Token) \dots 50$ Table 11: Results of mixed model with structure: Response ~ Visual stimulus * SOA * LTBW + $(1 + (Visual stimulus + SOA + LTBW)/Participant) + (1/Token) \dots 51$ Table 12: Results of Pearson's correlations between width of full, right-side and left-side Temporal Binding Window and AQ score53 Table 13: Results of Pearson's correlations between width of right-side Temporal Binding Window and AQ subfactor scores of Social Skill, Communication, Imagination, Attention to

List of Figures

Figure 1: Air puff set up17
Figure 2: Position of sine wave relative to vowel onset of a /pa/ production
Figure 3: Percentage of /ba/ and /pa/ responses in the visual-only condition
Figure 4: Percentage of /pa/ responses in the Oms and visual-only conditions
Figure 5: Percentage of /pa/ responses across Stimulus Onset Asynchronies, split by visual
stimulus
Figure 6: Percentage of a participant's "same" responses across the Stimulus Onset Asynchronies
Figure 7: Percentage of /pa/ responses across Stimulus Onset Asynchronies for participants with
low vs. high Autism Spectrum Quotient (AQ) scores43
Figure 8: Percentage /pa/ responses across Stimulus Onset Asynchronies for participants with
low vs. high Imagination scores, split by visual stimulus45
Figure 9: Percentage /pa/ responses across Stimulus Onset Asynchronies as a function of
Imagination score
Figure 10: Percentage of /pa/ responses across Stimulus Onset Asynchronies for participants with
narrow vs. wide right-side temporal binding windows, split by visual stimulus
Figure 11: Correlations between AQ score and full, and right-side temporal binding windows54
Figure 12: Correlations between the width of the right-side temporal binding window and the
subfactor scores of (from top left) Social Skill, Communication, Imagination, Attention to Detail,
and Attention Switching

Acknowledgments

I am extremely grateful to my supervisor Bryan Gick and committee members Donald Derrick and Eric Vatikiotis-Bateson for all of their helpful feedback and support throughout the project.

Special thanks also go to everyone in the Interdisciplinary Speech Reseach Lab at UBC for being great company and providing delicious snacks.

Thank you also to Kathleen Currie-Hall, Molly Babel and the 530 class on Attention and Salience in Phonetics and Phonology for all of their feedback, particularly in relation to Chapter 2, as well as Martina Wiltschko and the 518 Research Methods class for feedback and interesting discussion about research methodology.

Thank you to Michael Fry for being my speaker, Nick Romero for making an amazing switchbox, Oksana Tkachman for suggestions on an early version of the thesis, Bosko Radanov for advice on Chapter 3 methodology and Ryan Hill and Nicole Anger for useful discussions in the early design of the experiment.

Of course, a big thank you to my family and friends and a huge thank you to my husband Luke for his unending support and encouragement and just generally for being awesome.

Dedication

For Limey

1. Introduction

This thesis has two main goals. The first is to investigate perceptual integration of aero-tactile speech information during visual-tactile speech perception. Secondly, it examines individual differences in integration amongst perceivers, based on temporal acuity and behavioural attributes that have been associated with Autism Spectrum Disorder. Chapter 2 focuses on group perceptual behaviour during visual-tactile speech perception using synchronous and asynchronous stimuli and explores the temporal window within which individuals generally integrate visual-tactile speech stimuli. Results from the current study show that speakers use and integrate tactile speech information into the overall speech signal and that the perceptual behaviour exhibited for visual-tactile stimuli is similar to that found for audio-visual and audiotactile speech integration in terms of the size and asymmetry of the temporal window within which integration normally occurs. Chapter 3 explores individual differences in rates of integration of tactile speech information in relation to an individual's temporal acuity with regard to visual-tactile speech processing, as well as differing levels of behavioural traits thought to be associated with Autism Spectrum Disorder. Results show that temporal acuity and behavioural traits associated with Autism Spectrum Disorder both influence the range of the temporal window over which visual-tactile integration is likely to occur and further, that these factors can influence individuals' abilities to detect subtle differences between visual stimuli. Results also reveal direct correlations between temporal acuity and behavioural traits associated with ASD. Chapter 4 concludes with a summary of the major findings of the current research and discusses future directions. In the remainder of the present chapter, section 1.1 provides a brief review of the literature into multimodal speech research. Following this, individual differences in perception based on temporal acuity and behavioural traits are discussed in section 1.2. Finally, section 1.3 presents an outline of the current study.

1.1 Multimodal Speech Perception

Research into speech perception over the last half century has uncovered a wealth of evidence which demonstrates its multimodal nature. This multimodal research program has primarily focused on the combination of auditory and visual modes. Evidence for the importance of the visual mode is found in visual enhancement of speech perception (Sumby & Pollack, 1954), as well as the well-known perceptual speech illusion, the McGurk effect (McGurk & MacDonald, 1976), where the presentation of incongruent auditory and visual speech stimuli result in an integrated percept. Multimodal integration may be defined as the combination of information from different sources of stimuli so that what is perceived is a fusion of the two stimuli. In the case of McGurk stimuli, this often results in the perception of a novel third speech sound. The effect is robust; unlike many other illusions it is still maintained even when the perceiver is aware of what is happening. Further, it does not require synchrony of stimuli to be effective (Munhall, Gribble, Sacco, & Ward, 1996; van Wassenhove, Grant, & Poeppel, 2007). Munhall et al. (1996) tested asynchronous stimuli ranging from -360ms (where the audio stimulus precedes the visual stimulus) to 360ms (where the visual precedes the audio) in increments of 60ms and found that the effect is maintained across a range of temporal asynchronies ranging from synchronous to 180ms. This range of asynchronies may be thought of as a window of multimodal *integration.* When the visual stimulus leads by more than 180ms, however, the rate of integration significantly declines. This result is predictable as stimuli combinations with greater asynchronies are more likely to be perceived as asynchronous and less likely to be considered components of the same event, therefore the information from the two stimuli are less likely to be integrated. However, when the audio stimulus precedes the visual by just 60ms, the rate of integration also significantly declines, demonstrating that the asynchronous durations over which the effect is maintained are asymmetric. This is explained by appealing to properties of the natural world, i.e. light travels faster than sound therefore people have experience perceiving events where the visual information is received before the audio. An extreme example of this is the perception of fireworks, which are normally viewed from a great distance and the visual 'explosion' is visible moments before the auditory 'explosion' is audible. However, Munhall et al. (1996) argue that people experience more subtle asynchronous events in daily life and through experience come to perceive the slightly asynchronous visual and auditory components

of an event as being simultaneous. Their results show that in fact when the visual stimulus leads the audio by 60ms the rate of integration is slightly higher than when the stimuli is synchronous.

A growing body of research, e.g. Fowler & Dekle (1991), has shown that the tactile mode also contributes information to the speech stream, so more recent research has therefore examined the audio-tactile modality combination. Gick & Derrick (2009) showed that when participants were presented with ambiguous audio stimuli of /pa/ and /ba/, they reported hearing more /pa/ stimuli when the audio stimulus was presented simultaneously with a puff of air to the skin. This result demonstrates that speakers use aero-tactile information available in speech, integrating it as if they were feeling aspiration from the speaker. Further, Gick, Ikegami, & Derrick, (2010) found that, similar to audio-visual speech events, participants integrate audio-tactile speech over a range of temporal asynchronies. When presented with stimulus asynchronies ranging from -300ms (tactile stimulus preceding audio stimulus) to 300ms (audio stimulus preceding tactile stimulus) participants exhibited a window of multimodal integration between -50ms and 200ms. Again the effect is asymmetric in a direction which suggests adherence to the relative speeds of physical properties of the world; sound travels faster than (aspiration) airflow, so integration is more likely to occur when the audio precedes the tactile stimulus. Mayer, Gick, Weigel, & Whalen (2013) explored potential integration of indirect aero-tactile information conveyed through the visual mode. Participants observed videos of a person producing voiced, or voiceless aspirated bilabial stop initial words in front of a candle, with the audio signal partially masked so that audio-only identification was at about 70% accuracy. They found an integration effect such that tokens which showed perturbation of the candle flame resulted in increased reporting of the aspirated stop.

There is a lack of research into visual-tactile speech perception, which is not surprising considering auditory information is traditionally thought to be primary in the speech signal. However, Gick, Jóhannsdóttir, Gibraiel, & Mühlbauer (2008) examined the influence of tactile information on visual speech perception using the Tadoma method (Alcorn, 1932), a speech comprehension technique where perceivers place a hand in a particular position over the mouth and jaw of a speaker in order to perceive tactile speech information. They used untrained perceivers and found that some participants' perception of VCV syllables improved by around

ten per cent when they felt the speaker's face whilst watching them silently speak, as opposed to when they had access to only the visual speech information.

This research into multimodal speech perception shows that perceptual integration can occur with audio-visual, audio-(aero)tactile and visual-tactile modality combinations, meaning that information from all of these modalities is used by perceivers to comprehend speech. Regarding the research into audio-visual and audio-(aero)tactile speech integration, it also reveals that there is a general temporal window over which this integration is likely to occur and that this window of multimodal integration is asymmetric in a direction which reflects the relative speeds of certain physical properties. Currently unexplored is the potential for individuals to integrate speech information from the visual and aero-tactile modes, in the absence of the original auditory speech signal. Further, if integration involving this modality combination does occur, it is as yet unclear whether visual-(aero)tactile speech perception obeys the same principles in terms of temporal window properties as were observed for audio-visual and audio-(aero)tactile combinations of speech information. If visual-(aero)tactile speech integration does take place, this finding would contribute new knowledge to our understanding of multimodal speech perception and may also support the idea that speech is not perceived as an audio signal which is supplemented by information from other modes, but as more of a holistic event where speech primitives are modality neutral.

This section has considered multimodal perceptual integration in terms of group behaviour. There are, however, individual differences amongst speaker populations in terms of the degree to which they integrate stimuli from different modalities. Some potential reasons for these differences are discussed below in section 1.2.

1.2 Individual Differences in Multimodal Perception

Despite the robustness of McGurk effect, the effect is not experienced by all speakers and the rate of integration can vary amongst those who do perceive it. Several studies (e.g. Gelder, Vroomen, & Van der Heide, 1991; Mongillo et al. 2008; Irwin, Tornatore, Brancazio, & Whalen, 2011; Stevenson et al., 2014) have found differences in the perception of incongruent

synchronous audio-visual stimuli in children with Autism Spectrum Disorder (ASD) as compared to Typically Developed (TD) child populations. Their results have shown that children with ASD were significantly less likely to report the integrated speech sound and more likely to report the audio stimulus, as compared to TD children. Though there is evidence that this lack of integration may be at least in part due to poor speech reading skills (e.g. Williams, Massaro, Peel, Bosseler, & Suddendorf , 2004) and deficits in facial identity and expression recognition in ASD populations (Davies, Bishop, Manstead, & Tantam, 1994), Stevenson et al. (2014) reported no significant differences between ASD and TD populations in visual-only speech categorization tasks. They suggest that the lower rate of McGurk percepts in the ASD population may have been due to difficulties integrating two stimuli from any modalities. This is consistent with the weak central coherence theory (Frith, 1989); in TD individuals, low-level information components are combined to form higher-level meaning (global processing), but in individuals with ASD, information processing is thought to be characterized by a focus on these component parts (local processing), often resulting in exceptional skill when local processing is called for, but deficits related to processing global meaning.

Stevenson et al. (2014) also found that children with ASD exhibited significantly wider temporal binding windows (for audio-visual speech stimuli) than TD children. A temporal binding window is a window of temporal asynchronies over which two separate stimuli are perceived to occur simultaneously. This window is calculated based on rates of participants' simultaneity judgments of multimodal stimuli across a range of stimulus asynchronies (discussed in detail in section 3.3.1). This type of window is distinct from the *window of multimodal integration* discussed in section 1.1, which refers to a range of stimulus asynchronies over which speech information from more than one modality is integrated. Stevenson et al. (2014) also found that the width of temporal binding windows was negatively correlated with the rate of integration of incongruent stimuli for children with ASD; the wider a participant's window, the less likely they were to integrate synchronous audio-visual stimuli (a correlation which wasn't found in the TD population). These findings suggest that poor temporal acuity, exemplified by wider temporal binding windows, is connected with difficulties integrating multimodal speech. This is supported by other research into temporal binding window width in these two populations. Foss-Feig et al. (2010) showed that children with ASD have wider windows when tested on the flash-beep illusion (a perceptual illusion where a single flash of light is often perceived as two flashes when

it is accompanied by two audio beeps) and Kwakye, Foss-Feig, Cascio, Stone, & Wallace (2011) found this same difference in window widths when the two populations were tested on audiovisual, but not visual-visual stimuli in temporal order judgment tasks. Bebko, Weiss, Demark, & Gomez (2006) tested ASD and TD groups, as well as children with other developmental disabilities in an audio-visual preferential looking task and found that children with ASD did not detect temporal asynchrony in speech stimuli, again suggesting temporal processing difficulties unique to the ASD group.

Outside of clinical populations, studies involving simultaneity judgment tasks have demonstrated a great deal of individual variation in temporal binding window width (e.g. Miller & D'esposito, 2005; Stevenson, Zemtsov, & Wallace, 2012) and rates of integration in audio-visual speech perception (Stevenson et al., 2012). Stevenson et al. (2012) examined individual rates of McGurk percepts and temporal binding window widths within a TD adult population and found a correlation between integration rates and the widths of the right side of participants' temporal binding windows (where the visual stimulus leads the audio).

It is evident from this research that child ASD populations differ from TD groups in terms of their rates of audio-visual speech integration and widths of temporal binding windows. It is also apparent that audio-visual perceptual integration behaviour and window widths vary within an adult TD group. What remains unclear from this research is whether these findings can be related, specifically, whether the differential perceptual behaviour evident in child ASD populations may form one end of a continuum which also includes the variable perceptual behaviour of TD populations, and adults. The alternative would propose that this differential perceptual behaviour is confined to populations with a clinical diagnosis of ASD. A study which looked at the trajectory of development of audio-visual integration in child ASD and TD populations (Taylor, Isaac, & Milne, 2010) found that although the ASD population exhibited lower rates of integration compared to the TD group at the youngest age examined (7yrs), rates of integration by the oldest age tested (16yrs) were comparable in both groups, suggesting that audio-visual processing differences between ASD and TD populations that are present in childhood may disappear by adulthood.

The basis of the differences in perceptual integration in ASD populations is difficult to interpret due to a lack of a singular diagnosis for ASD. ASD is instead characterised by a range of atypical behaviours including difficulties in social interaction and communication, as well as repetitive behaviours. An attempt to measure certain behavioural attributes thought to be associated with ASD has been made with the design of the Autism Spectrum Quotient questionnaire (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). The questionnaire is comprised of fifty questions grouped into five subcategories that have been associated with ASD: social skills, communication skills, imagination, attention to detail, and attention switching/tolerance of change. These categories were developed based on research which found a triad of impairments (Wing & Gould, 1979) amongst children with ASD. This research showed that the factors of social impairment, language impairment and repetitive activities (rather than imaginative play) tended to cluster together amongst these individuals. Further categories were then developed based on existing evidence regarding cognitive impairments in ASD (Baron-Cohen et al., 2001). The questionnaire is not a diagnostic, but was originally designed as a screening tool for ASD. However, it has also been used as a tool in the investigation of individual variation in speech perception (Yu, 2010), production (Yu, Abrego-Collier, & Sonderegger, 2013), non-speech multimodal perception (Donohue, Darling, & Mitroff, 2012), visual perception (Grinter, Maybery, Van Beek, Pellicano, Badcock, & Badcock, 2009) and sensory illusion perception (Palmer, Paton, Hohwy, & Enticott, 2013; Chouinard, Noulty, Sperandio, & Landry, 2013) in TD adult populations, with the assumption that behavioural traits thought to be associated with ASD can be associated with differences in cognitive processing style. Yu (2010) examined normalization in coarticulation contexts and found that women with low Autism Spectrum Quotient scores (equated with levels of behavioural traits more distant from those of individuals with an ASD diagnosis) were less likely to normalize than men, and women with high Autism Spectrum Quotient scores. He looked at the context of /s/ and /ʃ/ occurring before /a/ and /u/. Though /s/ has lower, more $/\int/$ -like frequency before /u/ due to greater lip protrusion (than when occurring before unrounded vowels), speakers normalise for this effect and still perceive /s/ in this context. However in this study it was found that women with low Autism Spectrum Quotient scores were less likely to perceive /s/ before /u/. Further, poor attention switching (strong focus of attention) and good communication skills were interpreted as being correlated with more normalization. Yu et al., (2013) found that participants with poor attention switching/strong

focus of attention imitated the voice onset time of a speaker more than participants who were better able to switch attention. Exploring multimodal non-speech processing, Donohue et al. (2012) ascertained the temporal asynchrony at which participants were most likely to report simple, asynchronous, non-speech audio and visual pairings as being simultaneous and compared these peaks to the participants' Autism Spectrum Quotient scores. Results showed a significant correlation between the measures in that as the peak asynchrony moved leftward, Autism Spectrum Quotient scores increased, demonstrating that participants with higher Autism Spectrum Quotient scores were more likely to perceive audio-leading stimuli as simultaneous. The subfactors of social skills, communication skills, imagination, and attention switching were also correlated with the peak asynchrony and a regression analysis found that attention switching significantly predicted the peak asynchrony judged simultaneous by participants. These results reveal that in adult TD populations, an individual's behaviour in relation to speech production and perception, as well as non-speech multimodal perception, can be predicted by levels of behavioural traits associated with ASD, as measured by the Autism Spectrum Quotient. This questionnaire may therefore be a useful tool in the current study investigating individual variation in visual-tactile speech perception to explore potential connections between rates of visual-tactile speech integration, temporal binding window widths and levels of behavioural traits thought to be associated with ASD, in an adult TD population. It is important to note that the current study is not concerned with using the Autism Spectrum Quotient as a diagnostic, but rather to evaluate behavioural traits present in the general population as potential predictors of individual variation in multimodal integration. If correlations are found, this would suggest that differences in multimodal perception observed in child ASD populations are not specific to this clinical population, but are one end of a continuum on which the perceptual behaviour of TD adults also exists.

1.3 Outline of Current Study

The second chapter of this study investigates multimodal speech perception in the absence of the original audio speech signal. It uses visual and aero-tactile speech stimuli in a task that explores potential integration of visual-tactile speech. It also examines the shape of the potential window

of visual-tactile integration by employing both synchronous and asynchronous stimuli in the task. The purpose of the study is to examine whether participants use both the visual and aero-tactile speech information so that they are more likely to perceive the visual-tactile stimuli as containing an aspirated segment, as opposed to when they are presented with visual-only stimuli. If participants perceive more aspirated segments in the condition containing the tactile stimuli, this would suggest integration of the aero-tactile speech information. It should be noted that integration involving these stimuli does not result in a novel third percept, but combines the aspiration information with the visual information to create the perception of an English voiceless aspirated stop. If participants are found to exhibit integrated percepts, this would suggest that visual-(aero)tactile speech stimuli is perceived by speakers in a similar manner to that of other modality combinations, that is, speakers are able to utilise speech information from these two modalities and this information can be enough to affect categorization. Participants undertake a Visual-Tactile Integration task where silent videos of a speaker producing English voiceless aspirated and voiced bilabial stops in CV contexts (/pa/ and /ba/) are presented with or without a puff of air to the skin, which is intended to simulate the speaker's aspiration of the voiceless bilabial stop. The tactile stimulus is presented both synchronously and at various asynchronies with respect to the visual stimulus. English bilabial stops are generally considered to be visemes, speech sounds which are thought to be visually indiscriminable (Fisher, 1968) which should indicate that participants will perform at chance levels when presented with visualonly /pa/ and /ba/ stimuli. It should be noted however, that in contrast to this commonly accepted proposal a perception study which had participants identify p/, b/ and m/ from videos of a talking face (Abel, Barbosa, Mayer, & Vatikiotis-Bateson, 2011) showed a response bias against /b/, suggesting that participants may detect differences between bilabial segments based on visual information. In either case, the introduction of the tactile stimulus to the speech signal may provide enough additional speech information to shift categorization in favour of the aspirated segment.

The third chapter of the current study is concerned with individual variation in a TD population. It examines possible relationships between integration of visual-tactile speech, temporal acuity and cognitive processing style based on behavioural attributes associated with ASD. Participants undertake a Simultaneity Judgment task in which visual and aero-tactile stimuli are presented at various temporal asynchronies and participants are asked to judge whether or not the stimuli are simultaneous. Results from this task are used to determine each participant's temporal binding window. Participants then complete the Autism Spectrum Quotient questionnaire and overall scores and subfactor scores are calculated for each participant. The Autism Spectrum Quotient scores and temporal binding window calculations are then compared to participants' results on the Visual-Tactile Integration task to investigate potential correlations which may be predictable based on results of previous multimodal perception research on children with and without ASD. Firstly, it will examine whether there is a relationship between the Autism Spectrum Quotient scores and participants' rates of integration in the Visual-Tactile Integration task. Secondly, participants' temporal binding window widths will be assessed as predictors of integration rates of visual-tactile stimuli. Finally, Autism Spectrum Quotient scores will be compared to temporal binding window widths to explore possible correlations between these two factors. If relationships between the three factors are found, this would add support to notions that the multimodal perceptual behaviour of child ASD populations is just one end of a continuum on which the perceptual behaviour of TD adults also lies. Links between these factors may also suggest that traits associated with ASD may in some way be related to temporal acuity and the ability to integrate multimodal speech stimuli.

2. Visual-Tactile Integration

2.1 Overview

The present study aims to investigate the effects of aero-tactile information on visual speech perception. Previous research has shown that English bilabial stops can be considered visemes (Fisher, 1968), therefore without the addition of any other speech information, they should not be distinguishable solely through the visual modality (though see an alternative analysis presented in Abel et al. (2011) which suggests visual differences are present). In the audio-visual modality combination, the McGurk effect (McGurk & MacDonald, 1976) has shown that speech information from two separate modalities may be combined to form an integrated percept. This perceptual integration is maintained over a temporal window of slightly asynchronous stimuli (Munhall et al., 1996; van Wassenhove et al., 2007). Perceptual integration has also been found to occur in the audio-tactile modalities (Gick & Derrick, 2009) and is maintained over a similar temporal window (Gick et al., 2010). These windows of multimodal integration are asymmetric such that the visual-leading stimuli is integrated over a greater range of asynchronies than audioleading stimuli (in audio-visual combinations) and audio-leading stimuli is integrated over a greater range of asynchronies than tactile-leading stimuli (in audio-tactile combinations). This is due to the relative speeds of these properties, i.e. light travels faster than sound, which travels faster than airflow. Therefore, when perceiving a single event, individuals may detect information from different modalities at slightly different times, however, through experience they become accustomed to perceiving slightly asynchronous multimodal stimuli from one event as synchronous. Based on prior research into multimodal integration (discussed in detail in section 1.1), and extending these findings to visual-tactile modality combinations, the following hypotheses and predictions are made in section 2.1.1.

2.1.1 Hypothesis and Predictions

The present study hypothesizes that tactile speech information contributes to the speech signal in a comparable manner to that of auditory and visual speech information, so that this information is integrated as part of an overall speech percept. The temporal relationship between tactile stimuli and stimuli from different modalities does not need to be completely synchronous for this integration to occur, with integration commonly occurring over a temporal window of asynchronies. This window of multimodal integration is asymmetric due to the differing speeds of physical properties of the world. Considering visual and tactile stimuli, the speed of light is faster than the speed of airflow, therefore visual-tactile speech stimuli in which the visual stimulus precedes the tactile stimulus is more likely to be integrated that when the opposite order occurs. Regarding the current experimental study, the following predictions are made based on these hypotheses:

Following the notion that /b/ and /p/ are considered to be visemes (Fisher, 1968), it is firstly predicted that:

<u>Prediction 2.1</u>: Participants will perform at chance levels when presented with visual-only /pa/ and /ba/ stimuli.

Previous research into audio-visual (McGurk & MacDonald, 1976) and audio-tactile (Gick & Derrick, 2009) speech perception has demonstrated perceptual integration of speech information from each modality. Assuming that this is evidence for multimodal speech rather than bimodal speech specific to audio-visual and audio-tactile combinations, a second and primary prediction is the following:

<u>Prediction 2.2</u>: Participants will give more /pa/ responses when presented with synchronous visual-tactile stimuli than when visual-only stimuli are presented. Increased /pa/ responses would indicate that participants are integrating the tactile stimuli as perceived aspiration.

Further, previous research into both audio-visual (Munhall et al., 1996; van Wassenhove et al., 2007) and audio-tactile (Gick et al., 2010) speech perception has found that there is a window of asynchronies over which integration is maintained, with stimuli more likely to be integrated when the asynchrony is closer to 0ms. A third prediction is therefore:

<u>Prediction 2.3a</u>: Participants will give more /pa/ responses when the Stimulus Onset Asynchrony (SOA) is smaller (closer to synchronous in either direction). Increased /pa/ responses would again indicate that participants are integrating the tactile stimulus as perceived aspiration.

Based on findings related to the physical properties of the natural world in which the relative speeds of various modes of sensory information differ (Munhall et al., 1996; Gick et al., 2010), a caveat to the previous prediction (2.3a) is:

<u>Prediction 2.3b</u>: There will be an asymmetry in responses in that increased /pa/ responses to visual-lead stimuli will be sustained over greater SOAs than increased /pa/ responses to tactile-lead stimuli. This is because the speed of light is faster than the speed of airflow during speech.

2.2 Methods

2.2.1 Participants

Fifty-five speakers (University of British Columbia students) took part in the study and received course credit for their participation. As a result of the recruitment strategy, no restrictions could be placed on native speaker requirements, or whether participants had prior knowledge of the study. Due to the focus of the main task being dependent on a phonological contrast found in English but not necessarily present in other languages¹, the results of twenty-three non-native English speakers' data were not analysed in the present study. Of the thirty-two remaining native English speakers, one was excluded due to experimenter error (forgetting to turn on the babble audio) and five were excluded because they had knowledge of the study's purpose. Of the remaining twenty-six participants, the age range was 18 - 40 yrs, M = 21.23, SD = 4.67 yrs, with 19 females. Participants gave informed consent and reported no history of speech or hearing issues.

¹ Although the same contrast is present in many other languages its dimensions, especially Voice Onset Time, may differ and it was important to control for this factor.

2.2.2 Procedure

Each participant completed a Visual-Tactile Integration task, Simultaneity Judgment task, an Autism Spectrum Quotient (AQ) questionnaire and a language background questionnaire, in that order. The order of the tasks remained constant to avoid directing attention to the air puff during the Visual-Tactile Integration task, as the Simultaneity Judgment task contained instructions specifically related to the air puff. All tasks were administered using PsychoPy (Peirce, 2007) experimental software. Participants were seated in a sound-attenuated booth, with their heads positioned against a headrest to prevent excessive movement. An air tube was positioned ~7cm from the suprasternal notch (front of neck) of each participant. During pre-task instructions, participants were told they would feel puffs of air on their skin at some point during the experiment. Throughout both tasks involving the air puff, participants wore Direct Sound Extreme Isolation noise cancelling headphones through which they heard continuous English, multi-talker babble. This was to mask any sound coming from the air tube when the air puff was released and to create a more natural speech environment where the utterances seemed to be inaudible due to the babble, rather than being silent speech. The Visual-Tactile Integration task is described in section 2.2.2.1 below. The Simultaneity Judgment task and AQ questionnaire are discussed in sections 3.2.1 and 3.2.2.

2.2.2.1 Visual-Tactile Integration Task

The Visual-Tactile Integration task was a two-alternative forced-choice response task where participants watched silent videos of a person saying /pa/ or /ba/. Whilst watching the videos, participants received gentle puffs of air to their skin (in all but one condition) and heard English, multi-talker babble. They were instructed to watch the person on the screen speaking and respond via keyboard as to what he had said. The two response key options were the *z* and *slash* keys which were labelled "pa" and "ba". Response keys were counterbalanced across participants. Participants' responses triggered the next trial to automatically appear on screen. Participants completed four practice trials before the task began and no feedback was given. During the task, ten conditions were presented: a visual-only condition and the following SOA

conditions: 0ms, ± 50 ms, ± 100 ms, ± 200 ms, ± 300 ms, where "+" means that the visual stimulus precedes the tactile stimulus and "–" means the tactile stimulus precedes the visual. Each condition was presented ten times with both /pa/ and /ba/ visual stimuli for a total of 200 tokens which were completely randomized. The task took less than 15 minutes to complete.

2.2.3 Stimuli

Participants were presented with visual-only and paired visual and tactile stimuli. The visual stimuli consisted of videos of a talking face which had the original audio track removed. The tactile stimuli consisted of puffs of air to the skin. The presentation and timing of these stimuli was coordinated via a specially designed switchbox, which caused the release of the air puffs when a 10 kHz, 1db sinewave of 30ms duration was detected. This sinewave was added as an audio track to the video file and when the video file was played this audio signal was directed via audio cable to the switchbox. It was therefore inaudible to the participant.

2.2.3.1 Visual Stimuli

A 28 year old male native speaker of Vancouver English was instructed to produce eight repetitions of /pa/ and /ba/ in isolation, speaking naturally. The productions were recorded on a JVC camcorder (model GZ-E300AU), 48 kHz stereo PCM audio, 24 frames/second video, and 1280x720 pixel resolution. Editing proceeded in Adobe Premiere ProCC. Five productions of each syllable (/pa/ and /ba/) were chosen based on neutral facial expression, naturalness and consistency of production (as judged by the researcher, a native speaker of Australian English). Each production was cut into its own video file and trimmed to 1800ms so that the duration of each video was consistent.

2.2.3.2 Tactile Stimuli

The audio track from each video was extracted for analysis in Praat (version 5.4.08; Boersma & Weenink, 2015) and removed from the video files. In preparation for creating the tactile stimulus (air puff), the moment of the vowel onset for each production of /pa/ and /ba/ and the burst only for each /pa/ production were determined in Praat. The vowel onset was judged as the onset of the periodic portion of the waveform following the release of the stop. The burst for /pa/ productions was determined as the first spike in the waveform after the initial period of silence. The vowel onset was then subtracted from the burst for each production to determine the average Voice Onset Time (VOT) for the speaker's voiceless bilabial stop (M = 98.97ms, SD = 3.69). Average VOT for /ba/ syllables was ~10ms (M = 9.83ms, SD = 0.54ms), but the VOT for /ba/ was not considered in the creation of the tactile stimuli as the air puff was intended to be modelled on the speaker's aspiration in voiceless stops, simulating as closely as possible the same duration, and timing with respect to the vowel onset.

Tactile stimuli were produced via a system based on Gick & Derrick $(2009)^2$. A Jobmate air compressor set to ~6psi was connected via a 1/4 inch vinyl tube to a custom-made switchbox which housed a solenoid valve. This equipment was situated outside the sound-attenuated booth in which the tasks took place. A second 1/4 inch vinyl tube ran from the switchbox, through the wall of the sound booth and was attached at the other end to a microphone stand with a flexible head. This end released the puff of air which was directed towards the participant. The microphone stand was placed to the left of the participant and the end of the vinyl tube was positioned ~7cm from the suprasternal notch of each participant. This is illustrated in Figure 1. An audio cable ran from the computer which played the visual stimuli to the switchbox. When each stimulus was presented, the sine wave from the video file was detected by the switchbox and triggered the switch, which in turn triggered the solenoid valve to open. This process took ~45ms and resulted in a gentle puff of air being released towards the suprasternal notch of the participant. Due to the time it took for the solenoid valve to close again after opening, a 30ms sine wave caused an air puff of 100ms duration.

² The system for the present study differed in that the former focused on audio-tactile stimuli and used a different switchbox. Different experimental software was also used for stimulus presentation.



Figure 1: Air puff set up. A 1/4 inch vinyl tube is attached to a flexible microphone stand and positioned ~7cm in front of the suprasternal notch (neck) of the participant

2.2.3.3 Coordination of Stimuli

To create the synchronous stimuli, the sine wave was positioned so that its onset occurred 100ms prior to the vowel onset (determined from the original speech audio) of each production, both /pa/ and /ba/. It was then shifted another 45ms to the left to account for the switchbox system latency, therefore in total the sine wave onset occurred 145ms before the vowel onset in each production, as illustrated below in Figure 2(b). This resulted in the onset of the air puff occurring 100ms prior to the original vowel onset and ending at the vowel onset, which simulated the timing of the original period of aspiration in the speaker's aspirated syllables. Due to the difference in VOT between /pa/ and /ba/, this meant that the puff was differentially aligned for the unaspirated syllables as compared to aspirated syllables. For example, when the sine wave was positioned for the 100ms SOA condition, it was actually closer to being synchronous with the burst for /ba/ productions; the /ba/ bursts occurred ~10ms before the vowel onset so the difference in alignment between the sine wave onset and /ba/ burst onset was only ~10ms, whereas in the synchronous condition, the sine wave onset occurs ~90ms before the burst for /ba/. It should however be noted that visually the mouth is not always clearly open at the moment of

the burst and so the visual cues may differ slightly to the (absent) audio cues. Nevertheless, it was important to maintain consistency of puff duration and onset position across syllables, therefore the voiceless, aspirated syllable was chosen as the model for simulating aspiration. Appropriate adjustments for the position of the sine wave were made for the various SOA conditions. For example, in the -200ms condition, the onset of the sine wave was positioned 345ms prior to the vowel onset (see Figure 2(a)), whereas the onset of the sine wave for the 200ms condition was positioned 55ms following the vowel onset (see Figure 2(c)). In total, 100 stimuli were created for the Visual-Tactile Integration task (which were repeated once during the task for a total of 200 tokens), with the five original video productions for both /ba/ and /pa/ being used to create stimuli for each of the nine SOA conditions, as well as the visual-only condition (which did not have a sinewave added). Two extra SOAs (\pm 150ms and \pm 250ms) were created for the Simultaneity Judgment task.



Figure 2: Position of sine wave relative to vowel onset of a /pa/ production (a) -200ms, (b) 0ms condition and (c) 200ms condition. Note that the audio tracks these images are based on are used only for illustrative purposes. All original speech audio was removed from the experiment stimuli.

2.3 Analysis and Results

2.3.1 Visual-only Condition

Based on the idea that /p/ and /b/ are generally considered to be visemes, prediction 2.1 stated that participants would perform at chance levels in the visual-only condition. Figure 3 illustrates the percentage of /pa/ and /ba/ responses in this condition.



Figure 3: Percentage of /ba/ and /pa/ responses in the visual-only condition

If participants perform at chance levels, they should show a 50% rate of /pa/ and /ba/ responses. As can be seen, the rate for /pa/ responses is lower (34%). This is a significant deviation from chance, as shown by a binomial test, p = <0.001, 95% CI [0.61, 0.70], thus demonstrating that participants exhibit a /ba/ bias in the visual-only condition.

2.3.2 Synchronous Condition

The introduction of the tactile stimulus in the synchronous puff condition was predicted to produce an increase in /pa/ responses, as compared with the visual-only condition (prediction 2.2). Figure 4 illustrates the total percentage of /pa/ responses in the visual-only and synchronous puff (Oms) conditions. The figure shows that in the visual-only condition, /pa/ responses to the stimuli are 34%, whereas in the synchronous puff condition /pa/ responses increase to 59%.



Figure 4: Percentage of /pa/ responses in the 0ms and visual-only conditions

To investigate whether the percentage of /pa/ responses differs significantly between the visualonly and synchronous conditions, a logistic mixed effects model was fit using the glmer function (Bates, Maechler, Bolker, & Walker, 2014) in R (R Core Team, 2014) with response as the dependent variable, visual stimulus type (/pa/ or /ba/) and puff condition (visual-only, synchronous) and their interaction as fixed effects, a random effect for participant and a byparticipant random slope for the interaction of visual stimulus type and puff condition, and a second random slope for each visual token. The model structure as a formula is as follows:

Response ~ *Visual stimulus* * *Puff condition* + (1 + (*Visual stimulus* * *Puff condition*)/*Participant*) + (1/*Token*)

Results (see Table 1) show a significant main effect of intercept ($\beta = -0.73$, SE = 0.30, z = -2.47, p = <0.05), and puff condition ($\beta = 0.95$, SE = 0.42, z = 2.27, p = <0.05), indicating that participants report significantly more /pa/ responses during trials where there is a puff of air presented synchronously with the visual stimulus.

 Table 1: Results of mixed model with structure: Response ~ Visual stimulus * Puff condition

 + (1 + (Visual stimulus * Puff condition)/Participant) + (1/Token)

	β	SE	Z	p
Intercept	-0.73	0.30	-2.47	<0.05 *
Visual stimulus (pa)	-0.08	0.37	-0.20	0.84
Puff condition (0ms)	0.95	0.42	2.27	<0.05 *
Visual stimulus (pa): Puff condition (0ms)	0.55	0.52	1.06	0.29

2.3.3 Asynchronous Conditions and Asymmetry

Prediction 2.3a stated that when the SOA was smaller (closer to synchronous), participants would report more /pa/ responses. Related to this, prediction 2.3b asserted that the /pa/ responses would be asymmetric; that is, participants would be more likely to perceive a token as /pa/ when the visual stimulus preceded the audio, than when the opposite order occurred. Figure 5 shows the mean percentage of /pa/ responses as a function of the SOA and visual stimulus. As illustrated, when the tactile stimulus leads the visual by 300ms (-300ms SOA) /pa/ responses are around 40% for both /pa/ and /ba/ visual stimuli, with /pa/ visual stimuli having slightly higher

percentage /pa/ responses (42%) as compared to /ba/ visual stimuli (37%). As the SOA progresses towards synchronous (0ms SOA), there is a rise in /pa/ responses which reaches around 60% at 0ms SOA (64% when the visual stimulus is /pa/ and 55% when the visual stimulus is /ba/). As can be seen, there is a clear asymmetry in responses. The figure also shows that the highest rate of /pa/ responses does not occur when the two stimuli are synchronously presented. When the visual stimulus is /pa/, the highest /pa/ response (65%) is at 50ms SOA and when the visual stimulus is /ba/ the highest /pa/ response (67%) is further rightwards, at 200ms SOA. The rate of /pa/ responses can be seen to drop off after the respective peaks.



Figure 5: Percentage of /pa/ responses across Stimulus Onset Asynchronies, split by visual stimulus

To investigate whether the degree of asynchrony of the SOAs significantly affects responses and whether this happens in a symmetrical fashion, a logistic mixed effect model was fit with response as the dependent variable, visual stimulus and SOA (0ms, \pm 50ms, \pm 100ms, \pm 200ms, \pm 300ms) and their interaction as fixed effects (SOAs were converted from factors to a scale as

the SOAs were a related continuum), a random effect for participant and a by-participant random slope for the interaction of visual stimulus and SOA, with a second random slope accounting for each visual token. The model structure as a formula is as follows:

Response ~ *Visual stimulus* * *SOA* + (1 + *Visual stimulus* * *SOA*/*Participant*) + (1/*Token*)

This model (see Table 2) shows a significant main effect of intercept ($\beta = 0.27$, SE = 0.12, z = 2.20, $p = \langle 0.05 \rangle$ and SOA ($\beta = 0.42$, SE = 0.12, z = 3.40, $p = \langle 0.001 \rangle$, indicating that as the SOA increases, there is also an increase in /pa/ responses.

 Table 2: Results of mixed model with structure: Response ~ Visual stimulus * SOA + (1 + Visual stimulus * SOA/Participant) + (1/Token)

	β	SE	z	р
Intercept	0.27	0.12	2.20	<0.05 *
Visual stimulus (pa)	-0.05	0.15	-0.32	0.75
Stimulus Onset Asynchrony (SOA)	0.42	0.12	3.40	< 0.001 ***
Visual stimulus (pa): SOA	-0.22	0.14	-1.52	0.13

A polynomial model with the structure described below was run to investigate whether the various SOAs significantly affected responses in symmetrical fashion.

Response ~ Visual stimulus *Poly(SOA, degree = 2) + (1/Participant) + (1/Token)

Results are shown in Table 3 below:

	β	SE	Z.	р
Intercept	0.24	0.11	2.11	<0.05 *
Visual stimulus (pa)	-0.03	0.12	-0.24	0.81
Poly(SOA, degree = 2)1	25.84	6.18	4.18	<0.001 ***
Poly(SOA, degree = 2) 2	-9.80	6.59	-1.49	0.14
Visual stimulus (pa) : Poly(SOA, degree = 2)1	-12.71	8.93	-1.42	0.15
Visual stimulus (pa) : Poly(SOA, degree = 2)2	-8.13	9.45	-0.86	0.39

Table 3: Results of mixed model with structure: Response ~ Visual stimulus *Poly(SOA, degree = 2) + (1/Participant) + (1/Token)

The results show that like the linear model, the 2nd order polynomial model demonstrates a significant main effect of the intercept and SOA.

To test whether responses at each individual SOA differ significantly from the visual-only condition a model was run with fixed effects for SOA, visual stimulus and their interaction, as well as a random variable for participant and token. SOAs were converted to factors as this allowed for significance testing at each SOA. That model failed to converge. The model was simplified until the only remaining random variable was the one for token.

Response ~ (Visual Stimulus * SOA) + (1/Token)

Results (see Table 4) show that the temporal window of visual-tactile integration is asymmetric, with significantly higher /pa/ responses from -100ms to 300ms, compared to the visual-only condition.

	β	SE	z	р
Intercept	-0.64	0.24	-2.71	<0.01 **
Visual stimulus (pa)	-0.07	0.34	-0.24	0.81
SOA -300ms	0.11	0.34	0.31	0.75
SOA -200ms	0.41	0.34	1.21	0.22
SOA -100ms	0.66	0.34	1.97	<0.05 *
SOA -50ms	0.91	0.34	2.71	<0.01 **
SOA 0ms	0.85	0.33	2.56	<0.05 *
SOA 50ms	1.12	0.34	3.35	< 0.001 ***
SOA 100ms	1.34	0.34	3.98	<0.001 ***
SOA 200ms	1.39	0.34	4.09	< 0.001 ***
SOA 300ms	1.13	0.34	3.37	<0.001 ***
Visual stimulus (pa): SOA -300ms	0.31	0.47	0.64	0.52
Visual stimulus (pa): SOA -200ms	0.05	0.47	0.11	0.92
Visual stimulus (pa): SOA -100ms	0.08	0.47	0.18	0.86
Visual stimulus (pa): SOA -50ms	0.05	0.47	0.10	0.92
Visual stimulus (pa): SOA 0ms	0.49	0.47	1.00	0.32
Visual stimulus (pa): SOA 50ms	0.25	0.48	0.53	0.60
Visual stimulus (pa): SOA 100ms	-0.06	0.48	-0.13	0.89
Visual stimulus (pa): SOA 200ms	-0.20	0.48	-0.43	0.67
Visual stimulus (pa): SOA 300ms	-0.48	0.47	-0.99	0.32

 Table 4: Results of mixed model with structure: Response ~ (Visual Stimulus * SOA) + (1/Token)

2.4 Discussion

This study examined the potential integration of aero-tactile speech information during visualtactile speech perception. Based on previous research into audio-visual and audio-tactile speech perception, several predictions were made with respect to participants' perceptual behaviour when presented with visual-tactile speech. Prediction 2.1 assumed that participants would perform at chance levels when presented with visual-only bilabial syllables, but found that when presented with visual-only /pa/ or /ba/, they were significantly more likely to report the unaspirated syllable, counter to the prediction. This was an unexpected result as the speech sounds $\frac{p}{and}$ and $\frac{p}{and}$ are generally considered to be visemes (Fisher, 1968) and also considering that in a similar perception study, Abel et al. (2011) found a response bias against /b/ (however a three-way versus two-way identification task may not be a legitimate comparison). It is possible that the /ba/ bias in the present study was a result of the weighting of conditions in the task; nine out of ten conditions involved presentation of both visual and tactile stimuli, whereas only one condition lacked the tactile stimulus (recall that tokens were completely randomized so the /ba/ bias would not have been the result of an ordering effect). Arguably over the course of the experiment participants became more accustomed to perceiving tokens including the tactile stimulus as /pa/ and therefore were more inclined to report a /ba/ when the token lacked the puff of air. In any case, it must be acknowledged that this result of a /ba/ bias in the visual-only condition then became the baseline for comparison with all of the other conditions containing the tactile stimulus. This obviously affected all subsequent analysis as this imbalance in the baseline then limited the statistical power of the other results.

A major purpose of this study was to examine whether people integrate aero-tactile information during visual-tactile speech perception in such a way that it affects their categorization. A similar finding has been made for audio-visual (McGurk & MacDonald 1976) and audio-tactile (Gick & Derrick, 2009) modality combinations. The results comparing the visual-only and synchronous conditions supported the major prediction (2.2), demonstrating that the puff of air significantly affected the perception of the visual stimulus so that in the synchronous puff condition, participants were more likely to perceive a voiceless aspirated stop (as opposed to a voiced stop)
than in the conditions where the visual stimulus was presented in isolation. As noted above however, one limitation of this finding was that participants did not perform at chance levels in the visual-only condition, therefore the significance of the effect in the synchronous puff condition could have been driven by the bias against /pa/ in the visual-only condition. Nevertheless, these findings appeared to show that people integrate this aero-tactile information as the sensation of aspiration. This suggests that, whether speakers are consciously aware of it or not, the aero-tactile mode provides information to the speech stream and that this information is integrated with information from other modalities to contribute to the overall perception of the speech signal. This result also showed that speech integration can occur without participants receiving any information from the original audio speech signal. This was an informative result considering the audio signal is generally considered to be the primary source of speech information. The findings from this study also show that combinations of visual and aero-tactile speech information contribute enough to the speech stream to affect perceivers' categorization of speech segments. Admittedly, participants did receive auditory information from the multi-talker babble. Though this auditory information was unrelated to the target speech stimuli, information from this signal could have potentially been extracted and affected participants' perception. White noise rather than babble may have lessened any of this type of effect as it lacks real speech information, however the decision to use a babble track was based on the desire to create a relatively natural speech environment in which utterances seemed inaudible because of crowd noise.

This study also examined whether the degree of asynchrony of the stimuli affected participants' responses. Based on previous research (Munhall et al., 1996; Gick et al., 2010), it was predicted that the closer the SOA was to synchronous, the more likely participants would be to consider the visual and tactile stimulus combination as being part of the same event, and therefore the more likely they would be to integrate the tactile stimulus and report a /pa/ (prediction 2.3a). Figure 5 showed that at -300ms where the tactile stimulus leads the visual stimulus, participants reported perceiving /pa/ at around 40% and this rate grew steadily as the SOA grew closer to synchronous, peaking at 50ms when the visual stimulus was /pa/ and 200ms when the visual stimulus was /ba/. The difference in responses when the visual stimulus was /pa/ compared to /ba/ (though not significant based on statistical tests) may have been due to the position and duration of the tactile stimulus, which as previously discussed in section 2.2.3.3 was designed to

28

be aligned with /pa/ visual stimuli, but not /ba/ stimuli which has a shorter VOT. The /pa/ responses declined again towards 300ms, though they were noticeably higher at 300ms than -300ms. Participant responses at SOAs of -100ms through to 300ms showed a significant difference from responses in the visual-only condition, with the strength of the significance peaking at 100 – 200ms. These findings showed that although /pa/ responses increased towards the synchronous condition, they peaked later than this and this is interpreted as a partial validation of prediction 2.3a. These findings also suggest a group window of visual-tactile integration of -100ms to 300ms, slightly wider to that found for audio-visual stimuli (0ms to 180ms for Munhall et al., 1996 and -30ms to 170ms for van Wassenhove et al., 2007) and audiotactile stimuli (-50ms to 200ms for Gick et al., 2010) but the same in terms of the direction of asymmetry, as discussed below.

Prediction 2.3b asserted that participants would integrate speech information from the visual and tactile modalities over a greater range of SOAs when the visual stimulus preceded the tactile stimulus, as opposed to when the stimuli were presented in the opposite order. This prediction was based on previous research which found asymmetry in the windows of multimodal integration (Munhall et al., 1996; Gick et al., 2010), as well as on knowledge of physical properties of the world such as speeds of light and speech airflow. Results showed that in support of this prediction, participants were more likely to report the voiceless aspirated stop when presented with visual-leading tokens. These results were seen in Figure 5 where the rate of /pa/ responses (to both /pa/ and /ba/ visual stimuli) was generally higher when the visual stimulus preceded the tactile stimulus. This suggests that, similarly to perception of audio-visual and audio-tactile speech stimuli, individuals perceive visual-tactile speech stimuli in a manner consistent with relative speeds of physical properties of the world. The range of integration on the right side is noticeably wider than was found for audio-visual and audio-tactile combinations and this may be due to the relatively slow transmission speed of speech airflow as compared to the speed of light. The perceptual behaviour observed for the asynchronous stimuli was apparent even though speakers would not have experience perceiving aspiration from a great distance as airflow from aspiration is found to dissipate at around 30-40cm from the mouth (Derrick, Anderson, Gick, & Green, 2009). Aspiration airflow is known to be delayed by 25ms at 17cm and by 100ms at between 30 and 35cm distance (Derrick et al., 2009), so speakers may have some experience perceiving slightly delayed aspiration with respect to the rest of the speech

signal at 50ms and possibly 100ms, two of the SOAs tested. Further, aspiration is often audible so speakers could plausibly gain knowledge about aspiration onset timing with respect to the rest of the speech signal through auditory aspiration cues. Speakers also arguably experience delays in perceiving aero-tactile stimuli, as compared to visual stimuli from the same event, from other non-speech sources, e.g. feeling the delayed air from an oscillating fan on the other side of a room when the fan has already turned in another direction. This more general airflow information may contribute to speakers' knowledge of how airflow from aspiration behaves.

Regarding speaker knowledge of aspiration, based on various communicative experiences speakers arguably do have considerable understanding of aspiration behaviour. Several communicative situations provide direct tactile stimuli in the form of aspiration. Whispering is one example; in this situation the listener receives strong aspiration information to the side of the face. Simply speaking in close proximity to a conversation partner may also be a situation where aspiration can be felt. Other situations provide a visual representation of aero-tactile speech information which can be incorporated as general knowledge of aspiration, as was shown by Mayer et al., (2013) with the visible perturbation of a candle. Further examples include speaking in very cold temperatures, where speakers produce visible puffs of air, and speaking while smoking. The use of microphones is a situation in which aspiration is represented audibly. All of these experiences may contribute to speakers' awareness of the behaviour of aspiration. Speakers may also receive aero-tactile feedback to their own lips when they speak. Results of the current study suggest that naïve speakers can map this aspiration information to particular speech segments in their language.

Findings from the present study have shown that speakers have either conscious or unconscious awareness of aero-tactile speech information, an underexplored area in speech research. Speakers have shown the ability to, in the absence of the original audio speech signal, make use of information from the aero-tactile mode to distinguish speech sounds when they are presented with an ambiguous visual speech signal. This ability shows that aero-tactile speech information is influential in multimodal speech integration and contributes enough information to the signal to shift categorization of speech segments. This finding also suggests that speech may be perceived as a modality neutral event in which signals from all relevant modes are perceived

30

equally (though the weighting of speech relevant information may differ), rather than it being perceived primarily as an auditory event.

The current chapter has dealt with group behaviour in visual-tactile speech perception. The following chapter will explore individual differences in perceptual behaviour of visual-tactile speech related to temporal acuity and behavioural traits associated with ASD.

3. Individual Differences in Integration Based on Behavioural Attributes and Temporal Binding Windows

3.1 Overview

Chapter 2 was concerned with integration of visual-tactile speech stimuli and examined participant behaviour as a group. The current chapter explores individual differences in integration to test whether rates of integration can be predicted by two factors: Autism Spectrum Quotient (AQ) scores and widths of participants' temporal binding windows. Children with ASD have been found to integrate synchronous audio-visual stimuli less than TD children (Gelder et al., 1991; Irwin et al., 2011; Mongillo, et al., 2008; Stevenson, et al., 2014). Variation in AQ and AQ subfactor scores has been reported amongst TD populations (Baron-Cohen et al., 2001) and found to influence speech production (Yu et al., 2013), speech perception (Yu, 2010) and nonspeech audio-visual perception (Donohue et al., 2012). Temporal binding window width has also been shown to be wider in child ASD populations as compared to TD groups (Foss-Feig et al., 2010; Kwakye et al., 2011; Stevenson, et al., 2014), vary within adult TD populations (Miller & D'esposito, 2005; Stevenson et al., 2012), and influence rates of multimodal integration in child ASD (Stevenson, et al., 2014) and TD adult (Stevenson et al., 2012) populations. Based on this previous research (discussed in detail in section 1.2) and extending these findings to the visualtactile modality combination and a TD adult population, the following hypotheses and predictions are made in section 3.1.1.

3.1.1 Hypothesis and Predictions

The present study hypothesizes that TD individuals vary with regard to their rates of integration during visual-tactile speech perception. This variation can be connected with levels of behavioural traits (social skills, communication skills, imagination, attention to detail and attention switching/tolerance of change) associated with ASD, a population which in child studies has demonstrated significantly lower rates of audio-visual speech integration as

compared to TD children. Differences in integration behaviour are proposed to vary along a continuum from child ASD to adult TD populations and be linked to these behavioural traits associated with ASD. TD adults who exhibit levels of behavioural traits more similar to those of individuals with an ASD diagnosis will exhibit perceptual behaviour reflective of that previously found for child ASD populations (assessed on audio-visual integration), that is, they will demonstrate lower rates of visual-tactile speech integration. As children with ASD have also been shown to exhibit poorer temporal acuity than TD children, individuals with levels of traits more similar to those of individuals with ASD will also integrate across a greater range of SOAs when stimuli are asynchronous.

Individual variation in multimodal speech integration can also be connected to an individual's temporal binding window width. The width of an individual's temporal binding window, particularly the full and right-side window, will be related to their rates of visual-tactile speech integration in that individuals with wider temporal binding windows are expected to exhibit lower rates of visual-tactile speech integration when presented with synchronous stimuli, as well as integration across a greater range of SOAs when stimuli are asynchronous. Further, as children with ASD have been shown to have wider temporal binding windows, temporal binding window width and behavioural traits associated with ASD will also be correlated in that individuals with wider windows will exhibit levels of behavioural traits more similar to those of individuals with ASD. Regarding the current experimental study and based on these hypotheses, the following predictions are made:

It has been demonstrated that children with ASD integrate synchronous audio-visual speech stimuli less than their TD age-matched counterparts (Gelder et al., 1991; Irwin et al., 2011; Mongillo, et al., 2008; Stevenson, et al., 2014). Extending these results to the visual-tactile modality combination, and the results from ASD child populations to an adult TD population, it is predicted that:

<u>Prediction 3.1</u>: Rates of integration of visual-tactile stimuli from the Visual-Tactile Integration task will be predictable from participants' scores on the AQ. A higher AQ score will be correlated with a lower rate of integration when stimuli are synchronous.

Research into ASD and TD child populations has found evidence that children with ASD have poorer temporal acuity (Bebko et al., 2006; Foss-Feig, et al., 2010; Kwakye et al., 2011; Stevenson, et al., 2014). This should mean that although overall rates of integration may be lower for participants with higher AQ scores, when they do exhibit integration, the window over which integration is exhibited will be wider than it is for those participants with low AQ scores. This is because visual-tactile stimuli at greater SOAs from 0ms (in either direction) may be more likely to be judged as synchronous by individuals with higher AQ scores, which may then mean that stimuli at these greater SOAs would be more likely to be considered part of the same event, and consequently integrated. Therefore:

<u>Prediction 3.2</u>: A higher AQ score will be correlated with integration which is maintained over a greater range of SOAs, as compared with lower AQ scores.

Widths of temporal binding windows in ASD child populations (Stevenson et al., 2014) and widths of right-side temporal binding windows in TD adult populations (Stevenson et al., 2012) have previously been found to be negatively correlated with rates of integration in audio-visual integration (McGurk) studies using synchronous stimuli. No connections have been made between left-side temporal binding windows and rates of multimodal speech integration. Extending these results to the visual-tactile modality combination, and the results from ASD child populations to an adult TD population, the following prediction is made:

<u>Prediction 3.3</u>: Rates of integration of visual-tactile stimuli from the Visual-Tactile Integration task will be predictable from widths of participants' temporal binding windows. A wider full window and right-side window will be correlated with a lower rate of integration when stimuli are synchronous.

Individuals with poorer temporal acuity (indicated by wider temporal binding windows) are less likely to detect slightly asynchronous stimuli, meaning that for these individuals, stimuli at greater SOAs may be more likely to be considered part of the same event and therefore integrated. It therefore follows that: <u>Prediction 3.4</u>: A wider full and right-side temporal binding window will be correlated with integration which is maintained over a greater range of SOAs, as compared with narrower full and right-side windows.

As mentioned above, children with ASD and individuals with wider temporal binding windows (in child ASD populations) or right-side temporal binding windows (in TD populations) have both been shown to predict lower of rates of synchronous audio-visual integration. It follows that:

<u>Prediction 3.5</u>: Higher AQ scores will be correlated with wider full and right-side temporal binding windows.

3.2 Methods

3.2.1 Simultaneity Judgment Task

Participants completed a Simultaneity Judgment task and the results of this task were used to calculate temporal binding windows for each participant. Methods for the calculation of the windows are discussed in section 3.3.1. Participants watched the same silent videos as those used in the Visual-Tactile Integration task whilst they received puffs of air to their neck and heard English, multi-talker babble. Participants were told they would see a person saying "pa" or "ba" and feel a puff of air on their skin. They were instructed to report whether they thought the puff occurred at the same time as the "p"/ "b", or not, by selecting "same" or "different" labels on the keyboard, positioned on the *a* and *apostrophe* keys. Response keys were counterbalanced across participants. Participant responses triggered the next trial to automatically appear on screen. Participants completed four practice trials before the task began and no feedback was given. The following SOAs were tested: 0ms, \pm 50ms, \pm 100ms, \pm 150ms, \pm 200ms, \pm 250ms, \pm 300ms. The extra SOAs (\pm 150ms and \pm 250ms) were added to increase the accuracy of the temporal binding window widths. Each condition was presented five times with both /pa/ and /ba/ visual stimuli resulting in a total of 130 tokens, which were completely randomized. The task took less than 10 minutes to complete.

3.2.2 Autism Spectrum Quotient Questionnaire

Each participant completed the Autism Spectrum Quotient (AQ) questionnaire (Baron-Cohen et al., 2001) via computer. This questionnaire was originally designed as a screening tool for ASD and has also been used in speech research to investigate individual variation (e.g. Yu, 2010; Yu et al., 2013; Donohue et al, 2012). The results from the questionnaire were used to determine an overall AQ score (out of fifty), as well as scores for the five subfactors (out of ten) of the questionnaire for each participant. The questionnaire is comprised of fifty statements, organised into five subcategories: Social Skill, Communication, Imagination, Attention to Detail, and Attention Switching/Tolerance of Change. An example statement is the following: "I often notice small sounds when others do not". Participants were instructed to read each statement and click on the number corresponding to how much they agreed or disagreed with the statement. Possible responses were: 1 = definitely agree, 2 = slightly agree, 3 = slightly disagree, and 4 = definitely disagree. Participants were not given a time limit to complete the task and were given one example statement before completing the questionnaire. Statements were counterbalanced so that reporting behaviour which signified levels of behavioural traits considered more similar to those of individuals with an ASD diagnosis corresponded to a definitely/slightly agree response for half of the statements and a definitely/slightly disagree response for the other half. Each of this type of response scored one point and was considered to mean poor social skills, poor communication skills, poor imagination, exceptional attention to detail, or poor attentionswitching/strong focus of attention (Baron-Cohen et al., 2001). See Appendix A for the full questionnaire.

3.3 Analysis and Results

3.3.1 Creation of Temporal Binding Windows

Results from the Simultaneity Judgment task were used to create a temporal binding window for each individual participant. The participant's "same" responses at each SOA were plotted as a

percentage in R using ggplot2 (Wickham, 2009) and a smooth line curve was fit to the responses using the *loess* function. The x, y coordinates of the peak of the curve (peak SOA) and the coordinates of the two points at which the line crossed the 75% threshold of participant "same" responses were determined (marked by crosses in Figure 6). The left-side temporal binding window was calculated by subtracting the left-side x, y coordinate from the peak SOA x, y coordinate and the right-side temporal binding window was determined by subtracting the peak SOA x, y coordinate from the right-side x, y coordinate. The left and right-side windows were added together to determine the value of the full temporal binding window. Applying these methods, 9 participants were excluded either because the peak of the curve did not rise above 75% (low "same" responses; 4 participants excluded) or the right-side measurement did not drop back down below 75% (high "same" responses on right-side SOAs; 5 participants excluded). Therefore, analysis involving window calculations included data from only 17 participants.



Figure 6: Percentage of a participant's "same" responses across the Stimulus Onset Asynchronies. Full, right-side and left-side temporal binding windows were created from these results. The figure is marked with crosses indicating the x, y coordinates used to calculate the windows (as described above in section 3.3.1).

3.3.2 Autism Spectrum Quotient Score as a Predictor of Integration Rates

3.3.2.1 Synchronous Condition

Participants' overall scores (out of 50) on the AQ questionnaire range from 9 - 25, M = 17.31, SD = 4.58. Scores for the subfactors (out of 10) are as follows: Social Skill: 0 - 6, M = 2.58, SD = 1.55; Communication: 0 - 6, M = 2.69, SD = 1.57; Imagination: 0 - 4, M = 2.23, SD = 1.37; Attention to Detail: 2 - 8, M = 5.27, SD = 1.95; Attention Switching: 2 - 7, M = 4.54, SD = 1.84. Higher overall and subfactor scores are equated with levels of behavioural traits more similar to those of individuals with an ASD diagnosis. An overall score of 32 or higher is considered a possible indication of symptoms of ASD (Baron-Cohen et al., 2001).

Prediction 3.1 stated that participants who exhibited higher AQ scores would integrate synchronous visual-tactile stimuli at a lower rate than those with lower AQ scores. To test whether AQ scores were a predictor of integration rates of synchronous visual-tactile stimuli, a logistic mixed effect model was run with the following structure:

Response ~ Visual stimulus * Puff condition * AQ + (1 + (Visual stimulus * Puff condition * AQ)/Participant) + (1/Token)

No significant main effects or interactions involving AQ score were found.

The AQ subfactors (Social Skill, Communication, Imagination, Attention to Detail, and Attention Switching/Tolerance of Change) were also investigated as potential predictors of integration of synchronous stimuli. Mixed effects models were run considering each subfactor as a fixed effect separately, with structures fitting this model:

Response ~ Visual stimulus * Puff condition * {subfactor} + (1 + (Visual stimulus + Puff condition + {subfactor})/Participant) + (1\Token)

No significant effects of the subfactors of Communication, Imagination, Attention to detail or Attention switching/Tolerance of change were found. Results of the model considering Social Skill as a fixed effect (see Table 5) show a main effect of Social Skill ($\beta = -0.37$, SE = 0.14, z = -2.68, p = <0.01) such that as Social Skill scores decrease (equalling higher social skill), /pa/

responses increase. There was a two-way interaction between visual stimulus and puff condition ($\beta = 1.38$, SE = 0.69, z = 2.01, $p = \langle 0.05 \rangle$) showing that when the visual stimulus is /pa/, /pa/ responses increase in the synchronous puff condition. A two-way interaction between puff condition and Social Skill ($\beta = 0.46$, SE = 0.18, z = 2.60, $p = \langle 0.01 \rangle$) suggests that a *higher* score for Social Skill (equally poorer social skills) increases /pa/ responses. This result appears to contradict the direction of the main effect of Social Skill. Therefore these results should be treated with caution.

 Table 5: Results of mixed model with structure: Response ~ Visual stimulus * Puff condition

 * Social Skill + (1 + (Visual stimulus + Puff condition + Social Skill)/Participant) + (1\Token)

	β	SE	z	p
Intercept	0.17	0.41	0.41	0.68
Visual stimulus (pa)	-0.56	0.50	-1.10	0.27
Puff condition (0ms)	-0.20	0.61	-0.33	0.74
Social Skill	-0.37	0.14	-2.68	<0.01 **
Visual stimulus (pa): Puff condition (0ms)	1.38	0.69	2.01	<0.05 *
Visual stimulus (pa): Social Skill	0.21	0.14	1.51	0.13
Puff condition (0ms): Social Skill	0.46	0.18	2.60	<0.01 **
Visual stimulus (pa): Puff condition (0ms):	-0.35	0.18	-1.94	< 0.053 .
Social Skill				

A second model was fit with an interaction in the random slope:

Response ~ Visual stimulus * Puff condition * Social Skill + (1 + (Visual stimulus * Puff condition * Social Skill)/Participant) + (1/Token)

In this model the only significant finding was the main effect of Social Skill (see Table 6).

	β	SE	Z.	p
Intercept	0.027	0.44	0.06	0.95
Visual stimulus (pa)	-0.42	0.54	-0.79	0.43
Puff condition (0ms)	0.05	0.70	0.08	0.94
Social Skill	-0.30	0.13	-2.22	<0.05 *
Visual stimulus (pa): Puff condition (0ms)	1.07	0.75	1.42	0.16
Visual stimulus (pa): Social Skill	0.13	0.16	0.83	0.41
Puff condition (0ms): Social Skill	0.35	0.19	1.90	0.06 .
Visual stimulus (pa): Puff condition (0ms): Social Skill	-0.20	0.24	-0.85	0.40

 Table 6: Results of mixed model with structure: Response ~ Visual stimulus * Puff condition

 * Social Skill + (1 + (Visual stimulus * Puff condition * Social Skill)/Participant) + (1/Token)

A likelihood ratio test (ANOVA) comparing the two models was performed with results showing no significant difference between the two models ($\chi^2(26) = 7.40$, p = 0.99).

	df	AIC	BIC	logLik	Deviance	ChiSq	df	р
No interaction in random variable	19	1342.3	1436.3	-652.15	1304.3			
Interaction in random variable	45	1386.9	1609.5	-648.45	1296.9	7.3985	26	0.9999

As discussed below, however, when considering all of the SOAs, the Social Skill subfactor does not prove to be a significant predictor of integration rates.

3.3.2.2 Asynchronous Conditions

To investigate whether higher AQ scores are correlated with integration maintained across a wider range of SOAs (as proposed in prediction 3.2), a mixed effects model with the following structure was run:

Response ~ Visual stimulus * SOA * AQ + (1 + (Visual stimulus + SOA + AQ)/Participant) + (1/Token)

SOA included all Stimulus Onset Asynchrony conditions from the Visual-Tactile Integration task (the visual-only condition was omitted). Results are reported below in Table 7.

	β	SE	z	p
Intercept	0.65	0.38	1.73	0.08 .
Visual stimulus (pa)	0.17	0.42	0.40	0.69
SOA	0.92	0.27	3.38	<0.001 ***
AQ	-0.02	0.02	-1.05	0.29
Visual stimulus (pa): SOA	-0.58	0.28	-2.07	<0.05 *
Visual stimulus (pa): AQ	-0.01	0.02	-0.50	0.62
SOA: AQ	-0.03	0.01	-2.09	<0.05 *
Visual stimulus (pa): SOA: AQ	0.02	0.01	1.59	0.11

Table 7: Results of mixed model with structure: Response ~ Visual stimulus * SOA * AQ +(1 + (Visual stimulus + SOA + AQ) | Participant) + (1|Token)

Results show a main effect of SOA ($\beta = 0.92$, SE = 0.27, z = 3.38, p = <0.001), indicating that as the SOA increases, the rate of /pa/ responses also increases. A two-way interaction between visual stimulus and SOA ($\beta = -0.58$, SE = 0.28, z = -2.07, p = <0.05) was also found, showing that when the visual stimulus is /pa/, this influences the slope of the SOA in a negative direction meaning that rates of /pa/ responses are generally more evenly spread over a wider range of SOAs. There was also a significant two-way interaction between SOA and AQ score ($\beta = -0.03$, SE = 0.01, z = -2.09, p = <0.05) such that an increase in AQ score influences the slope of the SOA in a negative direction, meaning that the rates of /pa/ responses are generally more evenly spread over a wider range of SOAs for participants with higher AQ scores.

A second model which included an interaction in the random slope failed to converge:

Response ~ Visual stimulus * SOA * AQ + (1 + (Visual stimulus * SOA * AQ)/Participant) + (1/Token)

Visualizations of /pa/ responses for participants with low vs, high AQ scores (see Figure 7 below) support the findings of the interaction between AQ score and SOA, illustrating that participants with lower AQ scores (indicating levels of behavioural traits more distant from those of individuals with an ASD diagnosis) have a steeper curve of /pa/ responses, whereas participants with higher AQ scores show a slightly flatter curve across the range of SOAs.



Figure 7: Percentage of /pa/ responses across Stimulus Onset Asynchronies for participants with low vs. high Autism Spectrum Quotient (AQ) scores. The figure on the left shows the curve of responses for participants with low scores on the Autism Spectrum Quotient (scores of 9 - 16) while the figure on the right shows responses for participants with high scores (scores of 17 - 25).

The AQ subfactors (Social Skill, Communication, Imagination, Attention to Detail, and Attention Switching/Tolerance of Change) were also investigated as potential predictors of integration in the asynchronous stimuli conditions (as proposed in prediction 3.2). Mixed effects models were run considering each subfactor as a fixed effect separately, with structures as follows:

Response ~ Visual stimulus * SOA * {subfactor} + (1 + (Visual stimulus + SOA + {subfactor})/Participant) + (1\Token)

No significant effects were found for the subfactors of Communication, Social Skill, Attention to detail or Attention switching/Tolerance of Change.

Results for the subfactor of Imagination show main effects of intercept ($\beta = 0.43$, SE = 0.20, z = 2.14, p = <0.05) and SOA ($\beta = 0.68$, SE = 0.15, z = 4.42, p = <0.001) indicating an increase in /pa/ responses as the SOA increases. A two-way interaction between visual stimulus and SOA ($\beta = -0.54$, SE = 0.17, z = -3.19, p = <0.01) shows that when the visual stimulus is /pa/, the SOA slope decreases, and a two-way interaction between SOA and Imagination ($\beta = -0.13$, SE = 0.05, z = -2.53, p = <0.05) shows a decrease in the SOA slope as scores on the Imagination subfactor increase (indicating poorer imagination). That is, the window of visual-tactile integration is flatter for those participants with poorer imaginations. A three-way interaction between visual stimulus, SOA and Imagination ($\beta = 0.16$, SE = 0.05, z = 3.25, p = <0.01) demonstrates that when the visual stimulus is /pa/, the SOA slope increases as the Imagination subfactor score decreases, meaning that participants with better imaginations give /pa/ responses more often when the visual stimulus is /pa/. Results are shown in Table 8 below.

	β	SE	Z	р
Intercept	0.43	0.20	2.14	<0.05 *
Visual stimulus (pa)	0.20	0.23	0.88	0.38
SOA	0.68	0.15	4.42	<0.001 ***
Imagination	-0.07	0.07	-0.93	0.35
Visual stimulus (pa): SOA	-0.54	0.17	-3.19	<0.01 **
Visual stimulus (pa): Imagination	-0.10	0.07	-1.35	0.18
SOA: Imagination	-0.13	0.05	-2.53	<0.05 *
Visual stimulus (pa): SOA: Imagination	0.16	0.05	3.25	<0.01 **

 Table 8: Results of mixed model with structure: Response ~ Visual stimulus * SOA *

 Imagination + (1 + (Visual stimulus + SOA + Imagination)/Participant) + (1/Token)

Figure 8 shows /pa/ responses across the SOAs for participants with low vs. high scores on the Imagination subfactor, split by visual stimulus. Higher Imagination scores indicate participants with poorer imaginations. As can be seen, participants with lower Imagination scores (from 0 - 2), which equal better imaginations, show more differences in their /pa/ responses depending on the visual stimulus. When the visual stimulus is /pa/, they show a steep curve of /pa/ responses which centres around synchronous, whereas for /ba/ visual stimuli, /pa/ responses peak higher and later. For participants with higher Imagination scores, a steeper curve is still evident when the visual stimulus is /pa/, however, the distinction between responses depending on the visual stimulus is not so great.



Figure 8: Percentage /pa/ responses across Stimulus Onset Asynchronies for participants with low vs. high Imagination scores, split by visual stimulus. The figure on the left illustrates the curve of responses for participants with Imagination scores of 0 - 2, while the figure on the right shows the curve of responses for participants with windows from 3 - 4. Lower Imagination scores are equated with higher levels of imagination

Figure 9(a) illustrates the rate of /pa/ responses across the SOAs as a function of participants' scores on the Imagination subfactor. As can be seen in the figure, participants with the two lowest Imagination scores (indicating better imaginations) generally exhibit the highest rates of

/pa/ responses (especially when the SOA is closer to synchronous), with participants who reported the lowest score (0) in particular also demonstrating the steepest curve across the range of SOAs. Comparatively, participants with poorer imaginations, especially those with the lowest score (4), show flatter overall curves and demonstrate generally lower rates of integration. Figure 9(b) and Figure 9(c) show rates of /pa/ responses when the visual stimulus is /pa/ and /ba/ respectively. As illustrated in Figure 9(b), when the visual stimulus is /pa/, participants exhibit a more reliable pattern of /pa/ responses as a function of their Imagination score, that is, participants with lower Imagination scores (better imaginations) generally give more /pa/ responses than participants with higher Imagination scores (poorer imaginations). Considering the overall degree of SOA curve, there is also an obvious difference between participants with the lowest and highest scores, with participants with the lowest (0) score exhibiting the steepest curve and participants with the highest (4) score demonstrating the flattest overall curve. Comparing this pattern of responses to those of Figure 9(c), it is evident that responses when the visual stimulus is /ba/ are more varied. Participants with the highest Imagination scores still show the flattest SOA curve and participants with the lowest score seem to show the steepest curve (though scores on the right side do not drop off enough for this to be confirmed), however, rates of /pa/ responses as a function of Imagination score do not follow a clear pattern when the visual stimulus is /ba/.



Figure 9: Percentage /pa/ responses across Stimulus Onset Asynchronies as a function of Imagination score. (a) shows /pa/ responses with both visual stimuli grouped. (b) shows /pa/ responses when the visual stimulus is /pa/ and (c) shows /pa/ responses when the visual stimulus is /ba/. Lower Imagination scores are equated with higher levels of imagination

3.3.3 Temporal Binding Window Width as a Predictor of Integration Rates

3.3.3.1 Synchronous Condition

Prediction 3.3 stated that a wider full and right-side temporal binding window would be correlated with lower rates of integration when the stimuli were synchronous. To investigate whether this was the case, three mixed effects models were run to test the full temporal binding window (TBW), right-side temporal binding window (RTBW) and left-side temporal binding window (LTBW). The model structures were as follows:

Response ~ Visual stimulus * Puff condition * TBW + (1 + (Visual stimulus * Puff condition * TBW)/Participant) + (1/Token)

Response ~ Visual stimulus * Puff condition * RTBW + (1 + (Visual stimulus + Puff condition + RTBW)/Participant)

Response ~ Visual stimulus * Puff condition * LTBW + (1 + (Visual stimulus + Puff condition + LTBW)/Participant) + (1/Token)

Full, right-side and left-side TBW values were converted to a scale. Due to convergence issues, the right-side and left-side TBWs did not include an interaction in the random slope and the right-side TBW did not include a random effect for token. Contrary to prediction 3.3, results of the full and right-side TBW models reveal no significant effects of TBW, nor were significant effects found for the left-side TBW model.

3.3.3.2 Asynchronous Conditions

It was also predicted that a wider TBW would be correlated with integration which was maintained over a greater range of SOAs (prediction 3.4). To examine whether participants' full, right-side or left-side TBW widths were accurate predictors of the range of SOAs over which they integrated, three models were fit with the following structure:

Response ~ Visual stimulus * SOA * {window type} + (1+ (Visual stimulus + SOA + {window type})/Participant) + (1/Token)

Results are reported in Table 9, Table 10 and Table 11 respectively.

	β	SE	z	р
Intercept	0.40	0.16	2.57	<0.05 *
Visual stimulus (pa)	-0.10	0.17	-0.60	0.55
SOA	0.48	0.13	3.72	<0.001 ***
TBW	-0.29	0.19	-1.54	0.12
Visual stimulus (pa): SOA	-0.20	0.14	-1.38	0.17
Visual stimulus (pa): TBW	0.19	0.15	1.27	0.20
SOA: TBW	-0.38	0.12	-3.25	<0.01 **
Visual stimulus (pa): SOA: TBW	0.31	0.09	3.64	<0.001 ***

 Table 9: Results of mixed model with structure: Response ~ Visual stimulus * SOA * TBW +

 (1+ (Visual stimulus + SOA + TBW)/Participant) + (1/Token)

Results for the full TBW model (see Table 9 above) show significant main effects of intercept ($\beta = 0.40$, SE = 0.16, z = 2.57, p = <0.05), and SOA ($\beta = 0.48$, SE = 0.13, z = 3.72, p = <0.001) indicating that a higher SOA produces more /pa/ responses. A significant two-way interaction

between SOA and TBW ($\beta = -0.38$, SE = 0.12, z = -3.25, p = <0.01) demonstrates that a wider TBW influences the SOA slope in a negative direction, meaning that participants with wider TBWs exhibit rates of /pa/ responses which are generally more evenly spread over a wider range of SOAs. A three-way interaction between visual stimulus, SOA and TBW ($\beta = 0.31$, SE = 0.09, z = 3.64, p = <0.001) shows that when the visual stimulus is /pa/, the SOA is influenced in a positive direction when the TBW is relatively narrow, meaning that participants with narrower TBWs give more /pa/ responses when the visual stimulus is /pa/.

 Table 10: Results of mixed model with structure: Response ~ Visual stimulus * SOA *

 RTBW + (1+ (Visual stimulus + SOA + RTBW)/Participant) + (1/Token)

	β	SE	Z.	р
Intercept	0.44	0.17	2.53	<0.05 *
Visual stimulus (pa)	-0.11	0.17	-0.62	0.53
SOA	0.48	0.13	3.72	<0.001 ***
RTBW	-0.46	0.18	-2.57	<0.05 *
Visual stimulus (pa): SOA	-0.20	0.14	-1.40	0.16
Visual stimulus (pa): RTBW	0.27	0.16	1.72	0.08 .
SOA: RTBW	-0.37	0.13	-2.81	<0.01 **
Visual stimulus (pa): SOA: RTBW	0.33	0.09	3.79	<0.001 ***

Results for the RTBW model (see Table 10 above) show a significant main effect of intercept ($\beta = 0.44$, SE = 0.17, z = 2.53, p = <0.05), SOA ($\beta = 0.48$, SE = 0.13, z = 3.72, p = <0.001) and RTBW ($\beta = -0.46$, SE = 0.18, z = -2.57, p = <0.05), suggesting that a narrower RTBW is equated with more /pa/ responses. A two-way interaction between SOA and RTBW ($\beta = -0.37$, SE = 0.13, z = -2.81, p = <0.01) and a three-way interaction between visual stimulus, SOA and RTBW ($\beta = -0.37$).

0.33, SE = 0.09, z = 3.79, p = <0.001) are also evident, both in the same direction as observed for the full TBW.

Similarly, results for the LTBW model (see Table 11) show a significant effect of intercept ($\beta = 0.40$, SE = 0.16, z = 2.61, p = <0.01) and SOA ($\beta = 0.47$, SE = 0.13, z = 3.74, p = <0.001), a twoway interaction between SOA and LTBW ($\beta = -0.37$, SE = 0.11, z = -3.46, p = <0.001) and a three-way interaction between visual stimulus, SOA and LTBW ($\beta = 0.30$, SE = 0.09, z = 3.49, p = <0.001) all in the same direction as that observed for TBW and RTBW.

 Table 11: Results of mixed model with structure: Response ~ Visual stimulus * SOA *

 LTBW + (1+ (Visual stimulus + SOA + LTBW)/Participant) + (1/Token)

	β	SE	Z.	p
Intercept	0.40	0.16	2.61	<0.01 **
Visual stimulus (pa)	-0.10	0.17	-0.58	0.56
SOA	0.47	0.13	3.74	<0.001 ***
LTBW	-0.24	0.16	-1.51	0.13
Visual stimulus (pa): SOA	-0.20	0.14	-1.37	0.17
Visual stimulus (pa): LTBW	0.15	0.15	1.01	0.31
SOA: LTBW	-0.37	0.11	-3.46	< 0.001 ***
Visual stimulus (pa): SOA: LTBW	0.30	0.09	3.49	<0.001 ***

Visualizations of /pa/ responses for participants with narrow vs. wide RTBWs, split by visual stimulus (see Figure 10) show that participants with narrower RTBWs make a clear distinction between responses when the visual stimulus is /pa/ vs /ba/. When the visual stimulus is /pa/, they exhibit a curve of /pa/ responses which centres around synchronous whereas responses when the visual stimulus is /ba/ peak higher and later. Comparatively, participants with wider RTBWs

show flatter overall curves of /pa/ responses across the range of SOAs and do not distinguish as much between responses depending on whether the visual stimulus is /pa/ or /ba/.



Figure 10: Percentage of /pa/ responses across Stimulus Onset Asynchronies for participants with narrow vs. wide right-side temporal binding windows, split by visual stimulus. The figure on the left illustrates the curve of responses for participants with right-side temporal binding windows from (68 to 123) while the figure on the right shows the curve of responses for participants with windows from (124 to 168). These groupings contained eight temporal binding window widths in each (two participants had a right-side temporal binding window of the same width).

3.3.4 Relationship Between Autism Spectrum Quotient Score and Temporal Binding Window Width

Prediction 3.5 stated that higher AQ scores would be correlated with wider full and right-side TBWs. To test this prediction, Pearson's product-moment correlations were run with full, right-side and left-side TBWs against the AQ scores. Results reveal weak correlations between AQ

score and TBW, RTBW and LTBW, however, *p*-values do not reach significance³. These results are shown in Table 12 below.

				C	ĊI	
	t	df	r	lower	upper	p
TBW: AQ	1.34	15	0.33	-0.18	0.70	0.20
RTBW: AQ	1.15	15	0.28	-0.23	0.67	0.27
LTBW: AQ	1.47	15	0.35	-0.15	0.71	0.16

 Table 12: Results of Pearson's correlations between width of full, right-side and left-side

 Temporal Binding Window and AQ score

Figure 11 illustrates the correlation between AQ score and TBW, and RTBW. As shown, as AQ scores increase, there is also a slight increase in temporal binding window width, however as *p*-values are not significant, these results should be treated with caution.

³ Further analysis using a larger sample size (by doubling the original data) revealed weak correlations between AQ score and TBW (r(32) = 0.33, p = 0.059), RTBW (r(32) = 0.28, p = 0.10) and LTBW (r(32) = 0.35, p = 0.039), with *p*-values reaching, or trending towards significance, suggesting that a larger sample size may result in more reliable correlations.



Figure 11: Correlations between AQ score and full, and right-side temporal binding windows. Note that *p*-values are not significant.

Pearson's correlations were also run between the AQ subfactors and full, right-side and left-side TBWs. Results are reported in Table 13 below. The subfactor of Imagination shows a strong positive correlation with the TBW (r(15) = 0.56, p = <0.05), RTBW (r(15) = .52, p = <0.05) and LTBW (r(15) = 0.59, p = <0.05), revealing that as participants scores for Imagination increase (indicating poorer imagination), TBWs of all three types become wider. Likewise, the Communication subfactor shows strong correlations for TBW (r(15) = 0.52, p = <0.05), RTBW (r(15) = 0.49, p = <0.05) and LTBW (r(15) = 0.54, p = <0.05), demonstrating that as scores for Communication increases (indicating poorer communication skills), TBWs again become wider. The Social Skill subfactor shows moderate positive correlations for TBW (r(15) = 0.39, p = 0.12) and LTBW (r(15) = 0.38, p = 0.13), showing that as scores for Social Skills increase (indicating poorer social skills), TBWs become wider, however p-values are not significant for this result. These patterns are illustrated in Figure 12. It should be noted that the range of subfactor scores in these correlations was very small, therefore the significance of these tests should be treated with caution.

	t	df	r	lower	upper	p
TBW: Social Skill	1.62	15	0.39	-0.12	0.73	0.13
RTBW: Social Skill	1.63	15	0.39	-0.11	0.73	0.12
LTBW: Social Skill	1.59	15	0.38	-0.12	0.73	0.13
TBW: Communication	2.37	15	0.52	0.06	0.80	<0.05 *
RTBW: Communication	2.20	15	0.49	0.09	0.79	<0.05 *
LTBW: Communication	2.46	15	0.54	0.08	0.81	<0.05 *
TBW: Imagination	2.65	15	0.56	0.12	0.82	<0.05 *
RTBW: Imagination	2.36	15	0.52	0.05	0.80	<0.05 *
LTBW: Imagination	2.85	15	0.59	0.16	0.84	<0.05 *
TBW: Attention to Detail	-0.51	15	-0.13	-0.58	0.37	0.61
RTBW: Attention to Detail	-0.56	15	-0.14	-0.58	0.36	0.58
LTBW: Attention to Detail	-0.47	15	-0.12	-0.57	0.38	0.65
TBW: Attention Switching	-0.80	15	-0.20	-0.62	0.31	0.44
RTBW: Attention Switching	-0.93	15	-0.23	-0.64	0.28	0.37
LTBW: Attention Switching	-0.69	15	-0.17	-0.60	0.33	0.50

Table 13: Results of Pearson's correlations between width of right-side Temporal Binding Window and AQ subfactor scores of Social Skill, Communication, Imagination, Attention to Detail, and Attention Switching



Figure 12: Correlations between the width of the right-side temporal binding window and the subfactor scores of (from top left) Social Skill, Communication, Imagination, Attention to Detail, and Attention Switching

3.4 Discussion

This chapter explored whether individual differences in behavioural traits associated with ASD or widths of temporal binding windows could be considered reliable predictors of rates of integration of visual-tactile stimuli. Research comparing rates of audio-visual integration in child ASD and TD populations (Gelder et al., 1991; Mongillo et al., 2008; Irwin et al., 2011; Stevenson et al., 2014) found lower rates of integration of synchronous stimuli in the ASD populations. Based on these findings, prediction 3.1 stated that rates of visual-tactile speech integration with synchronous stimuli would be predictable from participants' AQ scores in a TD population, such that a higher AQ score would be correlated with lower rates of integration in the synchronous stimuli condition. Results of mixed models showed that the overall AQ scores were not significant predictors of rates of synchronous visual-tactile speech integration.

AQ subfactor scores were also examined as potential predictors of integration rates of synchronous visual-tactile stimuli. Significant effects of the subfactor of Social Skill were found, with a main effect showing that participants with higher Social Skill scores, indicating poorer levels of social skills, demonstrated lower levels of integration, in support of prediction 3.1. However, a significant interaction between SOA and the Social Skill subfactor was interpreted as showing that participants with poorer social skills exhibited *higher* rates of integration. A second model was run which included an interaction in the random slope and this time only the main effect result remained. As the results of the first model were contradictory they are not considered to be strong indicators of the Social Skill subfactor having a significant influence on rates of integration in either direction. Further, the model only found the Social Skill subfactor to be a significant predictor of integration rates when considering the synchronous stimuli condition, suggesting a somewhat less robust finding than if it had influenced integration rates across the asynchronous conditions (as described in prediction 3.2).

Due to research demonstrating that children with ASD have poorer temporal acuity than TD children (Bebko et al., 2006; Foss-Feig et al., 2010; Kwakye et al., 2011; Stevenson et al., 2014), prediction 3.2 proposed that participants in the present study with higher AQ scores would integrate across a wider range of SOAs as compared to participants with lower AQ scores, who would exhibit steeper curves of integrated responses. This was because individuals with higher levels of traits associated with ASD may be more likely to judge stimuli of greater SOAs as synchronous, as compared to those with lower levels of the relevant traits. Consequently these individuals may be more likely to consider stimuli with greater SOAs as part of the same event and therefore be more likely to integrate information from them. Results evaluating AQ score as a potential predictor in the asynchronous conditions found a significant interaction between AQ score and SOA, demonstrating that, in support of prediction 3.2, individuals with higher AQ scores exhibited integrated responses over a wider range of SOAs, suggesting that each specific SOA did not influence their perception as much as it did for participants with lower AQ scores. Visualizations of the data in Figure 7 revealed that participants who scored lower on the AQ demonstrated steeper curves of integrated responses, whereas participants who scored higher had slightly flatter curves over the range of SOAs, as well as slightly lower peaks of integrated responses (though there was no main effect difference). Incidentally, this finding was apparent with only a relatively small sample size which exhibited a narrow range of AQ scores from 9 to

25 (with scores of 32 or higher thought to be an indicator of possible ASD symptoms; Baron-Cohen et al., 2001). It is plausible that a wider range of AQ scores may have demonstrated stronger patterns of integration consistent with prediction 3.2.

Investigating the AQ subfactor scores in relation to prediction 3.2, a significant interaction between the Imagination subfactor and SOA revealed that participants with poorer imaginations (indicated by higher Imagination scores) had flatter overall responses across the SOAs, again suggesting that the individual SOAs did not affect their perception as much as it did for those with better imaginations (indicated by lower Imagination scores). These results were also evident in the visualizations, Figure 9(a) illustrating that participants with better imaginations generally had higher rates of integrated responses (though there was no main effect difference) and, primarily considering the curves for participants with the best and poorest imaginations, those with the better imaginations also exhibited steeper curves. As the Imagination subfactor is considered to be a component of overall AQ score, these findings can also be interpreted as supporting prediction 3.2, demonstrating that participants with higher scores on a subfactor of the AQ exhibited integrated responses which spanned a wider range of SOAs, as compared to participants with a lower scores on a subfactor of the AQ. Further, three-way interaction results demonstrated that participants with better imaginations were better at distinguishing between the /pa/ and /ba/ visual stimuli in that when the visual stimulus was /pa/, they were more likely to give integrated responses than those with poorer imaginations. These results were illustrated in Figure 9(b) which showed a clear pattern of integrated responses as a function of Imagination score when the SOA was between -100ms and 50/100ms. Figure 9(c) showed that integrated responses when the visual stimulus was /ba/ were more variable with regard to levels of imagination, though those participants with the best imaginations (a score of 0) showed a trend towards a steeper SOA curve and those with the poorest imaginations (a score of 4) exhibited the flattest curve of integrated responses. An alternative view of the data in Figure 8 illustrated the differential integrated responses for participants with relatively good vs. poor imaginations. Those participants with better imaginations exhibited a clear difference in integrated responses depending on visual stimulus, as opposed to those with poor imaginations for whom the visual stimulus did not have as much influence their responses. These results suggest that participants with better imaginations were more easily able to detect differences between the visual stimuli and more reliably reported integrated responses when the tactile stimulus was combined with the

visual stimulus of an aspirated stop and when the SOA was closer to synchronous. Intriguingly, this finding also suggests that varying levels of imagination affect individuals' multimodal speech perception, however it is currently unclear how imagination plays a part in perceptual integration behaviour.

These results showed that in an adult TD population, individual rates of visual-tactile integration vary and can be linked to participants' levels of behavioural traits thought to be associated with ASD, as measured by the AQ. The previous research on which prediction 3.2 is based showed that children with ASD exhibit differential integration behaviour from that of TD children, as well as demonstrating poorer temporal acuity in perception of multimodal stimuli. The current findings may suggest, in accordance with those findings, that TD adults with higher levels of attributes associated with ASD demonstrate perceptual behaviour similar to that of child ASD populations, and that this difference in perceptual behaviour may have its basis in differing levels of temporal acuity.

These results also suggest that the differences in integration rates previously observed in child ASD populations compared to TD groups are not specific to this clinical population, but extend along a continuum from individuals with an ASD diagnosis through to TD individuals. Further, the current results demonstrate that differences in rates of multimodal integration based on behavioural traits associated with ASD are apparent in adult, as well as child populations. These results differ from Taylor et al. (2010) who found that the differences in audio-visual integration between ASD and TD child populations disappeared by adulthood. Of course, the current study did not include children or individuals with an ASD diagnosis. To comprehensively test the notion of a continuum of multimodal integration behaviour ranging from children with ASD through to TD adults, it would be necessary to run a study which included participants from the various groups along the proposed continuum.

The present study also explored the potential relationship between temporal acuity and integration rates by comparing widths of temporal binding windows to rates of integrated responses for visual-tactile speech stimuli. Previous studies comparing audio-visual perception in child ASD and TD populations (Foss-Feig et al., 2010; Kwakye et al., 2011; Stevenson, et al., 2014) have shown that children with ASD exhibit wider temporal binding windows than TD children, and Stevenson et al. (2014) found a negative correlation between the width of these

windows and rates of audio-visual speech integration in the ASD population such that as window width increased, rates of integration of synchronous stimuli decreased. Stevenson et al. (2012) found a similar correlation between the widths of right-side temporal binding windows and integration rates in a TD adult population. Based on this research, it was proposed that wider temporal binding windows for visual-tactile stimuli would be equated with lower rates of integration in the synchronous condition (prediction 3.3). Contrary to this prediction, no significant results of a mixed model analysis were observed when considering full, right-side or left-side temporal binding window as a predictor. This was unexpected, especially considering the correlation found by Stevenson et al. (2012) within an adult TD population. There was however, a main effect result of right-side temporal binding window when considering the asynchronous conditions, such that narrower right-side windows were equated with higher rates of integrated responses. This result was in a direction which supports the previous findings by Stevenson et al. (2012), although they did not investigate asynchronous audio-visual stimuli.

The relationship between temporal acuity and integration over the full range of SOAs was also explored, with prediction 3.4 asserting that a wider full and right-side window would be correlated with visual-tactile integration which was maintained over a greater range of SOAs (showing a flatter curve of integrated responses). This was due to the likelihood of participants with poorer temporal acuity being again more inclined to consider stimuli at greater SOAs as part of the same event, and therefore integrate information from the stimuli at greater SOAs. A significant interaction between SOA and temporal binding window width of all window types showed that, similarly to the AQ and Imagination results (where participants with higher levels of behavioural traits associated with ASD exhibited flatter curves of integrated responses), participants with wider temporal binding windows demonstrated these same flatter curves of responses relative to participants with narrower windows. This result is consistent with prediction 3.4 and appears to show that for participants with wider temporal binding windows, the specific SOA did not have as much influence on their rates of integrated responses, mimicing results for participants who had generally higher AQ scores, and specifically poorer imaginations. This result also reveals a relationship between temporal acuity and rates of asynchronous visualtactile integration and is consistent with the notion that asynchronous stimuli which is judged as synchronous is also more likely to be considered part of the same event and consequently integrated.

Further, and again similarly to results for the Imagination subfactor, a significant three-way interaction between visual stimulus, SOA and temporal binding window (of all three types) showed that participants with narrower temporal binding windows responded differentially to the visual stimuli in that integrated reponses when the visual stimulus was /pa/ exhibited a steeper curve centred around synchronous, whereas when the visual stimulus was /ba/, integrated responses peaked higher and later. In contrast, participants with wider right-side temporal binding windows did not exhibit much difference in rates of integrated responses when the visual stimulus was /pa/ versus when it was /ba/ and these participants exhibited flatter overall SOA curves. This result was unexpected and suggests that individuals with better temporal acuity are somehow more easily able to detect differences between visually similar speech information. Moreover, the two-way and three-way interaction results for temporal binding window imitate those for imagination and therefore suggest a link between these two factors. These results showed that both better imaginations and narrower temporal binding windows were equated with steeper curves of integrated responses (two-way interaction) and also correlated with better detection of differences between the visual stimuli (three-way interaction). As imagination is considered a behavioural trait associated with ASD, this also shows a link between temporal acuity and ASD.

The relationship between AQ scores and temporal binding window widths was also directly explored. Based on prior research into temporal processing in ASD populations (Foss-Feig et al., 2010; Kwakye et al., 2011; Stevenson et al., 2014), prediction 3.5 stated that higher AQ scores would correspond to wider full and right-side temporal binding windows. Results of correlation tests showed weak relationships between AQ scores and temporal binding windows of all three types, however, these correlation results did not reach significance, which may be due to the small sample size used. These results are therefore unreliable, though they suggest that outside of diagnosed ASD populations, there appears to be a relationship between levels of behavioural traits associated with ASD and temporal acuity, such that higher levels of traits associated with ASD are equated with wider temporal binding windows.

Potential links between AQ subfactors and temporal binding window width were also considered and correlation tests revealed strong positive correlations between the Imagination and Communication subfactors and the three window types, demonstrating that participants with poorer imaginations and communication skills had wider temporal binding windows. Again considering the subfactors as components of the overall AQ score, these results lend support to prediction 3.5, showing a relationship between behavioural traits associated with ASD and temporal acuity. Further, the correlation result involving the Imagination subfactor is in concordance with the link between imagination and temporal acuity which was previously established.

Results from the individual differences aspect of the current study revealed a relationship between rates of integration of visual-tactile speech stimuli, behavioural traits associated with ASD and temporal acuity based on widths of temporal binding windows. Based on previous research which found a relationship between lower rates of synchronous audio-visual integration and wider temporal binding windows in an ASD population (Stevenson et al., 2014), it is unclear why no significant relationships were found between AQ score or temporal binding window width and integration rates when considering the synchronous visual-tactile stimuli conditions in the present study. However, results for the asynchronous conditions contribute new knowledge to our understanding in this research area, demonstrating that rates of multimodal integration were maintained over a greater range of asynchronies for individuals with higher levels of behavioural traits associated with ASD and individuals with poorer temporal acuity as measured by temporal binding window width. Lower rates of integration were also evident for participants with wider right-side temporal binding windows. Unexpected results showed a relationship between levels of imagination, widths of temporal binding windows and the ability to detect differences between visually similar speech stimuli. Finally, results of correlation tests showed strong correlations between two behavioural traits associated with ASD; imagination and communication skills, and widths of temporal binding windows. These results imply that differential multimodal integration behaviour exhibited in child ASD populations is not specific to this group but exists on a continuum which extends from child ASD to adult TD populations and may be linked to differing levels of an individual's behavioural traits associated with ASD, and temporal acuity. The current results also show that the relationships between integration rates, traits associated with ASD and temporal acuity are present during visual-(aero)tactile speech integration, suggesting that these factors influencing integration behaviour do not target specific modality combinations, but affect multimodal speech perception across a variety of modality combinations.

4. Conclusions

The present study investigated potential integration of aero-tactile speech information during visual-(aero)tactile speech perception. It was also concerned with exploring individual differences in rates of integration based on temporal acuity and behavioural traits associated with ASD.

Results from Chapter 2 showed that indeed, in the absence of the original auditory speech signal, when perceiving visual-tactile speech individuals take into account aero-tactile speech information and integrate it as aspiration information. This finding makes a significant contribution to research into multimodal speech perception by demonstrating perceptual integration involving a combination of modalities which is currently underexplored. Further, it demonstrates the possibility of speech integration which does not involve information from the original auditory speech signal. This result adds to our current understanding of multimodal speech integration and supports notions that speech is not perceived in terms of an auditory event which is supplemented by information from other modalities, but as a holistic, modality neutral event.

Results also revealed that integration occurs both when the visual and aero-tactile stimuli were presented simultaneously and when the stimuli were presented asynchronously, within a temporal window from -50ms (audio-lead) to 300ms (visual-lead). This window of visual-tactile integration is similar, though slightly wider, to that found for audio-visual and audio-(aero)tactile speech integration and demonstrates the same shape in terms of asymmetry with the right-side, where the visual stimulus precedes the audio, showing integrated responses over a greater range of asynchronies than the left-side, where the tactile stimulus precedes the visual. This is consistent with our knowledge of relative speeds of light and speech airflow. Light, which carries visual information travels much faster than the rate of speech airflow and as such may be perceived by the brain slightly earlier than airflow information from the same speech event. Through experience, we learn to synchronize stimuli that our brains perceive slightly
asynchronously if it is derived from the same event, presumably so that (speech) events can be perceived efficiently.

Chapter 3 investigated individual differences in visual-tactile speech integration based on behavioural attributes associated with ASD as measured by Autism Spectrum Quotient scores, and temporal acuity as measured by temporal binding window width. Results showed that levels of behavioural traits associated with ASD significantly influenced the range of SOAs over which an individual was likely to integrate, with individuals who scored higher on the AQ exhibiting integrated responses over a greater range of asynchronies than those with lower AQ scores. One trait in particular, imagination, significantly predicted integration rates such that a poorer imagination was correlated with integration which was maintained over a greater range of asynchronous stimuli. Similarly, the width of an individual's temporal binding window (of all three window types) was also shown to influence an individual's rates of visual-tactile speech perception in the same direction, with wider windows being correlated with integration which was maintained across a greater range of stimulus asynchronies. These results suggest that individuals with higher levels of behavioural traits associated with ASD, in particular, poorer imaginations, as well as individuals with relatively poor temporal acuity are less influenced perceptually by the time alignment of stimuli when integrating information from the aero-tactile and visual modalities.

Further results revealed an intriguing link between levels of imagination, temporal acuity and the ability to detect differences between visual speech stimuli, with individuals with better imaginations and narrower temporal binding windows appearing to be better able to differentiate between the two bilabial visual stimuli, more reliably reporting integrated responses when tactile information was presented with the visual stimuli of an aspirated stop. These results show that the same factors (imagination and temporal acuity) which affected the strength of influence of the time alignment of stimuli during visual-tactile integration also affected the ability to differentiate subtle differences between visually similar speech stimuli. Finally, the current study compared temporal binding window widths and AQ scores, finding strong correlations between two traits associated with ASD, imagination and communication skills, and widths of temporal binding windows, such that higher levels of these traits were equated with wider windows of all

three window types. These results demonstrate a direct link between behavioural traits associated with ASD, and temporal acuity.

Together, the findings from investigations into individual differences show that visual-tactile integration, behavioural traits associated with ASD, and temporal acuity are all interrelated and suggest that the differential integration behaviour between child ASD and TD groups reported in previous research is not specific to populations with a clinical diagnosis of ASD, but exists on a continuum that extends to TD populations, and from child to adult groups.

The current study explored perceptual integration using visual and aero-tactile speech information for nonsense CV syllables as stimuli. Real speech in context is obviously perceived differently from nonsense syllables presented in isolation. Future research may investigate visual-tactile speech integration in real speech contexts, i.e. using real words in sentences. If integration occurs in this speech environment too, this would provide even stronger evidence of speech integration involving visual and tactile information in the absence of information from the original auditory speech signal. Considering that *aero*-tactile information is just one type of tactile speech information, another avenue of interest for future research may be to explore integration using other types of tactile speech information. For example, vibro-tactile information may be used to simulate voicing, and a paradigm similar to the one in the present study could be set up to test the potential perceptual influence this type of tactile information has on multimodal speech integration.

Considering the investigation into individual differences, the present study has established links between rates of visual-tactile integration, behavioural traits associated with ASD, and temporal acuity. The underlying reasons for these links are uncertain and this is unquestionably due to the complex nature of ASD, which is characterized by a range of atypical behaviours rather than having a singular diagnosis. Future research may explore the basis of some of these relationships more closely. The link between imagination and rates of visual-tactile integration, for example, is an intriguing result with a non-transparent explanation which deserves closer examination. Though the current results appear to support the notion of a continuum of multimodal perceptual behaviour which includes both ASD and TD populations, as well as children and adults, the current study did not involve participants from all of these groups. Future work could more comprehensively test the notion of a continuum by sampling the full range of populations proposed to exist on the continuum. If validated, this could both lend support to the accuracy of the current self-reporting methodology, as well as leading to the development of multimodal speech perception tasks to assist in testing for symptoms of ASD.

References

- Abel, J., Barbosa, A. V., Mayer, C., & Vatikiotis-Bateson, E. (2011). The labial viseme reconsidered: Evidencce from production and perception. In Y. Laprie, & I. Steiner (Ed.), *9th International Seminar on Speech Production (ISSP)* (pp. 337-344). Montreal: PQ.
- Alcorn, S. (1932). The tadoma method. Volta Rev, 34, 195-198.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autismspectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders*, 31(1), 5-17.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). Retrieved from _lme4: Linear mixedeffects models using Eigen and S4_. R package version 1.1-7: http://CRAN.Rproject.org/package=lme4
- Bebko, J., Weiss, J., Demark, J., & 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.
- Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.08, retrieved 24 March 2015 from http://www.praat.org/.
- Chouinard, P., Noulty, W., Sperandio, I., & Landry, O. (2013). Global processing during the Müller-Lyer illusion is distinctively affected by the degree of autistic traits in the typical population. *Experimental brain research*, 230(2), 219-231.
- Davies, S., Bishop, D., Manstead, A., & Tantam, D. (1994). Face perception in children with autism and Asperger's syndrome. *Journal of Child Psychology and Psychiatry*, 35(6), 1033-1057.

- Derrick, D., Anderson, P., Gick, B., & Green, S. (2009). Characteristics of air puffs produced in English "pa": Experiments and simulations. *The Journal of the Acoustical Society of America*, 125(4), 2272-2281.
- Donohue, S., Darling, E., & Mitroff, S. (2012). Links between multisensory processing and autism. *Experimental brain research*, 222(4), 337-387.
- Fisher, C. (1968). Confusion among Visually Perceived Consonants. *Journal of Speech and Hearing Research, 11*, 796-804.
- Foss-Feig, J., Kwakye, L., Cascio, C., Burnette, C., Kadivar, H., Stone, W., et al. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, 203(2), 381-389.
- Fowler, C., & Dekle, D. (1991). Listening with eye and hand: cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 17(3), 816.
- Frith, U. (1989). Autism: Explaining the enigma. Oxford: Basil Blackwell.
- Gelder, B., Vroomen, J., & Van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology*, *3*(1), 69-86.
- Gick, B., & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*, 462(7272), 502-504.
- Gick, B., Ikegami, Y., & Derrick, D. (2010). The temporal window of audio-tactile integration in speech perception. *The Journal of the Acoustical Society of America*, 128(5), EL342-EL346.
- Gick, B., Jóhannsdóttir, K., Gibraiel, D., & Mühlbauer, J. (2008). Tactile enhancement of auditory and visual speech perception in untrained perceivers. *The Journal of the Acoustical Society of America*, 123(4), EL72-EL76.

- Grinter, E., Maybery, M., Van Beek, P., Pellicano, E., Badcock, J., & Badcock, D. (2009).
 Global visual processing and self-rated autistic-like traits. *Journal of autism and developmental disorders*, 39(9), 1278-1290.
- Irwin, J., Tornatore, L., Brancazio, L., & Whalen, D. (2011). Can children with autism spectrum disorders "hear" a speaking face? *Child development*, 82(5).
- Kwakye, L., Foss-Feig, J., Cascio, C., Stone, W., & Wallace, M. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in integrative neuroscience*, 4.
- Mayer, C., Gick, B., Weigel, T., & Whalen, D. (2013). Perceptual integration of visual evidence of the airstream from aspirated stops. *Canadian Acoustics*, *41*(3), 23-27.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746–748.
- Miller, L., & D'esposito, M. (2005). Perceptual fusion and stimulus coincidence in the crossmodal integration of speech. *The Journal of neuroscience*, 25(25), 5884-5893.
- 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 and Developmental Disorders*, *38*(7), 1349-1358.
- Munhall, K., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics*, 58(3), 351-362.
- Palmer, C., Paton, B., Hohwy, J., & Enticott, P. (2013). Movement under uncertainty: the effects of the rubber-hand illusion vary along the nonclinical autism spectrum. *Neuropsychologia*, 51(10), 1942-1951.
- Peirce, J. (2007). PsychoPy Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8-13.
- R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

- Stevenson, R., Siemann, J., Schneider, B., Eberly, H., Woynaroski, T., Camarata, S., et al. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, 34(3), 691-697.
- Stevenson, R., Zemtsov, R., & Wallace, M. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1517.
- Sumby, W., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212-215.
- Taylor, N., Isaac, C., & Milne, E. (2010). A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. *Journal of autism and developmental disorders*, 40(11), 1403-1411.
- van Wassenhove, V., Grant, K., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, 45(4), 598-607.
- Wickham, H. (2009). ggplot2: elegant graphics for data analysis. New York: Springer Science & Business Media.
- Williams, J., Massaro, D., Peel, N., Bosseler, A., & Suddendorf, T. (2004). Visual–auditory integration during speech imitation in autism. *Research in developmental disabilities*, 25(6), 559-575.
- Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of autism and developmental disorders*, 9(1), 11-29.
- Yu, A. C. (2010). Perceptual compensation is correlated with individuals'"autistic" traits: Implications for models of sound change. *PLoS ONE*, 5(8), e11950.
- Yu, A., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individualdifference perspective: Subjective attitude, personality and "autistic" traits. *PLoS one*, 8(9), 1-13.

Appendix A: Autism Spectrum Quotient Questionnaire

Below is a copy of the paper version of the Autism Spectrum Quotient (AQ) (Adult) sourced from the Autism Research Centre website <u>http://www.autismresearchcentre.com/arc_tests</u>

Note that in the current study, this test was administered via computer and did not contain the title information or require that the participant give their personal details. It also included only one example statement.

The Adult Autism Spectrum Quotient (AQ)

<u>Ages 16+</u>

SPECIMEN, FOR RESEARCH USE ONLY.

For full details, please see:

S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin and E. Clubley, (2001) The Autism Spectrum Quotient (AQ) : Evidence from Asperger Syndrome/High Functioning Autism, Males and Females, Scientists and Mathematicians Journal of Autism and Developmental Disorders 31:5-17

Name:	Sex:
Date of birth:	Today's Date

How to fill out the questionnaire

Below are a list of statements. Please read each statement <u>very carefully</u> and rate how strongly you agree or disagree with it by circling your answer.

DO NOT MISS ANY STATEMENT OUT.

Examples

	E1. I am willing to take risks.	definitely	slightly	slightly	definitely
(agree	agree	disagree	disagree
/	E2. The playing board games.	definitely	slightly	slightly	definitely
(agree	agree	disagree	disagree
	E3. I find learning to play musical instruments easy.	definitely	slightly	slightly	definitely
		agree	agree	disagree	disagree
	E4. I am fascinated by other cultures.	definitely	slightly	slightly	definitely
	(agree	agree	disagree	disagree

I prefer to do things with others rather than on	definitely	slightly	slightly	definitely
my own.	agree	agree	disagree	disagree
2. I prefer to do things the same way over and over	definitely	slightly	slightly	definitely
again.	agree	agree	disagree	disagree
3. If I try to imagine something, I find it very easy	definitely	slightly	slightly	definitely
to create a picture in my mind.	agree	agree	disagree	disagree
4. I frequently get so strongly absorbed in one	definitely	slightly	slightly	definitely
thing that I lose sight of other things.	agree	agree	disagree	disagree
5. I often notice small sounds when others do not.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
6. I usually notice car number plates or similar	definitely	slightly	slightly	definitely
strings of information.	agree	agree	disagree	disagree
7. Other people frequently tell me that what I've	definitely	slightly	slightly	definitely
said is impolite, even though I think it is polite.	agree	agree	disagree	disagree
8. When I'm reading a story, I can easily imagine	definitely	slightly	slightly	definitely

what the characters might look like.	agree	agree	disagree	disagree
9. I am fascinated by dates.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
10. In a social group, I can easily keep track of	definitely	slightly	slightly	definitely
several different people's conversations.	agree	agree	disagree	disagree
11. I find social situations easy.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
12. I tend to notice details that others do not.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
13. I would rather go to a library than a party.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
14. I find making up stories easy.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
15. I find myself drawn more strongly to people than	definitely	slightly	slightly	definitely
to things.	agree	agree	disagree	disagree
16. I tend to have very strong interests which I get	definitely	slightly	slightly	definitely
upset about if I can't pursue.	agree	agree	disagree	disagree
17. I enjoy social chit-chat.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
18. When I talk, it isn't always easy for others to get	definitely	slightly	slightly	definitely
a word in edgeways.	agree	agree	disagree	disagree
19. I am fascinated by numbers.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
20. When I'm reading a story, I find it difficult to	definitely	slightly	slightly	definitely
work out the characters intentions.	agree	agree	disagree	disagree
21. I don't particularly enjoy reading fiction.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
22. I find it hard to make new friends.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree

23. I notice patterns in things all the time.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
24. I would rather go to the theatre than a museum.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
25. It does not upset me if my daily routine is	definitely	slightly	slightly	definitely
disturbed.	agree	agree	disagree	disagree
26. I frequently find that I don't know how to keep a	definitely	slightly	slightly	definitely
conversation going.	agree	agree	disagree	disagree
27. I find it easy to "read between the lines" when	definitely	slightly	slightly	definitely
someone is tarking to me.	agree	agree	disagree	disagree
28. I usually concentrate more on the whole picture,	definitely	slightly	slightly	definitely
rather than the small details.	agree	agree	disagree	disagree
29. I am not very good at remembering phone	definitely	slightly	slightly	definitely
numbers.	agree	agree	disagree	disagree
30. I don't usually notice small changes in a	definitely	slightly	slightly	definitely
situation, or a person's appearance.	agree	agree	disagree	disagree
31. I know how to tell if someone listening to me is	definitely	slightly	slightly	definitely
getting bored.	agree	agree	disagree	disagree
32. I find it easy to do more than one thing at once.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
33. When I talk on the phone, I'm not sure when it's	definitely	slightly	slightly	definitely
my turn to speak.	agree	agree	disagree	disagree
34. I enjoy doing things spontaneously.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
35. I am often the last to understand the point of a	definitely	slightly	slightly	definitely
јоке.	agree	agree	disagree	disagree
36. I find it easy to work out what someone is	definitely	slightly	slightly	definitely
uninking or reeling just by looking at their face.	agree	agree	disagree	disagree
37. If there is an interruption, I can switch back to	definitely	slightly	slightly	definitely

what I was doing very quickly.	agree	agree	disagree	disagree
38. I am good at social chit-chat.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
39. People often tell me that I keep going on and on	definitely	slightly	slightly	definitely
about the same thing.	agree	agree	disagree	disagree
40. When I was young, I used to enjoy playing	definitely	slightly	slightly	definitely
games involving pretending with other children.	agree	agree	disagree	disagree
41. I like to collect information about categories of	definitely	slightly	slightly	definitely
train, types of plant, etc.).	agree	agree	disagree	disagree
42. I find it difficult to imagine what it would be like	definitely	slightly	slightly	definitely
to be someone else.	agree	agree	disagree	disagree
43. I like to plan any activities I participate in	definitely	slightly	slightly	definitely
carefully.	agree	agree	disagree	disagree
44. I enjoy social occasions.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
45. I find it difficult to work out people's intentions.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
46. New situations make me anxious.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
47. I enjoy meeting new people.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
48. I am a good diplomat.	definitely	slightly	slightly	definitely
	agree	agree	disagree	disagree
49. I am not very good at remembering people's date	definitely	slightly	slightly	definitely
of birth.	agree	agree	disagree	disagree
50. I find it very easy to play games with children	definitely	slightly	slightly	definitely
that involve pretending.	agree	agree	disagree	disagree

Developed by:

The Autism Research Centre

University of Cambridge

© MRC-SBC/SJW Feb 1998