AUDIOVISUAL INTEGRATION IN ADULTS: USING A DYNAMIC TASK TO MEASURE DIFFERENCES IN TEMPORAL BINDING WINDOWS ACROSS STIMULI

MELISSA FERLAND

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF ARTS

GRADUATE PROGRAM IN CLINICAL DEVELOPMENTAL PSYCHOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

SEPTEMBER 2017

© MELISSA FERLAND, 2017

ABSTRACT

Being able to integrate information from multiple sensory modalities, such as hearing and sight, is essential for everyday functioning. The temporal binding window (TBW) refers to how out of synch these two modalities can be before they are considered asynchronous. In the current study a new method to measure the TBW of audiovisual stimuli with varying social and linguistic contents was developed. Participants manually adjusted the soundtrack of a video by varying increments until it was synchronous to the visual information (*toward-synch*) or asynchronous (*from-synch*). The newly developed task with increments of 50ms produced the smallest TBW compared to all other versions of the task and to a commonly used method. Smaller windows were found for speech stimuli compared to both non-social and non-speech ones, and for adjustments *toward-synch* versus *from*-synch. Giving participants the ability to control the soundtrack proved to be a superior methodology over prior commonly used ones.

ACKNOWLEDGEMENTS

Je voudrais tout d'abord remercier mes parents de leur amour inconditionnel. Je n'aurais pu m'imaginer poursuivre mes études aux cycles supérieurs sans leur appui et leurs encouragements. Je vous aime énormément.

I would also like to thank my supervisor, Dr. James M. Bebko, for his guidance and support throughout this project. This project would never have seen the light of day if it was not for his creative outlook on the methodologies of audiovisual integration and his vast knowledge in the field. Thank you for bearing with me through all the ups and downs of this project.

I would also like to thank my committee member, Dr. Christine Till for her valuable input and suggestions throughout this writing process. I am deeply appreciative for all the time you took to carefully edit my drafts. I would also like to express my gratitude towards the other members of my thesis committee, Dr. Laurence Harris and Dr. Mary Desrocher, for agreeing to be part of my committee and for your feedback.

Lastly, I would like to express my gratitude toward my friends and especially my roommate, Elisea De Somma, for keeping me sane throughout this process and for always being there to offer advice. Thank you also to Charles Burger for reading through my entire thesis and fixing my spelling, and encouraging me to keep going. Finally, this project would not have been possible without the cooperation of all the participants who took part in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Audiovisual Integration in Adults: Using a Dynamic Task to Measure Differences in Temporal Binding Windows Across Stimuli

Ongoing events in everyday life are often perceived and registered through multiple sensory organs. For example, when interacting with someone in a social situation, both auditory, sounds emitted from the speaker's mouth, and visual, movement of the lips and nonverbal facial expressions from the speaker, information are communicated to the listener at once. The ability of the perceptual system to coordinate and integrate information across these different sensory modalities is crucial in providing an accurate and coherent representation of the outside world. Furthermore, how the perceptual system integrates and organizes information is critical to what one perceives, learns, and remembers from the environment and therefore can shape one's social and adaptive behaviours (Wallace, 2004). Therefore, understanding the integration of multiple sensory information holds many real-world implications and is critical for everyday functioning.

In order to successfully navigate through this complex external world, the human brain must be able to differentiate multiple sensory cues that originated from one single physical event, and thus should be grouped as a single unified perceptual unit, from those that should not. Although many have studied and researched this topic, little advancement has been made in terms of the methodologies and techniques used to measure such events. The current study will attempt to develop new ways of examining multiple sensory integration by building on previous research and using an approach where participants are more involved.

Audiovisual Integration

One of the most frequently occurring multisensory integrations is that of audiovisual information (e.g., van Eijk, Kohlrausch, Juola, & van de Par, 2008). Individuals, often and unconsciously, perceive audiovisual information that originates from the same physical source as one single perceptual unit, such as when having a face-to-face conversation with someone or hearing and viewing the sirens on a firetruck as it passes by. Audiovisual integration holds many benefits such as improving reaction time (e.g., Colonius & Diederich, 2004; Diederich & Colonius, 2015; Gondan, Niederhaus, Rösler, & Röder, 2005; Hershenson, 1962), recognizing and identifying speech (e.g., Calvert, Brammer, & Iversen, 1998) or other stimuli (e.g., Lovelace, Stein, & Wallace, 2003), localizing stimuli (Hairston, Laurienti, Mishra, Burdette, & Wallace, 2003), and understanding speech in noisy environments (e.g., Bishop & Miller, 2009; Girin, Schwartz, & Feng, 2001; Grant & Walden, 1996; Sumby & Pollack, 1954; Summerfield, 1987). Furthermore, recent work by Zmigrod & Zmigrod (2016), found that greater sensitivity to audiovisual integration was related to better verbal and nonverbal problem solving abilities. In their study, healthy participants had to judge whether a flashing circle appeared at the same time as a beep sound, and the degree of offset of the sound in comparison to the visual information varied across trials. Their results showed that performance on the flash-beep task significantly predicted participants' performances on both verbal and nonverbal problem solving tasks (Zmigrod & Zmigrod, 2016). The authors thus put forth the theory that audiovisual integration ability could mirror other complex cognitive abilities, such as reasoning skills (Zmigrod & Zmigrod, 2016). As can be seen, an appropriate and accurate integration of audio and visual information can be extremely beneficial to navigate the external environment and can be linked to enhanced cognitive abilities.

On the other hand, atypical multisensory functioning, specifically difficulties in audiovisual integration, have been seen in several developmental disabilities, such as autism (e.g., Bahrick, 2010; Bebko, Weiss, Denmark, & Gomez, 2006; de Boer-Schellekens, Eussen, &

Vroomen, 2013; Keane, Rosenthal, Chun, & Shams, 2010; Mongillo et al., 2008; Taylor, Isaac, & Milne, 2010), dyslexia (Froyen, Willems, & Blomert, 2011; Hairston, Burdette, Flowers, Wood, & Wallace, 2005; Henry, 1998; Oakland, Black, Stanford, Nussbaum, & Balise, 1998; Wallace & Stevenson, 2014), as well as mental health disorders such as schizophrenia (de Gelder et al., 2005; de Jong, Hodiamont, Van den Stock, & de Gelder, 2009; Pearl et al., 2009; Ross et al., 2007; Szycik et al., 2009; Wallace & Stevenson, 2014; Williams, Light, Braff, & Ramachandran, 2010). For individuals with autism spectrum disorder (ASD) specifically, research has shown that the multisensory deficit in ASD goes beyond the deficits seen in individual sensory modalities (Ross et al., 2011; Stevenson et al., 2014b) as these individuals have been shown to have difficulties in matching voices to faces (Boucher, Lewi, & Colis, 1998), and perceiving perceptual unification illusion, especially in speech (e.g., Bebko, Schroeder, & Weiss, 2014; Foss-Feig et al., 2010). All these difficulties linked with audiovisual integration in individuals with ASD have been postulated to contribute to the inability to accurately represent the external environment and contribute to their social ability difficulties (Wallace & Stevenson, 2014). Therefore, a better understanding of the mechanism and properties involved in audiovisual integration could hold valuable information regarding perceptual processing and these disorders.

When visual and auditory information are not integrated, they are perceived as two distinct events, like a badly dubbed movie, where the voice of the actors and their lip movements are clearly out of synchrony. When one sees these two sensory modalities (i.e., audio and visual) as synchronous the visual information is integrated into a single perceptual event with what is heard (Dixon & Spitz, 1980). Whether auditory and visual information are considered synchronous is contingent on spatial and temporal factors (Lewald, Ehrenstein, & Guski, 2001;

Wallace, 2004; Wallace & Stevenson, 2014; Zmigrod & Zmigrod, 2016). Multisensory integration is more likely to occur if audio and visual information originate from the same or proximal sources (e.g., Bedford, 1989; Dixon & Spitz, 1980; Hillock, Powers, & Wallace, 2011; Meredith & Stein, 1986; Radeau, 1994; Spence, 2007, Teder-Sälejärvi, Di Russo, McDonald, & Hillyard, 2005; Welch, 1999; van Wassenhove, Grant, & Poeppel, 2007) and if the visual and auditory sensory information reaches the brain at approximately the same time (e.g., de Gelder & Bertelson, 2003; Vatakis & Spence, 2010; Vroomen & Keetles, 2010). However, given the fact that sound and light travel at different speeds and that their transduction speeds within the brain also differ; a certain degree of asynchrony between these two modalities is inevitable (Powers, Hillock, & Wallace, 2009).

Temporal Binding Window

The temporal binding window (TBW), also known as the temporal window of integration, refers to a window in which separate sensory information are perceived as one, synchronous event, despite some degree of asynchrony between these senses (Diederich & Colonius, 2009; Dixon & Spitz, 1980; Mégevand, Molholm, Nayar, & Foxe, 2013; Spence & Squire, 2003; Vroomen & Keetels, 2010). Many researchers have set out to measure the exact size of the audiovisual TBW and have discovered that the actual width is fairly broad (e.g. Hillock et al., 2011; van Wassenhove et al., 2007; Zmigrod & Hommel, 2011; Zmigrod & Zmigrod, 2016). Amongst the first ones to provide a definite measurement for the TBW of audiovisual integration were Dixon and Spitz (1980). They originally estimated the audiovisual TBW to span from -130ms to +250ms, where the negative sign represents auditory information presented before visual information, and the positive sign shows the opposite. Others have claimed that the TBW for audiovisual information could extend between -200ms to +350ms

(Zmigrod & Zmigrod, 2016). In both cases, a positively skewed TBW can be observed, where individuals seem better at detecting asynchrony when auditory information comes before visual information, as can be seen by the smaller difference from zero (which represents true synchrony) in the negative direction compare to the positive one. It has been stipulated that this positively skewed TBW for audiovisual information is a result of perceptual experience as visual information that precedes auditory information occurs more often in the “real world” (e.g., one sees lighting before hearing the thunder) (Hillock et al., 2011). Therefore, when the opposite happens where auditory information come before visual information, individuals are quicker to notice the violation (Hillock et al., 2011). In the “real world”, light travels faster than sound as the electromagnetic waves do not need a medium to travel through, while sound does, and therefore, visual information is more likely to reach an individual before sound, even though both events might be originating from the same place. Furthermore, Hillock et al., 2011 showed that the asymmetry in the TBW is not as drastic in children compared to adults and argues that this observation could be the result of adults having been more exposed to these perceptual experiences where sound occurs after visual information (Hillock et al., 2011). This demonstrates that the size of the TBW can also be influenced by certain factors such as the history of occurrence of asynchrony in one’s surrounding environment.

As research in the field of TBW has grown more popular over the last decade, researchers demonstrated that the TBW for audiovisual integration is neither stable nor consistent, but rather dynamic and flexible (e.g., de Boer-Schellekens & Vroomen, 2014; Hillock-Dunn & Wallace, 2012; Mégevand et al., 2013; Powers et al., 2009; Stevenson et al., 2014a; Stevenson, Zemtsov, & Wallace, 2012; van Wassenhove et al., 2007). The exact size of the TBW can be influenced by several factors, such as the sensory modalities being combined (Fujisaki & Nishida, 2009), the stimulus duration and intensity (Boenke, Deliano, & Ohl, 2009; Krueger Fister, Stevenson, Nidiffer, Barnett, & Wallace, 2016), age (Bates & Wolbers, 2014; Diederich, Colonius, & Schomburg, 2008; Hillock et al., 2011; Hillock-Dunn & Wallace, 2012; Lewkowicz & Flom, 2014), individual variability (e.g., Stevenson et al., 2012; Zmigrod & Zmigrod, 2016), and the task used to measure the TBW (Diederich & Colonius, 2015; van Eijk et al., 2008). More specifically, wider temporal binding windows have been associated with younger children (Hillock-Dunn & Wallace, 2012) and elderly (Diederich et al., 2008), and with lower intensity stimuli (Krueger et al., 2016). Furthermore, research has also shown that TBW might be narrower in males compared to females, demonstrating how gender may also play an influential role in multisensory integration (Zmigrod & Zmigrod, 2016). All these studies demonstrate that the sizes of the TBWs are highly influential.

The type of stimulus used has also been shown to impact the size of the TBW (e.g., Bebko, et al., 2006; Stevenson & Wallace, 2013; Van Eijk et al., 2008; Van Wassenhove et al., 2007; Vataski & Spence, 2006). Commonly occurring stimuli, also called complex stimuli, such as speech and object-action (i.e., a hammer hitting a nail), are associated with narrower TBW compared to more arbitrary stimuli (i.e., flash-beep) (e.g., van Eijk et al., 2008). More specifically, Segers (2012), from whom the current study is a continuation and extension of, has demonstrated that within commonly occurring stimuli, the TBW for audiovisual stimuli in adults is wider for social non-linguistic (i.e., a person making kissing sounds) and non-social non-linguistic (i.e., a hand playing the piano) stimuli, than for social linguistic ones (i.e., speech). These results demonstrate that adults are more sensitive to speech, and are less likely to combine audio and visual information together if the temporal discrepancy between the two is too large. Similar results were observed in Vatakis & Spence (2006), where they demonstrated that adults were more sensitive to asynchrony in speech than for either guitar or piano music videos. An increased sensitivity to asynchrony in speech prevents mispairing of audio and visual information, and helps one have a more accurate view of the social external world. Regardless of how flexible the sizes of the TBWs are, it is still important to attempt to measure the actual widths of these windows across different situations to have a better understanding of what can influence temporal binding of the external world.

Measuring the Temporal Binding Window

For over a century, scientists have been trying to measure the perception of sensory information. Ever since Gustave Theodor Fechner in 1860 put forward the term "psychophysics," the concept that psychological experiences of a stimulus such as sensory perception could be objectively measured, researchers have attempted to develop the most accurate way of measuring such phenomenon (Fechner, 1860/1966). Fechner and other pioneers in the field postulated the idea that one could quantify perception by using a specific set of experimental measures. Although many of these methods were first developed to focus on unimodal sensory systems, they have forged the way for the development of measuring multisensory integration. From trying to measure the absolute threshold, that minimum level of intensity needed to detect a sensory event, to using a method of adjustments, where individuals adjust the physical intensity of a stimulus until the point they can barely detect it, psychophysics influences are everywhere within the field of measuring audiovisual integration (Chaudhuri, 2011; Ehrenstein & Ehrenstein, 1999).

Specific to the idea of the measuring the temporal binding window and to one of the tasks used in this study is the method of constant stimuli. This method consists of exposing participants to a variety of set intensities of a stimulus that have been arranged in a random order

(Ehrenstein & Ehrenstein, 1990). What is specific about this method is that participants are unaware of the order of the presentation of the stimuli. For example, in contrast to moving in a specific direction where stimuli either get more or less intense, such as in the method of adjustments, participants do not know what the next stimulus will look like in comparison to the previous one (Chaudhuri, 2011). By using the method of constant stimuli and requiring the participants to make a forced-choice between whether they can or cannot detect the stimuli (or the targeted perception, which in the case of this study is the simultaneity between audio and visual information), researchers can develop a psychometric function of what an individual can and cannot perceive (Ehrenstein & Ehrenstein, 1999; Leek, 2001). With regards to audiovisual integration, the method of constant stimuli can be used to present participants with fixed stimuli in which audio and visual information have been put out of synchrony with each other at various degrees and where participants must make a decision regarding whether the two sensations are synchronous or not. Information gathered from such a task could then be used to produce a psychometric function of what the participants could and could not detect which in turn could be used to create a temporal binding window for audiovisual integration. As can be seen, principals of psychophysics play an influential role in the field of measuring temporal binding windows.

The two most common methods used to measure the TBW for audiovisual stimuli are the synchrony judgment (SJ) task and the temporal order judgment (TOJ) task. In the SJ task, the participant is presented with a pre-produced audio-visual file and is asked whether the video and audio information are in synch or not (Bushara, Grafman, & Hallett, 2001; Exner, 1875; Sugita and Suzuki, 2003). Meanwhile, in the TOJ task the participant is asked to identify which modality was presented first, the audio or the visual (Bald, Berrien, Price, & Sprague, 1942; Hamlin, 1895; Hirsh & Sherrick, 1961). As can be seen, both the TOJ and the SJ methods are based on the fundamental principles of the method of constant stimuli and forced-choice. However, the TOJ is more difficult as it requires one to be able to identify the order in which the components of the stimuli are presented and not just whether the stimuli are synchronous or not.

Task difficulty along with stimulus complexity have been postulated as potential reasons why conflicting results have been seen between TOJ and SJ tasks when it comes to measuring the size of audiovisual TBW (e.g., Allan, 1975; Aschersleben, 1999; Hirsh & Fraisse, 1964; van Eijk et al., 2008). A study by van Eijk and colleagues (2008), demonstrated that the TOJ and SJ tasks yield different results when it comes to the point of subjective simultaneity (PSS) of audiovisual stimuli, the point where one believes the audio and visual information are simultaneous. Participants in this study took part in both the SJ and the TOJ tasks and their PSS were compared across various stimuli, which ranged from commonly occurring (bouncing ball) to arbitrary (flash and beep). It was shown that within participants, the PSS estimates of the SJ task differed from those of the TOJ task, demonstrating that PSS is dependent on task. Moreover, the PSS estimates of the SJ task were correlated to a third task that asked participants to either identify the stimuli as “audio first” ,“visual first”, or “synch” (a merge of the TOJ and SJ tasks), while the TOJ was not. The authors postulated that because the TOJ task, in comparison to the SJ and the hybrid tasks, does not offer an option to indicate synchrony, the PSS is harder to determine as it can be anywhere within the calculated synchrony range. Each PSS for the TOJ task were determined based on the point where the proportion of audio first judgment equals the proportion of video first judgments for both audio delayed and advanced stimuli separately, which provided a synchrony range. It is possible that participants use different response strategy on TOJ when trying to determine the temporal order of a stimulus that they perceive as being synchronous. Van Eijk and colleagues (2008) concluded that when it comes to measuring perceived audiovisual synchrony, the SJ task should be preferred over the TOJ task, which is why the current study will use the SJ task as a comparison task.

Both the SJ and the TOJ tasks could be considered *"static"* as neither of them enables the participant to *dynamic*ally manipulate the size of the window. Instead, in both cases, participants are presented with a video that has a fixed audio delay or advance track, and they are asked to make a decision. Although both of these measures have been demonstrated to be quite useful in the past and many researchers in the field use them, it may be beneficial to use a different approach to measuring TBW size. A more *"dynamic"* task that would allow participants to adjust the audio file until they believe it is in sync with the video could allow researchers to have a more precise definition of the size of the TBW. The *dynamic* task for measuring the size of the TBW of various audiovisual stimuli was originally designed by Bebko & Segers for a study by Segers (2012). In this task, participants used arrow keys on a specially designed keyboard, to manipulate the audio track until they believed it was in-synch with the video track. Each key press was associated with either a 10ms delay or advance in the audio track relative to the video track, depending on which arrow key one pressed. Participants were instructed which arrow to press, and that they could use the opposite arrow key only if they believed they went beyond the point of synchrony. However, this new *dynamic* task could be improved in several ways which is what the current study aims to do.

Current Study

The current study aims to slightly alter and ameliorate the *dynamic* task invented by Segers (2012) and use it to measure the size of the TBW for commonly occurring audiovisual stimuli. Two major changes will be made to the Segers (2012) *dynamic* task. First, the 10ms increments were very small changes that needed repeated exposure to the asynchronicity of the stimuli and such repeated exposure to asynchronous stimuli has been found to lead to perceptual training and this can bias judgment in the direction of the repeated exposure (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Keetels & Vroomen, 2007; Navarra et al., 2005; Powers et al., 2009; Vroomen, Keetels, de Gelder, & Bertelson, 2004). In simpler terms, this means that the more exposed the participants are to asynchronous stimuli; the more likely they are to shift their TBW towards the asynchrony, meaning that their TBW become wider. Therefore, increasing the increments to 50ms could be advantageous and reduce the amount of repeated exposure, while at the same time making it easier to compare results from the *dynamic* task to the commonly used SJ task, which usually uses audio delays or advances of 50ms increments. Second, the current study aims to expand on Segers' (2012) methodology by incorporating an additional component where participants will not only make asynchronous stimuli synchronous (*toward-synch*) but also to make synchronous stimuli asynchronous (*from-synch*). This will allow us to determine whether the TBW is actually fixed or whether other components, such as in this case, the nature of the task, can influence it. No study to date has used this *dynamic* task in the opposite direction to see if the sizes of windows are constant in both directions. The *from-synch* task may also provide information regarding whether the repeated exposure to asynchronous stimuli does truly play a role in the TBW size by comparing window sizes across the *from-synch* and *toward-synch* tasks. The *from-synch* task should not be influenced by repeated exposure since participants are looking for the first point where they notice that the stimuli goes from being synchronous to not and therefore are not exposed to asynchrony repetitively.

Ultimately, this study will compare performance on both *dynamic* tasks (*toward-synch*, *from-synch*) using both 50ms and 10ms increments against a well-known fixed task (synchrony judgment, a version of the method of constant stimuli) in order to validate this new *dynamic* methodology of measuring audiovisual integration and its possibility of increasing precision for measuring temporal binding windows for audiovisual integration. In order to ensure that the current methodology is comparable to the synchrony judgment tasks, a variety of stimuli will be used to measure the TBW. Since previous studies have demonstrated that the sizes of the TBWs vary based on stimulus types (e.g., Dixon & Spitz, 1980; Stevenson & Wallace, 2013; van Eijk et al., 2008), it is important to check whether the new methodology is sensitive to these differences. The stimuli used for this study are the same that Segers (2012) used in her study. These stimuli varied in terms of social and linguistic content, ranging from social linguistic stimuli to non-social non-linguistic stimuli. The current study will also attempt to determine whether any additional intellectual and behavioural measures are correlated or could predict the size of the TBW across the various stimuli type. In the future, this paradigm may be used to measure the TWB for participants with ASD. Therefore, IQ and ASD characteristics are likely to differentiate future samples and therefore will be briefly examined here.

Developing a more precise way to measure the temporal binding window could be extremely beneficial for many reasons. First, it could provide a more accurate insight into how the sizes of the TBWs differ across individuals with and without sensory integration difficulties (i,e., individuals with ASD). Having a better understanding of how these windows differ across these individuals holds important implications at the level of interventions for individuals with sensory integration difficulties and how to help them discriminate between perceptual information that should and should not be integrated together. Second, as outline by Zampini, Shore, and Spence (2003), having a more precise measure of the TBW could further benefit fields such as the designing of hearing aids (McGrath & Summerfield, 1985), broadcasting, and improving video conferencing technology. Lastly, the current study also aims to determine what can impact and influence the TBW size. Understanding how and why TBWs vary can provide a greater understanding into the mechanism behind audiovisual integration.

Research Questions and Hypotheses

Research question 1. For the *toward-synchrony dynamic* tasks, how do the sizes of the temporal binding windows differ across the two *dynamic* task increments and how do they compare to the typical synchrony judgment task? Which one produces the narrowest windows?

Hypothesis 1. Although this is an exploratory research question and that no research to date has used the 50ms increment dynamic task, it is hypothesized that it will produce narrower windows than the other two tasks. The rationale behind developing a second dynamic task with increments of 50ms rather than 10ms was to reduce exposure to asynchrony in the hope of reducing the amount of perceptual training and therefore creating narrower windows as proposed by Segers (2012). Furthermore, it is hypothesized that the dynamic task with increments of 10ms will yield significantly larger windows than those produced by the synchrony judgment task as found in Segers (2012) as a result of repeated exposure to asynchrony.

Research question 2. For the *from-synchrony dynamic* tasks, how do the temporal binding window sizes differ from the ones obtained from the *toward-synch dynamic* tasks? Do they produce comparable values?

Hypothesis 2. No study, to the author's knowledge, has directly compared the widths of temporal binding windows between tasks that differ in terms of judging synchrony versus asynchrony, as Segers (2012) only used the *toward-synch* method to determine the TBW size. However, it is hypothesized that the TBW sizes will not differ between the *toward-synch* and *from-synch* tasks as both are looking for the point where participants first notice a change from interpreting the visual and auditory information as synchronous to asynchronous or vice versa.

Research question 3. How do the TBWs differ across the different types of commonly occurring stimuli? Which one produces a narrower and more precise window? Are these results replicated across all tasks?

Hypothesis 3. Based on previous studies and the ecological validity of the stimuli, it is hypothesized that commonly occurring stimuli, such as someone telling a story which is considered a social linguistic stimulus, will yield a smaller and more precise window than would a non-social non-linguistic stimulus (e.g., Segers, 2012).

Research question 4. Do measures of intelligence and autism-like traits, such as greater difficulty with relating to others' emotions and difficulty with social and communication skills, correlate with the size of the TBWs across various stimulus types? Do either the verbal or nonverbal abilities correlate with the TBW size?

Hypothesis 4. It is hypothesized that TBW size will be correlated with the level of autism-like traits demonstrated in participants, with participants with greater amounts of autism-like traits having overall wider TBWs. This would coincide with the autism literature which has demonstrated that individuals with ASD are more tolerant of audiovisual asynchrony, especially with speech stimuli (e.g., Bebko et al., 2006; Stevenson et al., 2014b; Taylor et al., 2010). Furthermore, it is hypothesized that verbal IQ ability might be correlated with the TBW of social linguistic stimuli, where one would see participants with better verbal IQ having narrower windows. This is based on the assumption that individuals with better verbal abilities are more likely to be sensitive to the linguistic portion of the stimuli. Previous studies involving clinical populations such as individuals with ASD have shown that children with ASD have greater difficulty noticing asynchrony for linguistic stimulus in comparison to non-linguistic ones (e.g., Bebko et al., 2006). Bebko and colleagues have postulated that these results may be link to the language impairment sometimes noticed in these individuals, which could indicate that verbal abilities may be linked to performance on audiovisual integration of linguistic stimuli.

Methods

Participants

Forty-four adults between 18 and 30 years of age, with the exception of a participant who was 43 years old, participated in this study (M = 20.86, SD = 4.43). The participants were recruited through the Undergraduate Research Participant Pool (URPP) of a large Canadian university where they received course credits for their participation in the study. They were all enrolled in a first year Introduction to Psychology class. As a result, participants were assumed to have average to above average intelligence. This specific age range was selected for multiple reasons. First, this population was readily accessible for this time sensitive study. Second, it has been shown that the size of the temporal binding window is still narrowing during adolescence (Hillock-Dunn & Wallace, 2012), and that much later on in life it widens again as it is wider in the elderly (Diederich et al., 2008). Therefore, participants in their early adult life were selected to ensure that they had stable window sizes. Inclusion criteria included normal or corrected to normal eye vision and normal hearing (as indicated by the participant) due to the fact that the stimuli used in this study all contain audio and visual information that were crucial to perform the task at hand. All participants needed to understand the English language as all the instructions and stimuli were in English, as were all the questionnaires.

Measures

Behavioural and cognitive measures were collected from all participants and were used to determine whether any of them were correlated with or linked in any ways to audiovisual integration skills. Participants also filled out the Adolescent Adult Sensory Profile (AASP; Brown & Dunn, 2002), a questionnaire that measures an individual's sensory processing patterns across different sensory modalities and how these affect every day functioning. However, results from the AASP were not included in the analyses as they are beyond the scope of the current study.

Wechsler Abbreviated Scale of Intelligence- 2nd Edition (WASI-II: Wechsler, 2011). The WASI was used to obtain a measure of the intellectual abilities of the participants in the study. The verbal IQ (VIQ) and the performance IQ (PIQ) were obtained for each participant. Due to time constraints, only one subtest per VIQ and PIQ was used. The VIQ was measured using a Vocabulary task where participants were asked to provide definitions for specific words. Meanwhile, the PIQ was measured using the Matrix Reasoning task in which the participants were asked to reason about shapes and designs and to find a pattern between visual images. Using only two of four subtests to measure IQ has been demonstrated to be a valid representation of VIQ and PIQ (Wechsler, 2011). The WASI-II is a reliable measure of intelligence that can be used for individuals between the ages of 6 up to 90 (Axelrod, 2002) and has been shown to have high correlations with items on the more comprehensive Wechsler Adult Intelligence Scale (WAIS-IV: Wechsler, 2008). The WASI-II was used to determine whether different aspects of intellectual abilities influence the participants' ability to integrate audiovisual stimuli. In all, the two tasks of the WASI-II, took about 15 minutes to complete.

Autism Spectrum Quotient (AQ-short; Hoekstra et al., 2011*).* The AQ is a 28-item self-administered instrument which assesses several key traits that are associated with autism, such as social and communication skills, attention to details, and imagination[1]. The AQ-short was derived from the original Autism Spectrum Quotient questionnaire (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) which had 50 items. The AQ has been used both with adults on the spectrum and non-ASD university students (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen et al., 2001; Hoekstra et al., 2011) and has shown good test-retest reliability and internal consistency (Baron-Cohen et al., 2001; Ruzich et al., 2015). The AQ-short contains two higher order factors that focus on "social behavioural difficulties", which include, social skills, fixation with routines, attentional switching, and imagination, and on "fascination for number and patterns" (Hoekstra et al., 2011). Scores can range between 28 to 112, with scores at or above 70 indicating high levels of autism-like traits. While validating the AQ-Short, Hoekstra and colleagues (2011) found that using scores of 70 or above as a cut-off to distinguish adults with high functioning autism from typically developing adults has a sensitivity of .94 and a specificity of .91. Furthermore, the AQ-short is a valid measure that had been shown to be reliable (Hoekstra et al., 2011). The AQ-short takes on average five minutes to complete. This questionnaire was administered to participants to determine whether there is a correlation between performance on the audiovisual integration tasks and the level of autism-like traits present in the participants. It has been demonstrated that when it comes to audiovisual integration, individuals with a diagnosis of autism spectrum disorder are known to have wider and therefore less precise, temporal binding windows (e.g., Bahrick, 2010; Bebko et al., 2006; de Boer-Schellekens et al., 2013; Keane et al., 2010; Mongillo

[1] Almost half of the participants (n = 21) were given the AQ 50-items questionnaires instead of the short version. However, only the 28 items of the short versions were scored for these participants. Due to time constraint, the short version of the questionnaire was adopted midway through testing to save time.

et al., 2008; Stevenson et al., 2014b; Taylor et al., 2010; Wallace & Stevenson, 2014). Determining whether a similar pattern can be seen in a sample without a formal diagnosis of ASD based on autism-like traits could help advance research in the field and point out specific correlations between symptoms of ASD and multisensory integration.

Empathy Quotient- short version (EQ; Baron-Cohen & Wheelwright, 2004). The EQ is a 40-item self-report questionnaire which measures how empathic one is and takes about ten minutes to fill out. The questionnaire records one's ability to understand and respond to others' emotions, and how one is affected by others' feelings. The EQ was first developed with the aim of identifying lack of empathy in psychopathology and in specific disorders such as ASD, where empathy was deemed difficult. Scores from 0 to 80 can be obtained on the EQ, with scores equal to or below 30 being indicative of low empathy and characteristic of individuals with ASD (Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004). The EQ has high test-retest reliability of .97 when tests are re-taken within a 12 months period (Lawrence et al., 2004). Furthermore, the EQ is inversely correlated with scores on the AQ, demonstrating the link between lower empathy and high levels of autism-like traits (Lawrence et al., 2004). This questionnaire was used as an additional behavioural measure to determine whether there were any links between the temporal binding window size and autism-like traits in the current sample.

Design

Experimental stimuli. Three different types of stimuli were used: a) Social Linguistic (Story) stimuli, b) Non-Social Non-Linguistic (Piano) stimuli and c) Social Non-Linguistic (Sounds) stimuli. The two social stimuli consisted of a video of a woman's face and neck making different sounds (*Figure* 1). In the linguistic stimuli, the actress was telling a story (e.g., "Goldilocks and the Three Bears") and in the non-linguistic stimuli she was making popping and kissing sounds with her mouth. The purpose of the Sounds stimuli was to isolate the social component from the linguistic one to determine whether it had a specific effect on the performance of the participants. Several studies has shown the benefits of audiovisual integration with speech (e.g., Bishop & Miller, 2009; Calver et al., 1998; Girin, et al., 2001; Grant & Walden, 1996; Sumby & Pollack, 1954; Summerfield, 1987) and therefore it was important to have a stimulus that was still social in nature (i.e., showing someone's face) but that did not have any speech components to it. Lastly, the Piano stimuli consisted of a video of an overhead view of a hand playing a simple song on a piano keyboard (*Figure 1*). Two versions of each of the Story and Piano condition were created to prevent participants from becoming bored of the stimuli. Meanwhile, only one version of the Sounds stimuli was used, as it was an exploratory condition and not the main focus of the current study. All stimuli were originally created by Hancock (2009) and were edited by Segers (2012) using the Abode Final Cut Pro program.

Figure 1. Screenshots of the three different stimuli: Story (left), Piano (center), and Sounds (right).

Experimental tasks. Two types of experimental tasks were used during this study: a dynamic task and a commonly-used fixed task. Every participant took part in both of the experimental tasks in order to determine whether the two tasks yield different results across and within participants.

Dynamic task: Manual adjustment task. The manual adjustment task consisted of two parts, going from asynchronous to synchronous (*toward-synch)* and vice versa (*from-synch*), and had two different versions based on the size of adjustable increments: 10ms or 50ms. During the *toward-synch* part, participants were presented with stimuli that have a specific degree of synchrony offset. The stimuli either had an "audio lead", where the audio information preceded the visual information by -400ms, or a "visual lead", where the audio information was lagging behind the visual information by 500ms. These specific starting points were selected because a large enough offset is needed to be present for the participant to want to start pressing the keys, and these offsets were known to be outside the temporal binding window for participants within the study's age group (Dixon & Spitz, 1980). Furthermore, the "audio lead" is smaller than the "visual lead" because individuals are better at detecting asynchrony when the audio information precedes the visual information (Dixon & Spitz, 1980). This increase sensitivity has been reported to be of at least a 100ms, which is why there was this asymmetry between "audio lead" and "video lead" stimuli. Participants were instructed to use arrow keys to adjust the audio track until the first point where they notice that what they were seeing and hearing were synchronous (for full instructions, please refer to the appendices). The video track played at a constant rate while only the audio track could be adjusted. Each key press adjusted the audio track by increments of either 10ms or 50ms, depending on the condition. Again, all participants took part in both the 10ms and 50ms increment conditions. In the *from-synch* task, participants started from a point of synchrony and were asked to adjust the audio track until they first noticed that what they were seeing and hearings were out of synch.

There were four dynamic task conditions, and each participant took part in all four. A single dynamic task condition, for example *Dynamic-50ms toward-synch*, contained a total of 24 trials divided into four blocks. Within each block there were two Story, two Sounds, and two Piano stimuli. All stimuli within one block were either all “audio lead” or all “visual lead”, and the blocks within a single dynamic task alternated between the two so that there was the same amount of audio and visual leading stimuli within one task. For the *Dynamic-10ms* tasks, only two blocks existed, one for each type of lead (visual or auditory). Each block was made up of the same number of stimuli as for the *Dynamic-50ms* conditions. The *Dynamic-10ms* tasks were shortened to only two Blocks because of time restrictions and because data from this condition had previously been collected by Segers (2012) and therefore were available if needed to supplement the current data.

Fixed task: Synchrony judgment task. The fixed task followed the general model of a synchrony judgment task, which has evolved from the field of psychophysics and has been also been referred to as the method of constant stimuli, where participants were shown a series of short videos (the same length and stimuli as the ones presented in the dynamic task) and for each, were asked whether they believed the audio and visual information was synchronous or not. A total of 96 trials divided across four blocks were presented. In all, the participants viewed each of the following audio asynchronies twice for each type of stimulus (Story, Sounds, Piano): ±50ms, ±100ms, ±150ms, ±200ms, ±250ms, ±300ms, ±350ms, and ±400ms. Participants were required to press one of two keys on the keyboard (yes/no) to indicate whether they believed the audio and visual information were synchronous or not. Participants were instructed to make their answer as quickly and accurately as possible.

Demonstration and practice trials. Before the participants were tested on the experimental tasks, they were first shown two demonstration videos and then performed a series of practice trials with slightly different stimuli. These demonstration and practice trials occurred at the beginning of each task, *Dynamic* and *Fixed*, and therefore occurred on five different occasions. All demonstration and practice stimuli were Social Linguistic stimuli created by using a different actress and performing a different task (e.g., counting to ten), than the one used in the experimental stimuli. Different actresses and stimuli were used in order to avoid participants getting familiarized with the experimental stimuli. The two demonstration videos were presented prior to the beginning of a task and the practice trials to show participants what was meant by synchronous and asynchronous videos. Each video was clearly identified beforehand as synchronous or asynchronous and showed either a woman counting to ten with the audio and video files perfectly synchronous or asynchronous by 500ms. Then, the participants did the practice trials. For the *Dynamic tasks*, participants were given two practice trials, one for each arrow key. During the practice trials, participants were first instructed on whether they had to make the audio and video files match or not match, and which arrow key to press. More specifically, for the asynchronous practice trials, a noticeably large asynchrony (e.g.,+500ms) was used to make it obvious to the participants that their task was to make them match and to initiate them to press the arrow keys. Meanwhile, for the *Fixed task*, the participant performed seven of the practice trials with the novel stimuli, where the level of asynchrony varied between -400ms and +500ms, including 0ms. Regardless of the task, the purpose of the practice trials were to ensure that the participants understood the task at hand and provided the experimenter with the opportunity to offer more information if the participant looked confused.

Apparatus. Participants were seated in front and approximately 60cm away from a 26-inch television screen with a resolution of 1280 x 720. The television screen height was adjusted until the participant's eye line fell in the middle of the screen. In front of and below the television screen, there was a Tobii X60 eye tracker, which monitored the participants' eye movements and to ensure that they were actually looking at the stimuli presented to them. For the current study, no analyses were performed regarding the looking pattern of participants while performing the task as these analyses fell outside the scope of this study. However, the Tobii software was used to run the study and to record the participants' answers throughout the experimental task and therefore played an important role in running the study.

For the *dynamic* tasks, the videos were presented in the Media Player Classic- Home Cinema (MPC-HC) program. The open source code of the program was altered from its original state to obtain a second version with 50ms increments in addition to the original, which was set for 10ms increments. For this reason, there was a break between the 10ms and the 50ms dynamic tasks in order for the examiner to change the key press increments. For the *dynamic* tasks, the Tobii program controlled the MPC-HC videos for display; for the fixed task, the videos were presented directly within the Tobii Studio program. For all conditions, the Tobii program monitored looking patterns while at the same time recording the keys being pressed and the frequency of the presses.

An altered computer keyboard, designed by Segers (2012), was used by the participants to enter their answers. Only five keys were available to the participants, the rest of them were removed and covered with a black cardboard (*Figure 2*). The five keys available were identified as follow: “next” (used to indicate that the participants were done adjusting the window and that they want to move on to the next stimuli), “yes” and “no” (to indicate during the fixed task whether they believed the audio and video of the stimulus were in synch), and left and right arrow symbols (used in the dynamic task to make the audio track advanced or delayed by the pre-set intervals).

Figure 2. A sketch of the modified keyboard used to perform the experimental tasks. Only five buttons were available, the two arrow keys, the "YES" and "NO", and the "NEXT" button.

Procedure

Participants came in on one occasion, for a 2 hours testing session. A general explanation of the study was provided before the start of the study and participants were explained the consent form (view *Appendix A*) before obtaining their signature. Each participant started with the experimental task. At the beginning of the experimental tasks, the eye tracker was adjusted to capture the participants' eyes and eye movements. A 5-point calibration was performed to ensure that eye movement would be encoded properly. Participants were administered 2 of the 5 experimental tasks before taking a break to answer the questionnaires (AQ, EQ, and SSP) and the WASI. Afterwards, participants completed the 3 remaining experimental tasks, which required re-calibrating the participant's eye movements. To control for any order effect, participants were randomly assigned to one of four starting condition (task order A to D)[2] (view *Figure 3).* Half of the starting conditions began with one of the 10ms *dynamic* task, while the

[2] During testing, it was noticed that participants seemed to perform differently on subsequent tasks depending on the set of tasks they started with. Since the purpose of the study was to examine the *Dynamic-50ms* tasks, once this was observed, participants were randomly assigned to either task order B or D to prevent task order effect. This will be discussed further in the results and discussion.

other half started with one of the 50ms *dynamic* task. In all cases, participants always ended with the fixed task.

Task Order A	Task Order B	Task Order C	Task Order D
1. Dynamic 10ms *from-synch*	1. Dynamic 50ms *from-synch*	1. Dynamic 10ms *toward-synch*	1. Dynamic 50ms *toward-synch*
2. Dynamic 10ms *toward-synch*	2. Dynamic 50ms *toward-synch*	2. Dynamic 50ms *from-synch*	2. Dynamic 50ms *from-synch*
3. Questionnaires	3. Questionnaires	3. Questionnaires	3. Questionnaires
4. Dynamic 50ms *from-synch*	4. Dynamic 10ms *from-synch*	4. Dynamic 10ms *toward-synch*	4. Dynamic 10ms *toward-synch*
5. Dynamic 50ms *toward-synch*	5. Dynamic 10ms *toward-synch*	5. Dynamic 10ms *from-synch*	5. Dynamic 10ms *from-synch*
6. Fixed task	6. Fixed task	6. Fixed task	6. Fixed task

Figure 3. The four different task order groups with a list of when each tasks were performed.

Results

A total of 44 participants took part in the study. Of those, thirty eight were used in the analyses. Five participants were dropped due to technical malfunctions caused by the eye tracker resulting in no data being recorded for each of those participants. An additional participant was dropped due to uncertainty regarding whether she understood the tasks at hand, which was further supported by her low WASI-II score, which was at the cut-off usually used to identify the clinical range (FSIQ = 70, 2^{nd} percentile). Since her score on the WASI was below the inclusion criterion assumed for the participants in this study (normal to above normal IQ), she was removed from all analyses. Lastly, it is important to note that full sets of data (complete data for all five experimental tasks) were not available for all of the thirty eight participants. Due to time constraints and computer malfunction, a few participants only completed certain tasks (n = 5, in which three participants are missing two tasks, and two are only missing one), while the majority completed all five experimental tasks (n = 33). The distribution of the number of participants who completed each task can be seen in *Table 2*.

Demographics

Demographic information regarding the participants, including gender distribution, and IQ, can be seen in *Table 1*. About 60% of the participants reported that English was their first language and 84% mentioned that the language they spoke the most frequently was English. However, a significant difference between native English speakers and non-native English speakers was observed on the verbal IQ scores, with non-native speakers scoring significantly lower than the native speakers, $F(1,36) = 4.78$, $p < 0.05$. These results are not surprising as the scores for the non-native speakers may be underestimates of their true verbal skills; however they may serve to explain why the sample included in this study had somewhat lower than expected IQ scores. Although, the average FSIQ of the sample is still within the average range, it is lower than what one would expect for a university based sample. However, since the tasks performed in this study were primarily perceptual tasks which focused on matching visual and auditory information, and not semantic based tasks, and that these students were enrolled in a major at an English speaking University, verbal skills were not seen as critical for this study.

Table 1

Demographic information of the participants and differences in IQ between native and non-native English speakers

	Total Sample	Native Speakers	Non-Native Speakers
Age (mean (SD))	20.86 (4.43)	-	-
Gender			
% Male	36.8	-	-
% Female	63.2	-	-
Native English Speakers			
% Yes	60.5	-	-
% No	39.5	-	-
Spoke English most frequently			
% Yes	84.2	-	-
% No	15.8	-	-
IQ			
FSIQ (mean(SD))	94.7(11.48)	96.5(13.29)	92.0(7.79)
VIQ (mean (SD))*	49.0(8.05)	51.3(9.09)	45.7(4.78)
PIQ (mean (SD))	45.0(9.06)	44.9(8.94)	45.3(9.56)

Note. The VIQ and PIQ scores are t-scores, while the FSIQ scores are composite scores, adjusted for age. Significant difference ($p < .05$) between native and non-native English speakers is identified by *.

Normality of Variables

All the dependent variables associated with the temporal binding windows were normally distributed based on the Skewness-Kurtosis normality test (*Table* 2), except for one (Piano stimuli of the *Dynamic-50ms from-synch*). For the *Dynamic-50ms from-synch* condition, there was a significant outlier for the width of the TBW on the Piano stimulus that stood more than two standard deviations above the mean. Once this outlier was removed, the variable met the assumption for normality and therefore the analyses were performed without this one data point. *Table 2* reports the data after the outlier was removed. The Skewness-Kurtosis normality test was selected over the Shapiro-Wilk and the Anderson-Darling tests for normality because of the presence of ties (when identical values are present within a data set) in the data (Pearson, D'Agostino, & Bowman, 1977). Furthermore, Skewness-Kurtosis tests have been shown to have good power properties for smaller samples (n < 50) (D'Agostino, Belanger, & D'Agostino, 1990; Kim, 2013; Razali & Wah, 2011).

The majority of the behavioural measures collected through questionnaires (AQ and EQ) and assessments (VIQ, PIQ, and FSIQ), were also normally distributed. Both of the autism trait measures, the AQ and the EQ, showed no significant skewness or kurtosis meaning that the normality assumption was met (*Table* 2). However, both the FSIQ and the VIQ did not meet the normality assumption. One participant was a significant outlier for both of these variables by scoring at least two standard deviations above the mean (FSIQ = 136). When this participant was removed from all the IQ variables (including the PIQ), all three variables met the normality assumption. Removing the participant from all the analyses did not impact the significance of the findings and therefore he was kept in the analyses. Furthermore, since the participants' cognitive abilities played a secondary role in the purpose of this study, it was deemed unnecessary to remove the participant from all other analyses based on the fact that he has a much higher IQ. Given the normality of the variables, parametric tests were used to analyze the data. For the non-normal data, the FSIQ and VIQ, non-parametric tests such as Spearman correlations were used when these variables were included in the analyses.

Table 2

Descriptive statistics and Skewness-Kurtosis of the dependent variables

	n	Mean (SD)	Min	Max	Skewness (SE)	Kurtosis (SE)
TBW width Toward-synch						
Story: Dynamic 50ms	37	299 (157)	17	583	-.05(.39)	.84(.76)
Sound: Dynamic 50ms	34	383(178)	50	725	-.18(.40)	-.86(.79)
Piano: Dynamic 50ms	36	374(174)	50	738	.18(.39)	-.47(.77)
Story: Dynamic 10ms	35	490(183)	50	810	-.60(.40)	-.22(.78)
Sound: Dynamic 10ms	35	571(167)	235	825	-.62(.40)	-.61(.78)
Piano: Dynamic 10ms	36	539(159)	270	850	.25(.39)	-.86(.77)
Story: Fixed ł	32	480(116)	150	650	-.88(.41)	1.32(.81)
Sound: Fixed	30	560(102)	300	700	-.50(.43)	-.01(.83)
Piano: Fixed	30	532(119)	300	800	.09(.43)	-.10(.83)
TBW width From-synch						
Story: Dynamic 50ms	36	659 (233)	160	1233	.11(.39)	.24(.77)
Sound: Dynamic 50ms	36	888 (327)	300	1700	.58(.39)	-.10(.77)
Piano: Dynamic 50ms	35	768(282)	50	1438	.16(.40)	.69(.78)
Story: Dynamic 10ms	36	684(245)	40	1310	.29(.39)	.93(.77)
Sound: Dynamic 10ms	37	771(342)	55	1585	.69(39)	.21(.76)
Piano: Dynamic 10ms	37	715(332)	15	1630	.49(.39)	.96(.76)
Point of Subjective Synchrony						
Story: Delayed Trials	37	233(115)	17	450	-.16(.39)	-.92(.76)
Sound: Delayed Trials	36	243(114)	25	450	-.20(.39)	-.99(.77)
Piano: Delayed Trials	37	245(106)	50	450	.05(.39)	-.67(.76)
Story: Advanced Trials	37	-65(57)	-200	0	-.67(.39)	-.50(.76)
Sound: Advanced Trials	35	-134(81)	-288	0	-.09(.40)	-.79(.78)
Piano: Advanced Trials	36	-128(80)	-288	0	-.38(.39)	-.57(.77)
Degree of Asymmetry						
Story: Asymmetry	37	168(92)	-75	358	-.41(.39)	.15(.76)
Sound: Asymmetry	34	119(79)	-63	263	-.51(.40)	-.47(79)
Piano: Asymmetry	36	117(73)	-25	304	-.11(.39)	.23(.77)
Behavioural Measures						
AQ	38	60.8(7.4)	45	76	-.25(.38)	-.25(.75)
EQ ł	38	42.6(11.6)	20	64	-.01(.38)	-1.07(.75)
VIQ*	37	49.0(8.1)	34	80	1.28(.39)	5.01(.76)
PIQ	37	45.0(9.1)	30	62	.38(.39)	-1.03(.76)
FSIQ*	37	94.7(11.5)	75	136	1.28 (.39)	3.28 (.76)

Note. * indicate variables that violate the assumption of normality. ł Indicate variables that despite having skewness and kurtosis values outside of +1.00 and -1.00, are still considered to have approximately normal distribution. All measures of audiovisual integration ability are mean values across similar trials. The Points of Subjective Synchrony and Degrees of Asymmetry are in reference to the *Dynamic-50ms toward-synch* task only.

Data Trimming

Data trimming was performed on single trials in order to have a more precise representation of the TBW. Across all five tasks, any single trial where a participant provided no key press was removed from the analyses. In terms of the *fixed* task, a no key press response provided no insight into whether the participant believed the stimuli presented during that trial was synchronous or not. Meanwhile, for both the *from-synch* and *toward-synch* tasks, it could be argued that no key press reflected the participant's opinion that the stimulus was already asynchronous (or synchronous if referring to the *toward-synch* tasks). However, given the fact that stimuli in the *from-synch* tasks started at true synchrony on the one hand, and the large offsets each stimuli started from in the *toward-synch* tasks, a no key press response was not interpretable unless the participant was not paying close attention. Therefore, each trial with no key press was removed in an attempt to obtain the most meaningful results.

Within the *toward-synch* tasks, single trials were removed if the participants passed the point of absolute synchrony (0 ms). The reasoning behind this data trimming was that previous research showed that typically developing adults should have perceive synchrony much before the absolute synchrony point (e.g., -130ms or +250ms (Dixon & Spitz, 1980)) and therefore participants who have gone beyond that are likely not paying close attention to what they were doing. By performing this strict data trimming, "speeders", participants who were carelessly performing the tasks as quickly as possible in order to be done faster, were removed and their assumed inaccurate results did not bias the rest of the data. Since the *from-synch* tasks started from the absolute synchrony point, the same data trimming did not apply to these two tasks; single trials were deleted only if the participant went the wrong way and pressed the wrong arrow key. *Table 3* shows how many trials were removed due to data trimming for each of the tasks. As can be noted, the *toward-synch Dynamic-50ms* task is the one where the greatest number of trials across all participants was lost. It is important to note that the *Dynamic-50ms* conditions had double the number of trials than the *Dynamic-10ms* conditions. Furthermore, all the values calculated from the experimental tasks are computed by averaging values across multiple trials and therefore losing some trials does not impact the results as heavily. None of the participants lost over half of their trials within the *Dynamic-50ms toward-synch* condition, with the exception of two participants who lost exactly half. However, removing these two participants from the analyses did not significantly impact any of the findings, with the exception of a few correlations and therefore, they were kept in the analyses. The affected correlations will be presented both with and without the two participants.

Table 3

Number of trials before and after data trimming across the different conditions

	Dynamic-50ms		*Dynamic-10ms*		*Fixed*
	from-synch	toward-synch	from-synch	toward-synch	toward-synch
Before data trimming	864	888	444	432	3264
After data trimming	839	730	439	420	3192

Point of Subjective Synchrony

Points of subjective synchrony (PSS) were operationally defined slightly differently based on the tasks. In terms of the *Dynamic toward-synch* tasks, the PSS was determined based on the degree of offset at which the participant first decided that the video and audio information were in sync with each other. Trials from the same stimulus type (e.g., within Story stimuli) were averaged together to obtain the PSS and these were done separately for delayed and advanced trials. Previous studies conducted in the lab using the same stimuli demonstrated that it was appropriate to combine stimuli within the same type as no significant difference existed between them (e.g., there is no significant difference in PSS between the stimuli of the woman telling the

"Goldilocks and the Three Bears" story versus the one telling an extract from "Clifford") (Hancock, 2009; Segers, 2012). For the *from-synch* task, the PSS was based on the last time the participant judged the audio and visual information to be synchronous; in other words, the point just before they said the stimulus was asynchronous.

For the fixed task, the PSS was calculated slightly differently. A PSS was determined for each of the delayed and advance trials. Each PSS was determined based on the following criteria: (1) the first time that the participant pressed the "yes" key on one of the two trials of a specific stimulus onset asynchrony (SOA) (recall that each SOA is presented twice within the fixed task), and (2) this must not be followed by "no" responses on both trials of an SOA that is closer to true synchrony point of 0ms (view *Figure 4*). These criteria have been commonly used for SJ tasks (e.g., Segers, 2012).

PSS (Advanced) ↓ (at -150) **PSS (Delayed)** ↓ (at 250)

SOA	**-400**	**-350**	**-300**	**-250**	**-200**	**-150**	**-100**	**-50**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**
SJ #1	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No
SJ #2	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No

Figure 4. Example of the PSS and TBW (shaded grey region) for the *Fixed* task.

Temporal Binding Window Width

The TBWs for each of the tasks were constructed separately using the PSS of the delayed and advanced trials. Each stimulus type within each task had its own TBW. This can be visualized as putting each of the delayed and advanced PSS for one type of stimulus on the opposite ends of a continuum, with zero in the centre (view *Figure 9*). The width of the TBW was obtained by summing the absolute values of the delayed and advanced PSS. This is the principal approach that was used to examine the differences in TBW size between the various conditions and stimuli.

Analyses of the Temporal Binding Window Width

Toward-synch tasks. A 3x3 repeated measures analysis of variance (ANOVA) was conducted to examine differences in TBW widths between conditions types (*Dynamic-10ms, Dynamic-50ms,* and *Fixed)* and stimulus type (*Story, Sounds,* and *Piano*) within the *toward-synch* task. Since this is a repeated measures model, only participants with full sets of data in each conditions were kept for the analysis (n =25). In order to correct for the violation of sphericity on the interaction between stimuli type and condition, a Greenhouse-Geisser adjustment on the degrees of freedom was adopted for that analysis only. Results yielded significant main effects for both stimulus type, $F(2, 48) = 15.73$ $p < .001$, $\eta_p^2 = .40$, and condition, $F(2, 48) = 27.39$, $p < .001$, $\eta_p^2 = .53$, but no significant interaction, $F(2.81, 67.41) = 0.94$, $p = .44$. These results showed that the width of the TBW is influenced by the task and the stimulus type. Post-hoc analyses using a Bonferonni correction showed that the *Dynamic-50ms* (*M*=379.51) condition produced smaller, more precise windows, than both the *Dynamic-10ms* (*M*=563.60) and *Fixed* (*M*=536.60) conditions, which did not significantly differ from each other (*Figure 5*). Furthermore, the post-hoc analyses also showed that the Story stimulus (*M*=445.49), generated more precise TBWs than both the Sounds (*M*=524.07) and the Piano (*M*= 509.56) stimuli, which did not significantly differ from each other (*Figure 5*). Together, these results demonstrate that regardless of the stimuli type, using a *Dynamic* task, where the participant can adjust the sound track by increments of 50ms, produces narrower and more precise windows than when they were using the 10ms increment or the classic synchrony judgment task, referred to in this study as the *Fixed* condition. Lastly, across all three conditions, the linguistic stimuli

(Story) yielded narrower and more precise windows in comparison to the other two non-linguistic stimuli (Sounds and Piano).

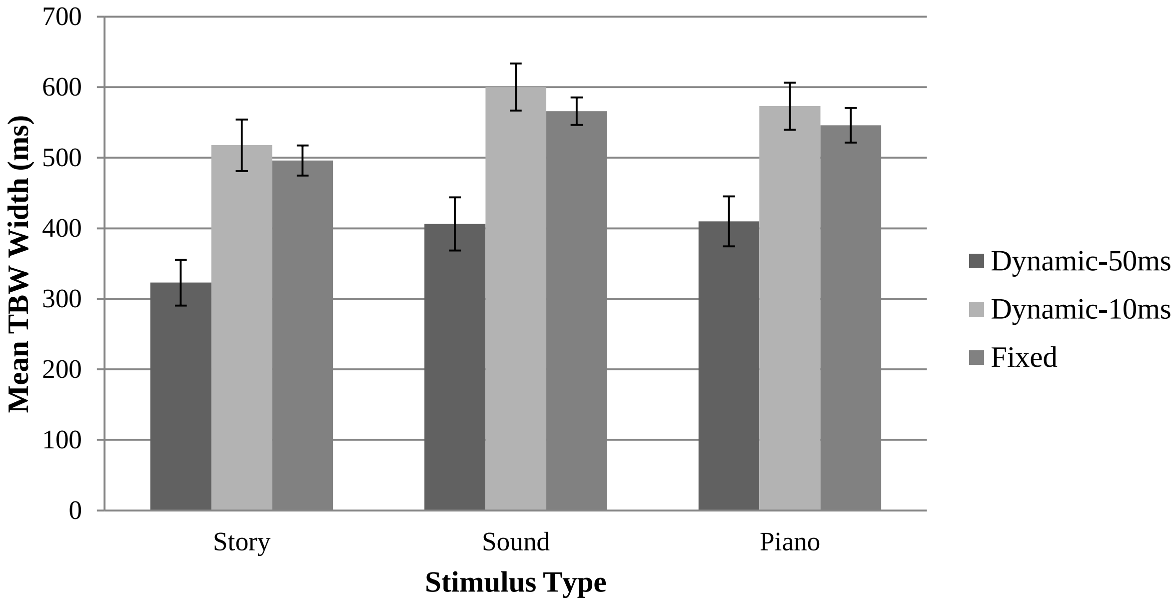

***Figure 5*.** The graph illustrates the mean temporal binding window width of *toward-synch* tasks for each of the stimuli (Story, Sound, and Piano) across the different conditions (*Dynamic-50ms, Dynamic-10ms, Fixed*). Error bars represent standard error of the mean.

Task order effect. The way the study was designed, in order to counterbalance for task order, participants either first started with the set of *Dynamic-10ms* tasks (*toward-synch* and *from-synch*) or the set of *Dynamic-50m* tasks, had a break and did all the behavioural and cognitive measures, then did the other *Dynamic* tasks set and always finished with the *Fixed* task. During testing it was noticed that participants' behaviour and performance on the second set of tasks varied based on which task they performed in the first set. More specifically, those participants who started with the *Dynamic-10ms* seem to use more key presses when doing the *Dynamic-50ms* task, creating larger windows, than those who started with the *Dynamic-50ms* task. It appeared that the greater number and rapid rate of key presses required in the *Dynamic-10ms* conditions primed a high rate of key pressing that carried over into the *Dynamic-50ms* conditions, when the *Dynamic-50ms* tasks were presented second. In order to optimize the number of subjects with unaffected *Dynamic-50ms* tasks, which were the focus of this study, more participants were assigned to start with one of the *Dynamic-50ms* (n=25) compared to the *Dynamic-10ms* (n=13) tasks, creating an unbalanced design. However, since the majority of the analyses are purely within subject designs, task order was controlled within participants in most analyses.

The main area of concern where the task order could have a greater impact on the results is when looking at data across multiple conditions. Therefore, the above 3x3 repeated measures ANOVA of condition and stimulus type was repeated separately for both the participants who started with one of the *Dynamic 10-ms* tasks, and those who started with one of the *Dynamic-50ms* tasks. Results for the individuals who started with one of the *Dynamic-10ms* (n = 7) task yielded a significant main effect of condition, $F(2, 12) = 16.74$, $p < .001$, $\eta_p^2 = .74$. Similar results with regards to the main effect of condition were found for the individuals who started with one of the *Dynamic-50ms* task (n=18), $F(2, 34) = 18.52$, $p < .001$, $\eta_p^2 = .52$. In both cases, the results demonstrated that the *Dynamic-50ms toward-synch* task generated the smallest TBW widths (*Figure 6*). Furthermore, the group that started with one of the *Dynamic-50ms* also had a significant main effect of stimuli, $F(2,34) = 19.69$, $p <.001$, $\eta_p^2 = .52$, where the Story stimuli were linked to smaller TBWs in comparison to the Sound and Piano stimuli which did not differ from each other. The purpose of these analyses is to confirm that the main effect of condition was still present and, therefore, it was justified to merge both groups together for the rest of the analyses. The potential implications of the task order effect will be discussed in more depth in the discussion.

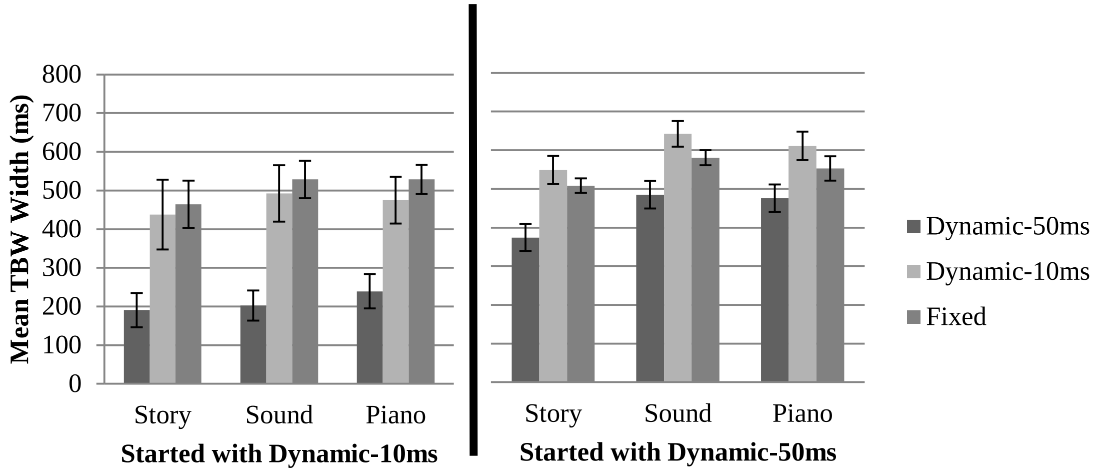

Figure 6. The graph illustrates the mean temporal binding window width of *toward-synch* tasks for each of the stimuli (Story, Sound, Piano) across the different conditions (*Dynamic-50ms*, *Dynamic-10ms*, *Fixed*) between participants who started with one of the *Dynamic-10ms* tasks (left) and those who started with one of the *Dynamic-50ms* tasks (right). Error bars represent standard error of the mean.

From-synch versus toward-synch tasks. The current study also investigated whether using the *from-synch* task would yield comparable TBW width than the *toward-synch* task. To analyse this, two 2x3 repeated measures ANOVA were conducted to examine the difference in TBW sizes between task types (*from-synch*, and *toward-synch)* and stimulus type (*Story, Sounds*, and *Piano*) within each condition (*Dynamic-10ms* and *Dynamic-50ms*) separately. The analyses were separated by condition as it had been shown that the *Dynamic-10ms* and the *Dynamic-50ms* differed from each other in the *toward-synch* task. Stimulus type was included in the analyses since previous analyses within this study showed differences in TBW size based on stimulus type and therefore it would be inaccurate to collapse TBW size across all three types of stimuli.

Dynamic-10ms. The results of the 2x3 repeated measures ANOVA yielded significant main effects of both stimulus type, $F(2, 64) = 5.93$, $p < .01$, $\eta_p^2 = .16$, and task, $F(1, 32) = 8.25$, $p < .01$, $\eta_p^2 = .21$, but not a significant interaction, $F(2, 64) = 0.34$, $p = .87$. No violation of sphericity was reported and therefore no adjustment was performed. Post-hoc analysis using a

Bonferroni correction showed that regardless of the task, the Story stimulus (*M*=595.68) produced smaller windows than the Sound stimulus (*M*=670.68), but did not differ from the Piano stimulus (*M*=627.12) (*Figure 7*). Meanwhile, the Piano and Sound stimuli did not differ from each other. Furthermore, the post-hoc analyses also showed that the *toward-synch* task yielded much smaller and more precise TBWs (*M*=548.23), then the *from-synch* task (*M*=714.09) (*Figure 7*). These results showed that the *from-synch* and *toward-synch* tasks are not comparable and that the *from-synch* task yields much wider windows.

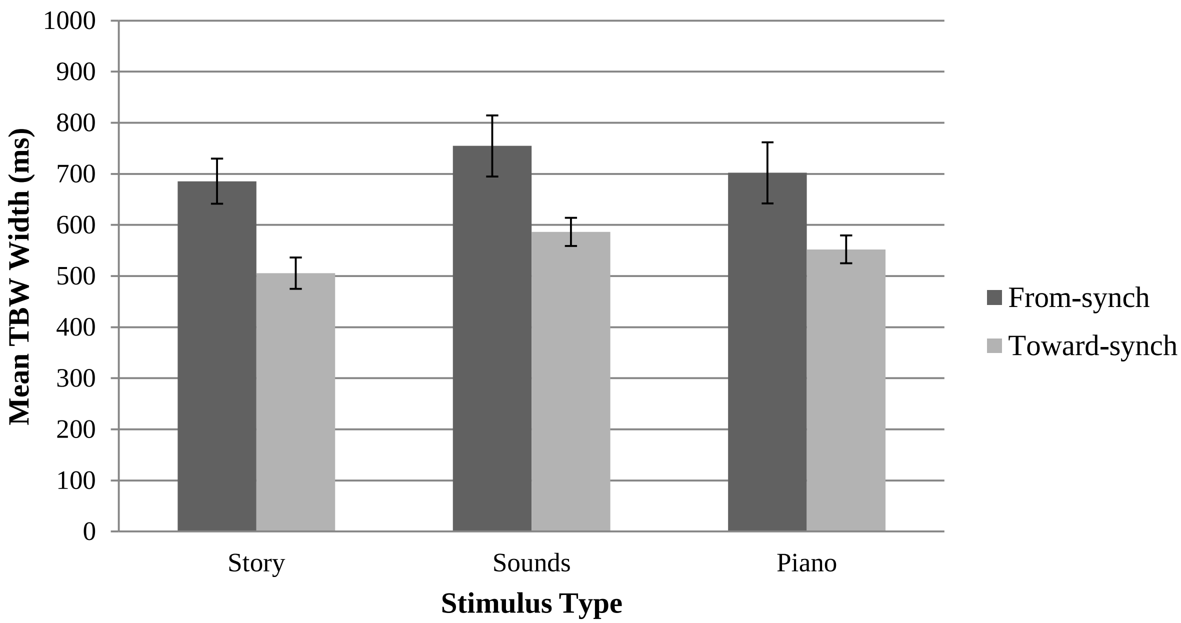

Figure 7. Graph comparing the mean TBW width of the *from-synch* and *toward-synch* tasks of the *Dynamic-10ms* condition across all three types of stimuli. Error bars represent standard error of the mean.

Dynamic-50ms. The 2x3 repeated measures ANOVA yielded a significant interaction of stimulus type and tasks, $F(2, 60) = 9.33$, $p < .001$, $\eta_p^2 = .24$, along with significant main effects of both stimulus type, $F(2, 60) = 32.03$, $p < .001$, $\eta_p^2 = .52$, and task, $F(1, 30) = 75.98$, $p < .001$, $\eta_p^2 = .72$. Since there was no violation of sphericity, no adjustment was performed. Post-hoc analyses using a Bonferroni correction showed that while all three stimuli generate significantly different windows within the *from-synch* task, with the Story stimuli (*M*=657.10) generating the narrowest TBWs, only the Story stimuli (*M*=312.50) within the *toward-synch* task significantly differed from the other two stimuli (Sounds: *M*=397.98; Piano: *M*=384.68), which did not differ from each other (*Figure 8*). Furthermore, the results showed that across all stimuli, the *toward-synch* task (*M* = 365.05) yielded much smaller windows than the *from-synch* task (*M*=770.42) (*Figure 8*), similar to the *Dynamic-10ms* condition. The difference in window sizes between the *from-synch* and the *toward-synch* was much larger in the *Dynamic-50ms* condition than in the *Dynamic-10ms* condition.

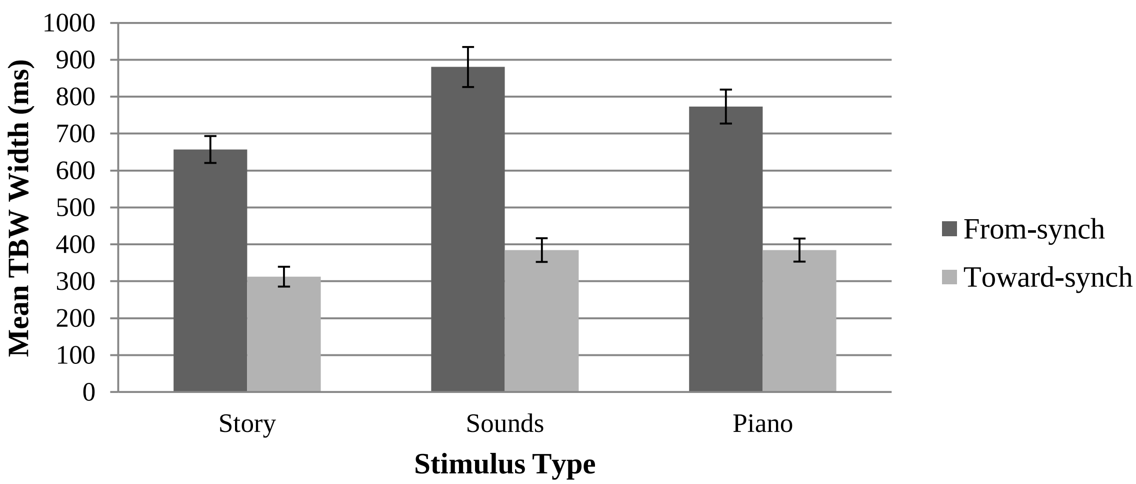

Figure 8. Graph of the TBW width differences for the *Dynamic-50ms* condition across both the *from-synch* and *toward-synch* tasks across all three types of stimuli. Error bars represent standard error of the mean.

Results from both the *Dynamic-10ms* and *Dynamic-50ms* analyses demonstrated that the dynamic *from-synch* task does not yield the most precise TBW width, and that the *toward-synch* task is better. Therefore, all future analyses will only focus on data obtained in *toward-synch* tasks. This difference is discussed in details in the discussion section. Furthermore, results from the previous section also demonstrated that the *Dynamic-50ms* condition produced the narrowest and most precise TBWs in comparison to the other two tasks. Therefore, future analyses, which focus on gathering more information regarding the new methodology and on audiovisual integration, will focus solely on data from the *toward-synch Dynamic-50ms* task.

Analyses of Individual PSS within Toward-synch Dynamic-50ms

The width of the TBWs is composed of the PSS of both the delayed and advanced trials. In order to obtain a deeper understanding of the TBW it is important to look at both of these trial types separately within the condition that yielded the smallest most precise TBWs. As past research has demonstrated, a certain level of asymmetry exists within the TBW, where individuals are better at detecting asynchrony when auditory information comes before the visual information (e.g., Hillock et al., 2011).

Audio advanced trials. A one-way repeated measures ANOVA was conducted to examine the difference in the PSS of advance trials within the *Dynamic-50ms toward-synch* task across the three types of stimulus (Story, Sounds, and Piano). The analysis was significant $F(2,68) = 41.60$, $p < .001$, $\eta_p^2 = .55$ showing that different stimuli yielded different PSS. A post-hoc analysis with a Bonferroni correction showed that the Story stimulus (M = -67.52) generated a PSS that was closer to the absolute point of synchrony (0ms), compared to the Sound (M = -134.05) and Piano (M = -130.24) stimuli, which did not differ from each other (*Figure 9*). These results with the PSS of advanced trials only replicated what was shown in the TBW width analyses for the *Dynamic-50m toward-synch* task.

Audio delayed trials. The same analysis used for the audio advance trials was replicated with the PSS of the audio delayed trials. The results were not significant, $F(2,70) = 0.98$, p =

.379, indicating that the PSS did not differ across stimulus type for the audio delayed trials (Story: (*M* = 227.43), Sound: (*M*=243.40), and Piano: (*M*= 240.63) (*Figure 9*). These results demonstrate that when the auditory information is delayed in comparison to the visual information, the PSS does not seem affected by stimulus type.

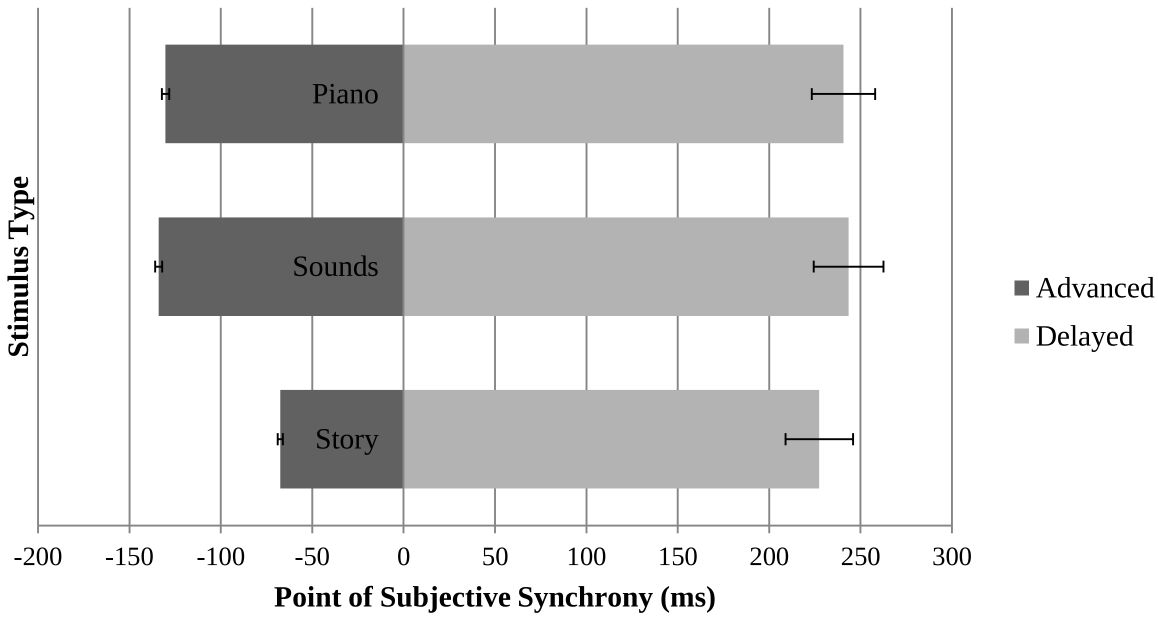

Figure 9. The graph depicts the PSS for both the advanced (left) and the delayed trials (right) across all three stimuli for the *Dynamic-50ms toward-synch* task. The zero represents the point of absolute synchrony between the audio and visual information. The error bars represent standard error from the mean.

Degree of asymmetry. The degree of asymmetry was obtained by subtracting the PSS of the delayed trial from the PSS of the advanced trials. This number demonstrate how positively skewed the TBWs are for the different stimuli. Again, a one-way repeated measures ANOVA was conducted to examine the difference in the degree of asymmetry across the three stimuli. There was a significant effect, $F(2, 66) = 10.14$, $p <0.001$, $\eta_p^2 = .26$, demonstrating that the degree of asymmetry varied based on stimulus type. A post-hoc analyses with a Bonferroni correction revealed that the Story stimulus had a more positively skewed asymmetry (M = 167.40) in comparison to both the Sound (M = 119.00) and Piano (M = 114.83), which did not differ from each other (view *Figure 9*). Taken together, the results from the degree of asymmetry analyses demonstrated that the smaller TBW width of the Story stimulus in comparison to the Sounds and Piano stimuli is mainly driven by participants' PSS on the advanced trials and not the ones from the delayed trials. This is consistent with previous findings and will be explained in greater details in the discussion.

Response Time

The time it took for participants to find the point of subjective synchrony on each trial were recorded, in milliseconds (ms), and analysed in order to determine whether a relationship exists between the time it took them to reach the PSS and how precise their audiovisual integration skills are. Response time refers to the time from the first key press until the time the participant pressed the "Next" key in order to move to the next trial. A response time for each trial was collected and then averaged across stimuli within the same type in order to obtain an average response time.

A 2x3 repeated measures ANOVA was performed to examine the differences in response time across trial type (Delayed or Advanced) and stimulus type (Story, Sound, Piano). It is important to note that the response times for the Sound stimuli of the advanced trials were not normally distributed (Table 4) and therefore results should be interpreted carefully as a parametric test was used on at least one non-parametric variable. In order to correct for the violation of sphericity, a Greenhouse-Geisser adjustment on the degrees of freedom was adopted. Results yielded a significant interaction between stimulus and trial types, $F(1.51, 46.76) = 8.68$, $p < 0.01$, $\eta_p^2 = .22$, and a significant main effect of stimulus type, $F(1.52, 47.01) = 8.94$, $p <$

0.01, η_p^2 = .22 , but not of trial type $F(1, 31) = 0.53$, $p = .474$. Post-hoc analyses revealed that the response time did not differ much across the three types of stimuli within the delayed trials (Story: (*M* = 13988.32 ms), Sounds: (*M* = 13928.32 ms), Piano: (*M* = 13162.40 ms)) (*Figure 10*). Meanwhile for the advanced trials, the participant took the least amount of time to find the PSS for Story stimuli (*M* = 10670.343 ms) and the most time for the Sounds stimuli (*M* = 16708.90 ms) and the Piano stimuli was in between the two (*M* = 15219.05 ms) (*Figure 10*). In all, participants were the fastest at finding the point of subjective synchrony on the Story stimuli of audio advanced trials. In terms of the main effect of stimulus type, across both trials, participants were faster at finding the PSS for the Story stimuli (*M* =12329.33 ms) in comparison to the Sounds (*M* =15318.61 ms) and Piano (*M* =14190.70 ms) stimuli, which did not differ from each other.

Table 4

Descriptive statistics and Skewness-Kurtosis of the response time of various dependent variables

	n	Mean (SD)	Min	Max	Skewness (SE)	Kurtosis (SE)
Response Time						
Story: Advanced Trials	35	10627(6272)	2120	25080	.59(.40)	-.70(.78)
Sound: Advanced Trials *	35	17098(10690)	3440	59101	1.69(.40)	5.75(.78)
Piano: Advanced Trials ł	36	14974(7066)	3776	29779	.11(.39)	-1.12(.77)
Story: Delayed Trials	37	14160(8358)	3669	36138	.80(.39)	.08(.80)
Sound: Delayed Trials	36	14946(9190)	2776	41146	.69(.39)	.18(.77)
Piano: Delayed Trials	37	13121(6656)	3824	28244	.47(.39)	-.49(.76)

Note. * indicate variables that violate the assumption of normality. ł Indicate variables that despite having skewness and kurtosis values outside of +1.00 and -1.00, are still considered to have approximately normal distribution. Response time are measured in milliseconds.

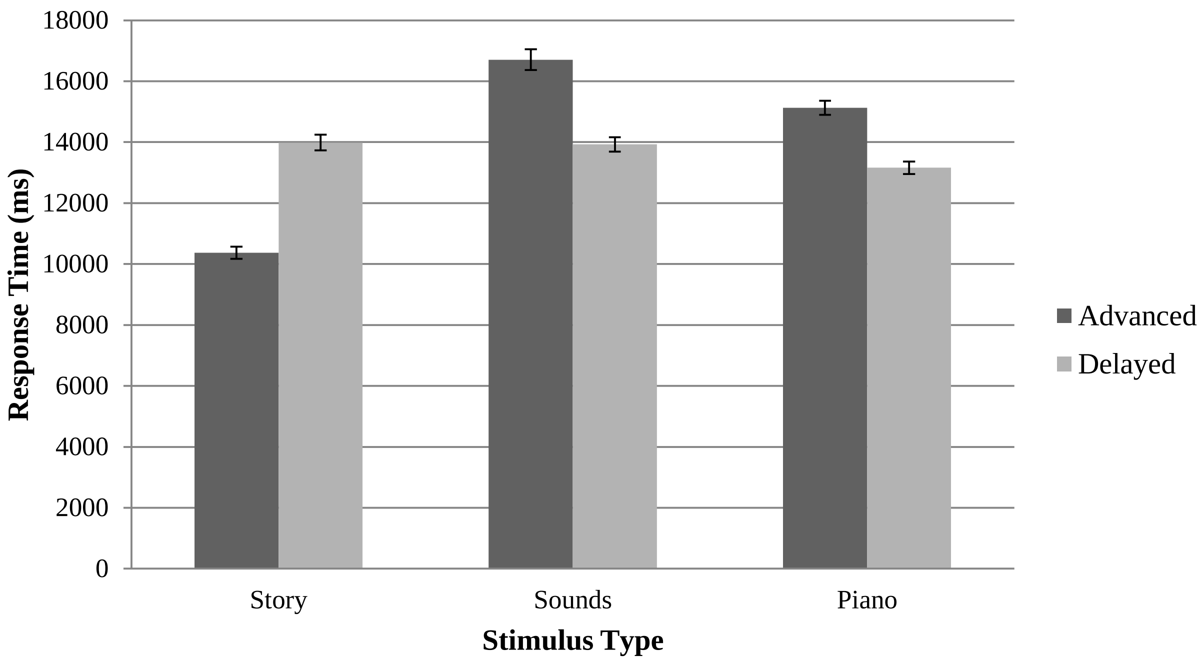

Figure 10. Graph of the average response time in milliseconds within advanced and delayed trials separately and across stimulus type for the *toward-synch Dynamic-50ms* task. Error bars represent the standard error means.

Correlations across response time. Correlations among response times across all three types of stimuli within and across advanced and delayed trials were conducted. The purpose of this analysis was to observe whether there was a consistent relationship between response times across all trials, which would indicate that regardless of the precision of the PSS, participants who took longer on one type of trials also took longer on all other types of trials. A mixed of Spearman and Pearson correlations were used to adjust for the response time of the Sound stimulus for advanced trials, which was not normally distributed. Results yielded significant interactions at the alpha significance level of .05 across all six variables, $r > 0.42$ (*Table* 5). Within the delay trials, correlations across the three types of stimuli ranged from $r = 0.65$ to $r = 0.80$. Meanwhile, within the advanced trials the correlations were even stronger and ranged from

$r = 0.81$ to $r = 0.87$. These strong correlations indicate that the amount of time the participants took to find the PSS are all highly correlated across trial type and stimuli.

Table 5

Correlations of response time across all stimulus and trial types

	1 (N)	2 (N)	3 (N)	4 (N)	5 (N)	6 (N)
1. Story: Advanced Trials	-					
2. Sound: Advanced Trials	.65(33)	-				
3. Piano: Advanced Trials	.78(34)	.80(35)	-			
4. Story: Delayed Trials	.49(35)	.80(35)	.74(36)	-		
5. Sound: Delayed Trials	.42(34)	.86(34)	.67(35)	.84(35)	-	
6. Piano: Delayed Trials	.52(35)	.90(35)	.78(36)	.87(37)	.81(36)	-

Note. Spearman's rho are reported for Sound advance trials variable, the remaining are Pearson correlation coefficients. All correlations were significant at $p < 0.05$

Correlations with PSS. Correlations between response times and varying measured of the audiovisual (AV) integration skill (i.e., PSS for delayed and advanced trials) were performed to determine whether there was any relation between accuracy and response time. Only correlations between variables measuring the same stimuli within the same types of trial were measured. Pearson correlations were used for all analyses with the exception of any correlation with the one variable that was not normally distributed (response time of Sound stimuli on advanced trial), in which case, Spearman correlations were used. The majority of the correlations were non-significant (*Table* 6). The only significant results found concerned the Story stimuli. First, it was found that the response times of the advanced trials were negatively correlated with the PSSs, $r(35) = -0.37$, $p < .05$, indicating that as participants' PSS moved further away from the true synchrony point on the Story stimulus, they took longer to respond (*Figure 11)*. However, the opposite was found on delayed trials where participants with PSSs closer to the true synchrony took longer, indicated by the significant negative correlation between the PSSs and response time on the Story stimulus of the delayed trials, $r(37) = -0.39$, $p < .05$, *(Figure 12)*. These opposite results will be discussed in greater details in the discussion.

Table 6

Correlations of response time (RT) and PSS

	RT of Advanced Trials		
	Story	Sound	Piano
PSS (n)	**-.37(35)***	-.04(35)	-.22(36)

	RT of Delayed Trials		
	Story	Sound	Piano
PSS (n)	**-.39(37)**	-.24(36)	-.11(37)

Note. * $p < 0.05$. Spearman's rho are reported for Sound advance trials variable, the remaining are Pearson correlation coefficient.

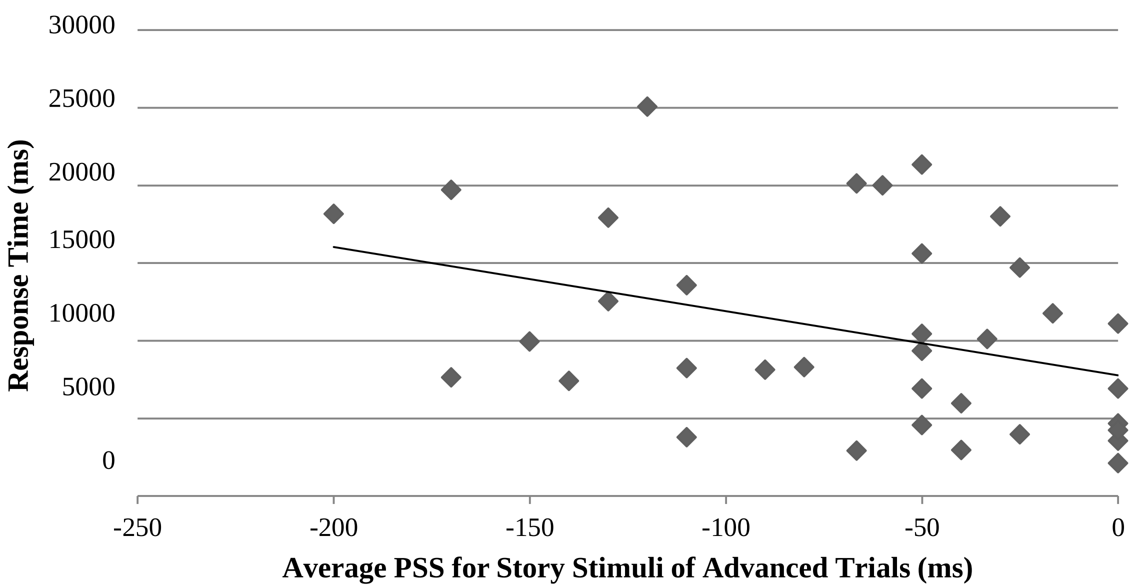

Figure 11. The graph shows the negative relationship between response time and the average point of subjective synchrony for the Story stimulus on the **advanced trials** of the *toward-synch Dynamic-50ms* task.

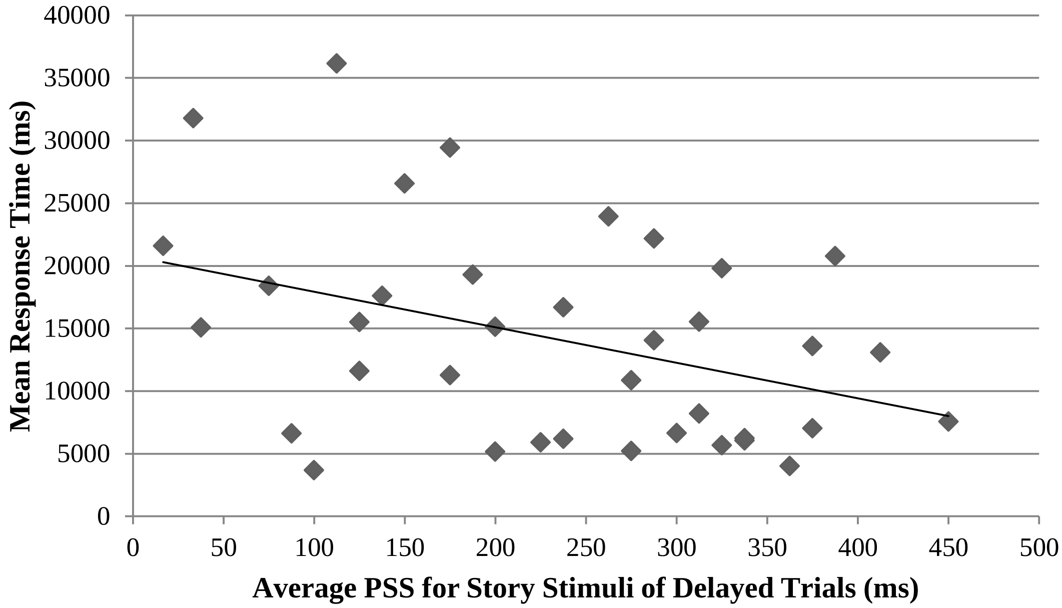

Figure 12. The graph shows the negative relationship between response time and the average point of subjective synchrony for the Story stimulus on the **delayed trials** of the *toward-synch Dynamic-50ms* task.

When the data from the two participants who lost half of their *toward-synch Dynamic-50ms* trials due to data trimming were removed, slightly different significant results were found. First, the significant negative correlation between response time and PSS on the Story stimulus of the advanced trials became non-significant, $r(33) = -0.34$, $p = .052$. However, this is most likely the results of a decrease in power from losing two data points as the results are still marginally significant. On the other hand, the correlation between response time and PSS on the Story stimulus of the delayed trials remained significant and became stronger despite the decrease in power indicating that this may a more robust correlation than the one found for the advanced trials, $r(35) = -0.41$, $p < .05$.

Cognitive Abilities and Autism Characteristic Traits

In addition to performing the two *Dynamic* and *Fixed* tasks, participants' intellectual ability and degree of autism-like traits were measured using an assessment and self-report questionnaires, respectively. Previous research had demonstrated that individuals with ASD tend to have larger TBWs, and therefore it was predicted that participants with a higher degree of autism-like traits would have larger temporal binding windows or PSSs that are further from the true synchrony point. Pearson correlations between the EQ and AQ and the various measures of audiovisual integration skills (i.e., PSS of delayed and advanced trials, and width of the TBW) yielded two significant results (Table 7). First, it was observed that there was a medium positive relationship between scores on the autism quotient and the width of the TBWs for the Sound stimuli, where participants who showed a higher degree of autism-like traits had wider and less precise windows, $r(34) = 0.38$, $p < .05$ (*Figure 13*). This correlation was not sustained once the two participants most greatly affected by the data trimming were removed, $r(32) = 0.33$, $p = .06$, probably due to the decrease in power from removing two data points. Furthermore, a medium negative correlation between the AQ score and the PSS of advanced trials for the Sound stimuli was also significant, again demonstrating that participants with higher levels of autism-traits were more likely to have PSS further away from the true synchrony, $r(35) = -0.42$, $p < .05$. These results remained even after removing the two participants, $r(33) = -.37$, $p < .05$. With regards to the intellectual measures, no significant correlations were found between any of the measures of audiovisual integration abilities and the VIQ, PIQ and FSIQ (Table 8).

Table 7

Correlations of autism traits and audiovisual integration skill

	TBW Width		
	Story	Sounds	Piano
AQ(n)	.24(37)	**.38(34)***	.28(36)
EQ(n)	.12(37)	-.10(34)	-.01(36)
	PSS of Delayed Trials		
	Story	Sounds	Piano
AQ(n)	.18(37)	.23(36)	.23(37)
EQ(n)	.15(37)	-.04(36)	.02(37)
	PSS of Advanced Trials		
	Story	Sounds	Piano
AQ(n)	-.30(37)	**-.42(35)***	-.28(36)
EQ(n)	-.01(37)	.17(35)	.04(36)

Note. *$p < 0.05$. AQ = Autism Quotient; EQ = Empathy Quotient, PSS = Point of Subjective Synchrony

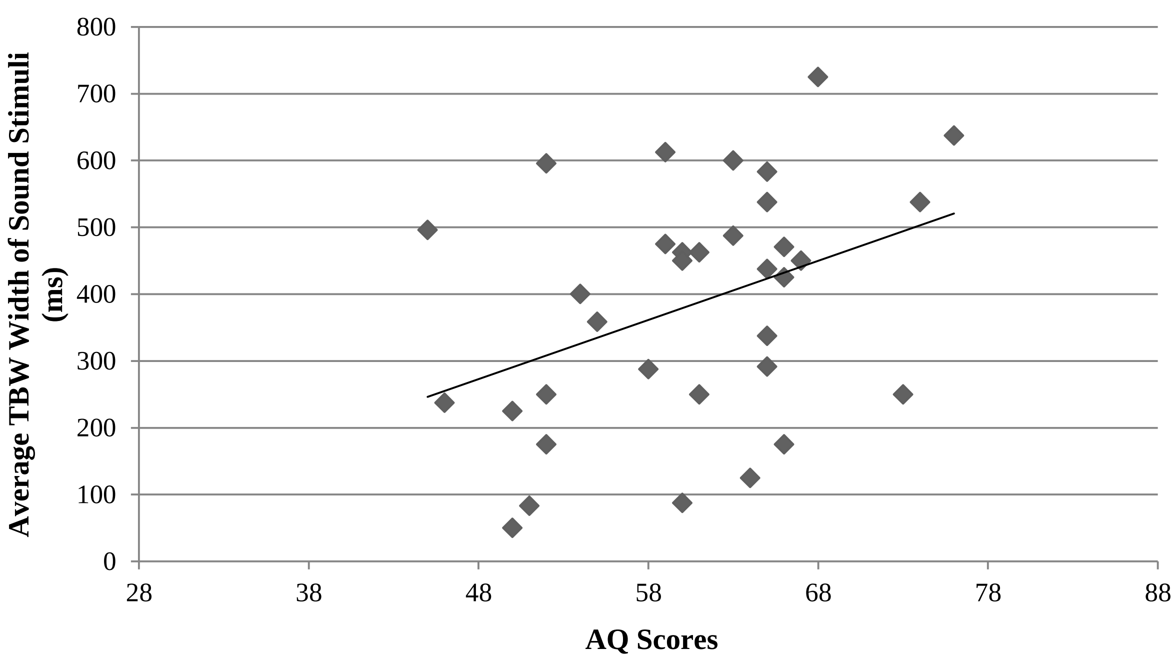

Figure 13. Scatterplot of the positive correlation between AQ scores and the average width of the temporal binding window for the Sounds stimulus in the *Dynamic-50ms toward-synch* task.

Table 8

Correlations of intellectual ability and audiovisual integration skills

		TBW Width	
	Story	Sounds	Piano
VIQ(n)	-.18(36)	-.17(33)	-.13(35)
PIQ(n)	-.23(36)	-.01(33)	-.22(35)
FSIQ(n)	-.20(36)	-.09(33)	-.18(35)
		PSS of Delayed Trials	
	Story	Sounds	Piano
VIQ(n)	-.11(36)	.03(35)	.00(36)
PIQ(n)	-.29(36)	-.07(35)	-.32(36)
FSIQ(n)	-.24(36)	-.02(35)	-.19(36)
		PSS of Advanced Trials	
	Story	Sounds	Piano
VIQ(n)	.23(36)	.25(34)	.25(35)
PIQ(n)	.05(36)	.00(34)	.05(35)
FSIQ(n)	.10(36)	.10(34)	.11(35)

Note. Pearson correlation coefficient are reported for the performance IQ (PIQ) and Spearman's rho for the verbal IQ (VIQ) and full scall IQ (FSIQ)

.

Discussion

The purpose of the current study was to investigate a new methodology for measuring the width of temporal binding windows of audiovisual stimuli in adults. More specifically, the study helped investigate whether using a "dynamic" task will produce more precise and narrower windows in comparison to the commonly used "fixed" task, and if so which variation (the *toward-synch* or the *from-synch*, and the 10ms or the 50ms increments) was the best. Furthermore, the current study aimed to determine what other factors, such as stimulus type and cognitive and behavioural abilities were linked to smaller temporal binding windows. The discussion will be broken up based on the hypotheses, and will be presented in the order the hypotheses were formulated. Then, discussion regarding exploratory analyses will be presented and lastly, implications of the findings, limitations of the current study, and future directions will be discussed.

Experimental Task with the Narrowest TBWs

One of the main goals of the current study was to determine whether giving the participants dynamic control over the timing of the audio information would yield more precise windows than the commonly-used synchrony judgment task that is a method of constant stimuli, where fixed audio advances or delays are presented to participants to judge. The results clearly showed that the *Dynamic-50ms toward-synch* task produced the smallest and most precise temporal binding windows in comparison to the *Fixed* task and all other variation of the "dynamic" tasks (e.g., *Dynamic-10ms toward-synch, Dynamic-10ms from-synch,* and *Dynamic-50ms from-synch*). Specifically, in comparison to all other *toward-synch* tasks, the *Dynamic-50ms toward-synch task* produced temporal binding windows that were almost half the size of the windows from the other two conditions (*Fixed* and *Dynamic-10ms*). Therefore, being able to manually adjust the audio file in relations to a video file with increments that are large enough (50ms but not 10ms) provided a more sensitive measure of audiovisual integration skills by generating narrower temporal binding windows than the commonly used synchrony of judgment task. The *Dynamic-50ms towards-synch* task differed from the others both in the size of adjustment and whether adjustments were moving toward or away from synchrony. Findings from both *toward-synch* and *from-synch* tasks will be discussed separately next.

Differences within the *toward-synch* tasks findings. The *Dynamic-50ms toward-synch* task differs from the *Dynamic-10ms toward-synch* task only in terms of the increments by which participants could move the audio file in relation to the video file. The 10ms increments, first used by Segers (2012), were very small and increased participant exposure to the asynchronicity of the stimuli, which could lead to a shift in temporal binding windows away from the point of true synchrony, resulting in wider windows. Therefore, the purpose of using two types of increments was to reduce the amount of repeated exposure and, in doing so, likely produce smaller windows. Results from the current study support this view as the *Dynamic-50ms toward-synch* task produced smaller temporal binding windows across all stimuli in comparison to the *Dynamic-10ms toward-synch* one. This phenomenon, where repeated exposure to asynchronous audiovisual stimuli can lead to a shift in the temporal binding window in the direction of the asynchrony, has been referred to as "temporal recalibration" (e.g., Fujisake et al., 2004; Vroomen et al., 2004) and has been repeatedly studied (e.g., Di Luca, Machulla, & Ernst, 2009; Keetels & Vroomen, 2008; Navarra et al., 2005; Vatakis, Navarra, Soto-France, & Spence, 2007). Although confounding results have been found in term of the generalization of temporal recalibration across different modalities, where some researchers found that it was generalizable (e.g., Hanson, Heron, & Whitaker, 2008; Odegaard & Shams, 2016) and others did not (e.g., Harrar & Harris, 2005; Harrar & Harris, 2008; Di Luca et al., 2009), all agreed with regards to its effect on audiovisual integration (for review see Vroomen & Keetels, 2010). The results from the current study seem to support this claim as well.

The temporal binding windows produced by the *Fixed* and *Dynamic-10ms toward-synch* tasks did not differ from each other. These results are considerably different from what was hypothesized and to those obtained by Segers (2012), who only compared the *Dynamic-10ms toward-synch* and *Fixed* tasks, as she found that the *Fixed* task produced smaller windows for two of the three stimuli she used (Sound, and Piano, but not Story) compared to the *Dynamic-10ms*. Several possible explanations for this finding exist. First, it is plausible that the order in which the tasks were presented had an effect. In the current study, the *Fixed* task was always performed last and, therefore, when it was time for participants to perform the *Fixed* task, they had been exposed to asynchrony stimuli for a longer period of time. This again could have shifted the participants' temporal binding windows away from the point of true synchrony, making the windows wider. Previous research has looked at the effect of temporal recalibration in both temporal order judgment task (e.g., Vatakis et al., 2007) and synchrony judgment task (e.g., Vatakis et al., 2007; Vroomen et al. 2004). Vatakis and colleagues (2007) demonstrated that temporal recalibration can occur across different domains, where repeated exposure to asynchrony in a speech stream (commonly occurring stimuli) with the audio information delayed by 300ms can cause a shift in PSS on a temporal order judgment task when using arbitrary stimuli such as flash-beep stimuli. Similar results have also been reported when using synchrony judgment tasks. Furthermore, Vroomen and colleagues (2004) found that when comparing the effects of temporal recalibration, the points of subjective synchrony were slightly more shifted in the direction of asynchrony when they were measured using a synchrony judgment (SJ) task in comparison to a temporal order judgment (TOJ) task. These results are important in terms of the current study where the *Fixed* task is a synchrony judgment task and therefore might have been more affected by the exposure to asynchronous audiovisual stimuli.

The current study attempted to reduce exposure to asynchrony by using both *toward-synch* and *from-synch* conditions. However, although, the current study counterbalanced which task participants did before the *Fixed* task, the simple act of being exposed to asynchronicity for an extended period of time within a certain timeframe before the *Fixed* task might be sufficient to shift the TBWs. Furthermore, the current study repeatedly exposed participants to asynchrony in both directions (audio advanced and audio delayed) within the *toward-synch* tasks, which potentially should have cancelled out the effect of the recalibration, or alternatively widen the TBW in both directions. Lastly, it is possible that since the *Fixed* task was at the end of a two hours study, participants' performance and motivation decreased, negatively impacting their ability to integrate audiovisual information which could give the illusion of temporal recalibration, when in fact it is simply a decrease in motivation.

From-synch tasks findings. The *from-synch* tasks were also originally developed to attempt to control for the amount of repeated exposure to asynchrony and to prevent the shift of the temporal binding window toward asynchrony that might have occurred in the Segers (2012) study. However, results demonstrated that the *toward-synch* and *from-synch* tasks did not yield comparable temporal binding windows or points of subjective synchrony. Specifically, for the *Dynamic-50ms* condition, the *from-synch* task yielded temporal binding windows that were more than double the size of the ones obtained in the *toward-synch* task. No specific hypotheses were made concerning how comparable the two conditions were, as it was the first time, to the author's knowledge, that these conditions were systematically compared. It is possible that the instruction of the *from-synch* task were not explicit enough. Although it was verbally stressed to each participants that the current study was interested in obtaining the first point where the participants believed that what they were seeing and hearing no longer matched, how asynchronous these two sources of information needed to be was not specified. It is possible that because the demonstration of the asynchronous stimuli used an offset of +500ms, participants were trying to replicate this wide asynchrony when in fact what the intent was the first noticeable mismatch. Furthermore, the way the *from-synch* task was designed, there was nothing stopping the participants from going beyond the first point of subjective asynchrony, as the video and audio files just became more asynchronous the more they pressed. However, in contrast, if participants went too far beyond the point of subjective synchrony in the *toward-synch* tasks, the audio and visual information would eventually become out of synch again, signalling to the participants to go back. No such consequence occurred in the *from-synch* tasks making it difficult to prevent participants from going well beyond their point of subjective asynchrony, which may explain the wider TBWs that were noticed.

Dixon and Spitz (1980), who were amongst the first to study the temporal binding window in audiovisual stimuli, used a somewhat similar task to the *from-synch Dynamic-50ms* tasks presented in this study where participants operated a switch control that could advance or delay the auditory information at a constant rate of 51 milliseconds per seconds. Participants were informed to hold down a key until they noticed that the visual and auditory information were out of synchrony. In comparison to the new tasks presented in this study, Dixon and Spitz (1980) only measured the points of subjective synchrony in relation to going toward asynchrony and not the opposite. They found that their task produced quite narrow temporal binding windows, ranging from about -130ms to +250ms. They were the only ones who ever used this alternate version of the "dynamic" task and therefore no replications of their findings using a "dynamic" task have been found. Unfortunately, the current study did not replicate their findings with the *from-synch Dynamic-50ms* task, but did find comparable and even narrower windows for the *toward-synch* task. A possible explanation for this lies in the differences in the designs across their task and the *from-synch Dynamic-50ms* task. First, Dixon and Spitz used a more continuous, instead of discrete, adjustment, where participants only had to let go of the control switch when they believed the audio and visual information were out of synch. This could have reduced the amount of perceived control the participants had in comparison to having to physically press a key every time they wanted to shift the audio file, which may have encouraged them to release the switch faster. Furthermore, Dixon and Spitz (1980) limited how far out of synch the participants could go (500ms). As mentioned previously, participants in the current study could go as far out of synch as they wanted and had no physical limits or cues to stop them which could have contributed to the much wider windows reported on these tasks compared to the *toward-synch* ones and to Dixon and Spitz's task.

Characteristics of the *Dynamic-50ms Toward-synch* Task

Since the smallest and most precise temporal binding windows were produced by the *toward-synch Dynamic-50ms* task, the majority of the results focused solely on describing various aspects of those windows. First, it was observed that the temporal binding windows generated by this task had a positively skewed asymmetry, where participants were better at detecting asynchrony when auditory information was presented before visual information. Similar results have been seen in many other studies (e.g., Dixon & Spitz, 1980; Hillock et al., 2011; Stevenson & Wallace, 2013; Zmigrod & Zmigrod, 2016), which goes to support that the current new task is sensitive in measuring and detecting this commonly occurring positively skewed asymmetry in perceptual integration. This ability to better detect asynchrony when auditory precedes visual information has been linked to the fact that these events rarely occur within the natural environment and, therefore, less tolerance to this type of asynchrony has been developed (Dixon & Spitz, 1980; Hillock et al., 2011).

Effect of stimulus type. The temporal binding windows obtained from the *Dynamic-50ms toward-synch* task also showed sensitivity to the different type of stimuli, where the more ecologically valid stimulus type (Story) produced smaller windows, and largest degree of asymmetry compare to the less commonly occurring stimuli (Sounds and Piano). These results are in line with the proposed hypothesis and replicate previous findings (e.g., van Eijk et al., 2008; Van Wassenhove et al., 2007; Vatakis & Spence, 2006). In accordance with the current study, Vatakis & Spence (2006) found that individuals are more sensitive to asynchrony in speech than they are in both a guitar and piano music stimuli. Furthermore, the average point of subjective synchrony for advanced trials of the Story stimulus (-65ms) was similar and, in some cases, closer to true synchrony in comparison to data reported in other studies, which ranged from -66ms (Vatakis & Spence, 2006) to -130ms (Dixon & Spence, 1980). Although, no differences between the points of subjective synchrony across the three stimulus types were observed on auditory delayed trials, the numbers obtained were still comparable to what other studies found (e.g., Conrey & Pisoni, 2006; Dixon & Spitz, 1980; Massaro & Cohen, 1993). Conrey and Pisoni (2006), who also looked at the differences in temporal binding window size across audiovisual speech and non-speech stimuli, found that the point of subjective synchrony for their auditory delayed trials did not differ across stimuli type. These comparable results in delayed and advanced trials support the claim that the *Dynamic-50ms toward-synch* task, a new method developed in the current study, is as sensitive if not better at measuring the audiovisual integration skills than other previously used methods.

It is important to note that across all three *toward-sync* conditions (*Fixed*, *Dynamic-10ms*, and *Dynamic-50ms*), the Story stimulus, where a woman is narrating a children's story, yielded the smallest temporal binding window. These results show the robustness of the human brain to detect asynchrony in social linguistic stimuli. More specifically the linguistic aspect of this type of stimulus seems to be the contributing factors to these findings as this was the only component that differentiates the Sound stimulus from the Story one. Previous research has explained a greater sensitivity to speech stimuli in reference to the ecological validity of the stimuli (Aschersleben, 1999; Vatakis & Spence, 2006; van Eijk et al., 2008). By adulthood, individuals on average have been exposed to a fairly large amount of social situations and speech, and therefore are able to detect more subtle asynchrony, especially when audio information precedes visual information. Meanwhile, although the Sound and Piano stimuli used in the current study still depicted somewhat naturally occurring life events, the likelihoods that the participants have been exposed to these events are not as prominent as for speech and so may not be as perceptually salient as speech. Therefore, it seems that greater sensitivity to audiovisual stimuli is strongly linked to events occurring in nature and the frequency of exposure to those events.

Autism-Trait and Intellectual Abilities Findings

Two measures of autism-traits were included in the current study to observe whether a link between audiovisual integration skills and autism-like traits existed and whether the new methodology could pick up on this link. Furthermore, many studies have demonstrated that individuals with autism spectrum disorders tend to have wider temporal binding windows in comparison to typically developing peers (e.g., Bebko et al., 2006; Stevenson et al., 2014; Taylor et al., 2010). However, very few significant correlations were found between the audiovisual integration skills and measures of autism-like traits. Only the Sound stimulus yielded significant results, showing that higher levels of autism-like traits, as indicated by higher scores on the Autism Quotient questionnaire, were linked to wider temporal binding window and points of subjective synchrony further from absolute synchrony. No other stimuli produced significant results and nothing was correlated with the Empathy Quotient questionnaire. These results are not surprising given that all the participants were typically developing university student and all scored quite low on autism-trait (mean of 60.8 on a possible score of 112, which is below the clinical cut-off of 70). Furthermore, there was very little variance in scores, making it very difficult to see significant correlations. However, the two significant correlations found do follow the predicted hypothesis and are in line with previous findings within the ASD population. It is possible that significant results were only seen for the Sound stimulus due to the level of difficulty associated with this stimulus. Anecdotally, the Sound stimulus was reported as being extremely difficult by participants and it yielded the least sensitive measures of audiovisual integration skills across most measures and fairly high variability. Therefore, maybe the greater level of variability in scores within the Sound stimulus was what made it possible to observe a significant correlation with not very well distributed autism-trait scores.

The link between audiovisual integration skills and intellectual ability was also analyzed. No significant correlations were found between any of the audiovisual integration measures and all three forms of intellectual abilities recorded. The absence of significant findings may be attributed to the small variance in intellectual abilities scores given that the current sample used only university students. However, it is important to note that the current sample had fairly large amount of English second language speakers who scored lower on verbal IQ, but yet performed similarly to native speakers on the tasks. Future research should attempt to obtain a wider variety of intellectual abilities to truly observe whether verbal abilities are linked to audiovisual integration skills within speech stimuli.

Response Time

Analyses of participants' response time were not originally planned and hypothesized for, however some of these preliminary findings are worth mentioning. First, it was observed that participants were faster at finding the points of subjective synchrony for advanced trials of the Story stimulus compared to the Piano and Sound stimuli, and that it took them the longest for the Sound stimulus. Response time could be interpreted as measure of difficulty, demonstrating that participants had a harder time finding the point of subjective synchrony for the Sound stimulus.

Furthermore, correlation between response time and point of subjective synchrony for the Sound stimulus yielded opposite results for the advanced trials and the delayed trials. Within Sound stimuli of the delayed trials, it was found that the closer participants got to true point of synchrony, the longer they took. This makes sense as in order to move closer to true synchrony participants needed to press the key more, which by default can result in more time. However, for advanced trial, it was found that the closer the participants got to the true point of synchrony, the faster they were. It is possible that the fact that individuals are better at detecting asynchrony when audio information is presented before visual information, triggered participants to press more quickly at the beginning of the advanced trials, especially for the Sound stimulus, getting them closer to true synchrony. However, it is important to note that these correlations were only found for one of the three types of stimulus and as such are not worth over analyzing. Lastly, it was found that response times across all stimulus type within delayed and advanced trials were highly correlated with each other. These results seem to indicate that the speed at which a participant performed the tasks was fairly consistent across tasks and stimuli types.

Limitations

Task order effect. As noted in the results section, a task order effect was noticed, where the first set of experimental tasks influenced how the participants behaved on the second set of tasks. However, due to the nature of the analyses performed (purely within-subject analyses) and the random assignment of participants to different task order groups, although not fully balanced, the impact of task order was reduced. Furthermore, since the preliminary analyses of individual task order groups yielded the same general finding, that the *Dynamic-50ms toward-synch* task produced the smallest windows; task order was no longer included in the results. However, the task order effect still has implication at the level of future design of the studies. The fact that the current study had an unequal number of participants who started with each task is a major limitation. Therefore, the next step for the current research would be to fully counterbalance the task order groups and to compare them to determine if the same results would be found. Another way to deal with this task order effect would be to only look at the first tasks that participants did. However, due to high variability across participants, it is not optimal to compare the size of temporal binding windows across participants and it is better to have a within subject design. A final way to deal with this effect would be to remove one of the *Dynamic* tasks; however this would make it hard to compare the two *Dynamic* tasks against each other.

The task order effect seen in the current study could be a result of the similarity between the two sets of tasks. The set of tasks were practically identical as the only difference lay in the increment each key press produced, either 10ms or 50ms. Although the participants were warned at the beginning of the second set of the *Dynamic* tasks to pay attention to the task and that they might have to press more or less on the arrow keys, it is possible that the instructions were not explicit enough or that participants simply did not pay close enough attention. Furthermore, although there was on average a forty-five minute break between the first and second set of *Dynamic* tasks, it might not have been long enough. These are all possible considerations to have for future design of the study. However, the main point remains that presenting the *Dynamic-10ms* tasks first influences the number of key presses for the *Dynamic-50ms* tasks.

Motivation and fatigue. Another major limitation of the current study originates from the length of the study. Most participants took up the entire two hours testing session to complete all the tasks and therefore it would not be surprising if their motivation and engagement in the task decreased over time. No measure of motivation or attention was included in the battery of tests performed. Future studies should take this into consideration and should add some form of attention checks or an engagement in task check. Previous studies have asked participants to keep track of a completely irrelevant occurrence of an outside stimulus to ensure that the participants were focused on the audiovisual integration tasks (Vroomen et al., 2004). However, dividing the participants' attention may also influence the size of the temporal binding windows and reduce the attention they put on the actual task of interest. Including more breaks or breaking up the study over two testing sessions are also possible ways to deal with the potential decrease in motivation and fatigue effect.

Furthermore, it is possible that the task order effect is also a result of a fatigue effect, where participants paid less attention during the second set of *Dynamic* tasks. If this is the case, it is possible that the larger windows observed in the *fixed* task have been influenced by participants attention and tiring. It was noticed that the widths of the temporal binding windows of the *fixed* task from the current study were slightly larger than the ones observed by Segers (2012). However, the widths for the fixed task that Segers (2012) reported were still much larger than the obtained widths for the *Dynamic-50ms toward-synch* tasks, demonstrating that the main results of the current study still hold. In the end, when looking at audiovisual integration skills for an extended period of time, it is important to take into consideration attention, motivation and the possibility of a fatigue effect.

Dynamic task. The *Dynamic-50ms toward-synch* task is not without limitations. First the task is subject to individual biases and judgements, as one participant's definition of synchrony may vary in terms of the degree of synchrony one is looking for. For example, one participant could interpret the *Dynamic* task as finding first point where the asynchrony is tolerable where another participant might be looking for perfect synchrony between audio and visual information. Although participants were all instructed to find that first point where they noticed that what they heard and what they saw was synchronous, how they defined synchrony was very personal and therefore there was no way to ensure that participants all had the same definition for it. In this regard, the current task is limited by this notion of response criterion that might have differed from one participant to the next. Methods within the field of psychophysics, such as the method of constant stimuli, forces participant to make a decision in order to minimize the impact of these varying response criteria between individuals. However methods that minimizes or eliminates subjective perceptual differences of what synchrony means is not without flaws either. Methods such as the synchrony judgment task which is fundamentally a method of constant stimuli, determines temporal binding windows based on a staircase approach where really it is the researcher who makes a judgement with regards to the rules he or she wants to use as a cut off or threshold. For example, is it when participants are answering 50 percent of the time that the information is synchronous and 50 percent that it is not synchronous? But a rationale could also be made that the criterion should be 40 or 60 or another percent. Furthermore, the fact that the current task embraces the notion of the differences in response criterion across participants could be seen as a strength of the current method. This is especially the case when thinking about using this alternative methodology with clinical populations where their definition of synchrony might be different from a typically developing population, such as is certain perceptual systems might work differently. Therefore, imposing a fixed definition of what a point of subjective synchrony is might limit the insight one could get into these individuals' worlds.

A second limitation of the *Dynamic-50ms toward-synch* task comes from another concept within the psychophysics field, and that is the presence of hysteresis. Hysteresis refers to the phenomenon where perception on a single trial within a task is influenced by performance on the preceding trial (e.g., Hock & Schoner, 2010; Martin, Kõsem, & van Wassenhove, 2015; Woodwroth & Schlosberg, 1972). In other words, the current perception of an event is depend partially on the current event but also partially on the residual effects of the previous event, specifically when these two events are close in time with each other (Odic, Hock, & Halberda, 2014). Although hysteresis was present in both the *Dynamic* and *Fixed* tasks, it would have been present to a higher degree within the *Dynamic* task as each key press within a single trial would be subject to hysteresis. What this means is that every time participants pressed the arrow key to move the audio file more in synchrony with the visual file, participants were making a judgement of whether or not they believed what they heard and what they saw was synchronous, and that judgment might have been influenced by what they saw before the current key press. Although the impact of hysteresis on performance within the *Dynamic-50ms toward-synch* task is important to acknowledge, it is important to note that the task itself was much more fluid than this idea that at every key press the participants stopped and took the time to decide how the previous stimulus before the key press compares to the current one. Furthermore, there is a movement within psychophysics that argues that the presence of hysteresis is not merely a reflection of a side effect of testing methods or participant bias but actual holds meaningful information (e.g., Odic et al., 2014). In this view, perceptual hysteresis is actually part of the sensory system and influences how one perceives the world in important ways. Therefore attempts at controlling such phenomena may alter what one is perceiving; therefore, it should be acknowledge as part of the perceptual system instead of trying to limit or eliminate it (Jones, Mozer, & Konoshita, 2009, Taylor & Lupker, 2001; Odic et al., 2014).

Future Directions

Future research could use this methodology with clinical populations, such as individuals with autism spectrum disorder. Data on autism-traits and intellectual abilities were collected with the goal of conducting the same study with individuals with autism spectrum disorder if conclusive results were obtained regarding the new methodologies. Previous research have demonstrated that individuals on the autism spectrum tend to have wider temporal binding windows than the typically developing population and these have been linked to their difficulties in social abilities (e.g., Bahrick, 2010; Bebko et al., 2006; Foss-Feig et al. ,2010; Wallace & Stevenson, 2014). However, these differences are not consistent across all types of stimuli. Specifically, Bebko and colleagues (2006) found that children with an autism spectrum disorder had difficulties with audiovisual integration for linguistic stimuli but not with non-linguistic stimuli. However, no study, to the author’s knowledge, has used a “dynamic” task to look at the differences between the two samples. Therefore, including a clinical population group would help to determine whether the *Dynamic-50ms* task is first sensitive enough to pick up the differences between groups. Secondly, since the *Dynamic-50ms toward-synch* task has been shown here to be a more precise measure of audiovisual integration skills in comparison to the classically use synchrony judgment task, it may reveal more in-depth information with regards to audiovisual integration skills for individuals with autism spectrum disorder. Therefore, future research including this clinical population could shed light on some of the social ability difficulties within this population and could potentially help inform future interventions in that domain.

Future research should make several modifications to the current design of the study. First, in order to shorten the duration of the study and to eliminate the task order effect, both of the *Dynamic-10ms* tasks should be removed. Both Segers (2012) and the current study found no supporting proof that the *Dynamic-10ms* yielded narrower and more precise windows than the classic synchrony-judgment task and therefore there is no need to keep using it. This modification could potentially take care of both the task order effect and the motivation issue discussed previously. Secondly, if the *from-synch* task of the *Dynamic-50ms* is kept, changes to the instructions must be made in order to discourage participants from over pressing and to encourage them to stop at their true point of perceived asynchrony.

Lastly, future directions should include collecting and analyzing eye tracking data while performing audiovisual integration tasks to observe where participants are looking while performing these tasks and to determine whether individuals who have smaller temporal binding windows differ in their looking patterns than those with larger windows. Secondly, the eye tracking data could be used as a measure of attention which could help inform the researcher to a certain degree about the motivation of the participants as time went by. This would particularly be useful in terms of determining whether having the *fixed* task at the end of the study impacted the results found.

Conclusions and Implications

In conclusion, the current study was able to demonstrate that the *Dynamic-50ms toward-synch* task is a valid methodology for measuring the width of temporal binding windows of audiovisual stimuli. This new task is sensitive to both the positively skewed asymmetry of the temporal binding windows and of the effect of different stimuli. Furthermore, the current study demonstrated that by having participants be more involved in determining the width of their temporal binding window, narrower windows were produced in comparison to using a more passive measure. These findings hold implications at the level of measuring audiovisual integration skills and in terms of more deeply understanding the integration of such multimodal sensory systems. First, having more precise measurements of the temporal binding window could greatly benefit fields such as the designing of hearing aids, broadcasting and video conferencing technology. Furthermore, it could help better understand and better inform interventions for individuals with multisensory integration difficulties such as autism. Lastly, having more precise measures for calculating points of subjective synchrony could eventually help better understand the mechanism behind audiovisual integration.

References

Allan, L. G. (1975). The relationship between judgments of successiveness. *Perception & psychophysics*, *18*(1), 29-36.

Aschersleben, G. (1999). On the asymmetry of the temporal contiguity window: Commentary on Lewkowicz. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 421-424). Amsterdam: Elsevier.

Axelrod, B. N. (2002) Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms of estimating intellectual functioning. *Assessment, 9*(1), 17-23. DOI: 10.1177/1073191102009001003

Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. In G. Bremner & T. D. Wachs (Eds.). *Blackwell handbook of infant development: 2nd ed.* (pp. 120-166), Oxford, England: Blackwell Publishing.

Bald, L., Berrien, F. K., Price, J. B., & Sprague, R. O. (1942). Errors in perceiving the temporal order of auditory and visual stimuli. *Journal of Applied Psychology*, *26*(3), 382-388. DOI: http://dx.doi.org/10.1037/h0059216

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, *34*(2), 163-175. DOI:10.1023/B:JADD.00000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry*, *42*(2), 241-251. DOI: 10.1111/1469-7610.00715.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17. DOI: 10.1023/A:1005653411471

Bates, S. L., & Wolbers, T. (2014). How cognitive aging affects multisensory integration of navigational cues. *Neurobiology of aging, 35*(12), 2761-2769. DOI:10.1016/j.neurobiolaging.2014.04.003

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

Bebko, J. M., Schroeder, J. H., & Weiss, J. A.(2014). The McGurk effect in children with autism and Asperger syndrome. *Autism Research*, *7*(1), 50-59.

Bedford, F. L. (1989). Constraints on learning new mappings between perceptual dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 232-248.

Bishop, C. W. & Miller, L. M. (2009). A multisensory cortical network for understanding speech in noise. *Journal of Cognitive Neuroscience, 21*(9), 1790-1804. DOI: 10.1162/jocn2009.21118.

Boenke, L. T., Deliano, M., & Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental brain research, 198*(2-3), 233-244. DOI: 10.1007/s00221-009-1917-z.

Boucher, J., Lewis, V., & Collis, G. (1998). Familiar face and voice matching and recognition in children with autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(2), 171-181.

Brown, C., & Dunn, W. (2002) *Adolescent/Adult Sensory Profile*. San Antonio, TX: Psychological Corporation.

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

Calvert, G. A., Brammer, M. J., & Iversen, S. D. (1998). Crossmodal identification. *Trends in cognitive sciences*, *2*(7), 247-253.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-grouping model. *Journal of Cognitive Neuroscience, 16,* 1000-1009.

Conrey, B., & Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *The Journal of the Acoustical Society of America*, *119*(6), 4065-4073.

D'Agostino, R. B., Belanger, A., & D'Agostino, R. B., Jr. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician*, *44*(4), 316-321.

de Boer-Schellekens, L., Eussen, M., & Vroomen, J. (2013). Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Frontiers in integrative neuroscience*, *7*(8),1-8. DOI: 10.3389/fnint.2013.00008.

de Boer-Schellekens, L., & Vroomen, J. (2014). Multisensory integration compensates loss of sensitivity of visual temporal order in the elderly. *Experimental brain research*, *232*(1), 253-262. DOI: 10.1007/s00221-013-3736-5.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in cognitive sciences*, *7*(10), 460-467. DOI: 10.1016/j.tics.2003.08.014

de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., & Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia research*, *72*(2), 195-203. DOI: 10.1016/j.schres.2004.02.013

de Jong, J. J., Hodiamont, P. P. G., Van den Stock, J., & De Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophrenia research*, *107*(2), 286-293. DOI: 10.1016/j.schres.2008.10.001.

Diederich, A., & Colonius, H. (2009). Crossmodal interaction in speeded responses: time window of integration model. *Progress in brain research, 174*, 119-135. DOI: 10.1016/S0079-6123(09)01311.9

Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: relationg reaction times and judgments of temporal order. *Psychological Review*, *122*(2), 232-241. DOI: 10.1037/a0038696

Diederich, A., Colonius, H., & Schomburg, A. (2008) Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia, 46*(10), 2556-2562. DOI: 10.1016/j.neuropsychologia.2008.03.026

Di Luca, M., Machulla, T. K., & Ernst, M. O. (2009). Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *Journal of vision*, *9*(12), 1-16. DOI: 10.1167/9.12.7.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception, 9*(6), 719-721. DOI: 10.1068/p090719

Ehrenstein, W.H. and Ehrenstein, A. (1999) *Psychophysical methods*. In Windhorst, U. and Johansson, H. (eds), *Modern Techniques in Neuroscience Research*.pp.1211 -1241. Springer, Berlin.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, *11*(1), 581-602. DOI : 10.1007/BF01659319.

Fechner, G. T. (1860/1966). *Elements of psychophysics.* H. E. Adler, D. H., Howes, E. G. Boring (Trans.). New York, NY: Holt, Rinehard and Winston.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*(2), 381-389. DOI: 10.1007/s00221-010-224-4

Froyen, D., Willems, G., & Blomert, L. (2011). Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter–speech sound processing. *Developmental science*, *14*(4), 635-648.

Fujisaki, W., & Nishida, S. Y. (2009). Audio–tactile superiority over visuo–tactile and audio visual combinations in the temporal resolution of synchrony perception. *Experimental brain research*, *198*(2-3), 245-259. DOI: 10.1007/s00221-009-1870-x

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. Y. (2004). Recalibration of audiovisual simultaneity. *Nature neuroscience*, *7*(7), 773-778. DOI: 10.1038/nn1268

Girin, L., Schwartz, J., & Feng, G. (2001). Audio-visual enhancement of speech in noise. *Journal of Acoustical Society of America, 109*(6), 3007-3020. DOI: 10.1121/1.1258887.

Gondan, M., Niederhaus, B., Rösler, F., & Röder, B. (2005). Multisensory processing in the redundant-target effect: a behavioural and event-related study. *Perceptual Psychophysics, 67*(4), 713-726. DOI: 10.3758/BF03193527.

Grant, K. W., & Walden, B. E. (1996). Evaluating the articulation index for auditory-visual consonant recognition. *Journal of Acoustical Society of America, 100*, 2415-2424.

Hairston, W. D., Burdette, J. H., Flowers, D. L., Wood, F. B., & Wallace, M. T. (2005). Altered temporal profile of visual–auditory multisensory interactions in dyslexia. *Experimental Brain Research*, *166*(3-4), 474-480. DOI: 10.1007/s00221-005-2387-6.

Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., & Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Experimental brain research*, *152*(3), 404-408.

Hamlin, A. J. (1895). On the least observable interval between stimuli addressed to disparate senses and to different organs of the same sense. *The American Journal of Psychology*, *6*(4), 564-575. DOI: 10.2307/1411193

Hancock, L. N. (2009). *The influence of background noise on the intermodal perception of speech in children with autism spectrum disorders.* Unpublished master's thesis, York University, Toronto, Ontario, Canada.

Hanson, J. V., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research*, *185*(2), 347-352. DOI: 10.1007/s00221-008-1282-3.

Harrar, V., & Harris, L. R. (2005). Simultaneity constancy: detecting events with touch and vision. *Experimental Brain Research*, *166*(3-4), 465-473. DOI: 10.1007/s00221-005-2386-7.

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental brain research*, *186*(4), 517-524. DOI: 10.1007/s00221-007-1253-0.

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

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, *63*(3), 289-293. DOI: 10.1037/h0039516.

Hillock, A. R., Powers, A. R., & Wallace, M. T. (2011). Binding of sights and sounds: age related changes in multisensory temporal processing. *Neuropsychologia*, *49*(3), 461-467. DOI: 10.1016/j.neuropsychologia.2010.11.041

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science, 15*(5), 688-696. DOI: 10.1111/j.1467-7687.2012.01171.x.

Hirsh, I. J., & Fraisse, P. (1964). Simultanéité et succession de stimuli hétérogènes. *L'Année psychologique*, *64*(1), 1-19.

Hirsh, I. J., & Sherrick Jr, C. E. (1961). Perceived order in different sense modalities. *Journal of experimental psychology*, *62*(5), 423-432. http://dx.doi.org/10.1037/h0045283

Hock, H. S., & Schöner, G. (2011). Measuring perceptual hysteresis with the modified method of limits: Dynamics at the threshold. In *Fechner's Legacy in Psychology* (pp. 63-86). Brill. DOI: 10.1163/ej.9789004192201.i-214.27

Hoekstra, R. A., Vinkhuyzen, A. A., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., . . . van der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of autism and developmental disorders, 41*(5), 589-596. DOI: 10.1007/s10803-010-1073-0

Jones, M., Kinoshita, S., & Mozer, M. C. (2009). Optimal response initiation: Why recent experience matters. In D. Koller, D. Schuurmans, Y. Bengion, and L. Bottou (eds.), *Advances in neural information processing systems* (pp. 785-792), Keane, B. P., Rosenthal, O., Chun, N. H., & Shams, L. (2010). Audiovisual integration in high functioning adults with autism. *Research in Autism Spectrum Disorders, 4*(2), 276-289. DOI: 10.1016/j.rasd2009.09.015

Keetels, M., & Vroomen, J. (2007). No effect of auditory–visual spatial disparity on temporal recalibration. *Experimental Brain Research, 182*(4), 559-565. DOI: 10.1007/s00221-007 1012-2.

Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics*, *38*(1), 52-54. https://doi.org/10.5395/rde.2013.38.1.52

Krueger Fister, J. K., Stevenson, R. A., Nidiffer, A. R., Barnett, Z. P., & Wallace, M. T. (2016). Stimulus intensity modulates multisensory temporal processing. *Neuropsychologia*, *88*, 92-100. DOI: 10.1016/j.neuropsychologia.2016.02.016

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychological medicine*, *34*(5), 911-920. DOI: 10.1017/S0033291703001624

Lewald, J., Ehrenstein, W. H., & Guski, R. (2001). Spatio-temporal constraints for auditory visual integration. *Behavioural brain research*, *121*(1), 69-79.

Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child development*, *85*(2), 685-694. DOI: 10.1111/cdev.12142.

Lovelace, C. T., Stein, B. E., & Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive brain research*, *17*(2), 447-453. DOI: 10.1016/S0926-6410(03)00160-5

Martin, J. R., Kösem, A., & Van Wassenhove, V. (2015). Hysteresis in audiovisual synchrony perception. *PloS one*, *10*(3),1-13. DOI: 10.137/journal.phone.0119365.

Massaro, D. W., & Cohen, M. M. (1993). Perceiving asynchronous bimodal speech in consonant-vowel and vowel syllables. *Speech Communication*, *13*(1-2), 127-134. https://doi.org/10.1016/0167-6393(93)90064-R

McGrath, M., & Summerfield, Q. (1985). Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *The Journal of the Acoustical Society of America*, *77*(2), 678-685. doi: http://dx.doi.org/10.1121/1.392336

Mégevand, P., Molholm, S., Nayak, A., & Foxe, J. J. (2013). Recalibration of the multisensory temporal window of integration results from changing task demands. *PloS one, 8*(8), e71608. DOI: 10.1371/journal.pone.0071608

Meredith, M. A., & Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain research*, *365*(2), 350-354.

Mongillo, E. A., Irwin, J. R., Whalen, D. H., Klaiman, C., Carter, A. S., & Schultz, R. T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *Journal of autism and developmental disorders*, *38*(7), 1349-1358. DOI: 10.1007/s10803-007-0521-y

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research*, *25*(2), 499-507. http://dx.xoi.org/10.1016/j.cogbrainres.2005.07.009

Oakland, T., Black, J. L., Stanford, G., Nussbaum, N. L., & Balise, R. R. (1998). An Evaluation of the Dyslexia Training Program A Multisensory Method for Promoting Reading in Students with Reading Disabilities. *Journal of learning disabilities*, *31*(2), 140-147.

Odegaard, B., & Shams, L. (2016). The brain's tendency to bind audiovisual signals is stable but not general. *Psychological science*, *27*(4), 583-591. DOI: 10.1177/0956797616628860

Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General*, *143*(1), 255-265. DOI: 10.1037/a0030825.

Pearl, D., Yodashkin-Porat, D., Katz, N., Valevski, A., Aizenberg, D., Sigler, M., . . . Kikinzon, L. (2009). Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Comprehensive psychiatry*, *50*(2), 186-192. DOI: https://doi.org/10.1016/j.comppsych.2008.06.004

Pearson, E. S., D'Agostino, R. B., & Bowman, K. O. (1977). Tests for departure from normality: Comparison of powers. *Biometrika*, *64*(2), 231-246. DOI: https://doi.org/10.1093/biomet/64.2.231

Power, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of Neuroscience*, *29*(39), 12265-12274. DOI: 10.1523/JNEUROSCI.3501-09.2009

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

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, *2*(1), 21-33.

Ross, L. A., Molholm, S., Blanco, D., Gomez-Ramirez, M., Saint-Amour, D., & Foxe, J. J. (2011). The development of multisensory speech perception continues into the late childhood years. *European Journal of Neuroscience, 33*(12), 2329-2337. DOI: 10.1111//j.1460-9568.2011.07685.x

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., & Foxe, J. J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia research*, *97*(1), 173-183. DOI: 10.1016/jschres.2007.08.008

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular autism*, *6*(1), 2-12. DOI: 10.1186/2040-2392-6-2.

Segers, M. (2012). A dynamic approach to temporal binding in audiovisual integration. Unpublished master's thesis, York University, Toronto, Ontaio.

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical science and technology*, *28*(2), 61-70. DOI: 10.1250/ast.28.61

Spence, C., & Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Current Biology*, *13*(13), R519-R521. DOI: 10.1016/S0960-9822(03)00445-7

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., ... & Wallace, M. T. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain topography*, *27*(6), 707-730. DOI: 10.1007/s10548-014-0365-7

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental brain research*, *227*(2), 249-261. DOI: 10.1007/s00221-013 3507-3

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S., & Wallace, M. T. (2014). Multisensory Temporal Integration in Autism Spectrum Disorders. *The Journal of Neuroscience*, *34*(3), 691-697. DOI: 10.1523/JNEUROSCI.3615—13.2014

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1517-1529. DOI: 10.1037/a0027339.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, *421*(6926), 911-911. DOI: 10.1038/421911a

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

Summerfield, A. Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp.3-51), London: Erlbaum.

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

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. http://dx.doi.org/10.1007/s10803-010-1000-4

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology-Learning Memory and Cognition*, *27*(1), 117-137. DOI: 10.1037/0278-7393.27.1.117

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of cognitive neuroscience*, *17*(9), 1396-1409. DOI: 10.1162/0898929054985383

van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments : Effects of experimental method and stimulus type. *Perception & Psychophysics*, *70*(6), 955-968. DOI: 10.3758/PP.70.6.955

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598-607. DOI: 10.1016/j.neuropsychologia.2006.01.001

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Temporal recalibration during asynchronous audiovisual speech perception. *Experimental brain research*, *181*(1), 173-181. DOI: 10.1007/s00221-007-0918-z

Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain research*, *1111*(1), 134-142. DOI: 10.1016/j.brainres.2006.05.078

Vatakis, A., & Spence, C. (2010). Audiovisual temporal integration for complex speech, object action, animal call, and musical stimuli. In *Multisensory object perception in the primate brain* (pp. 95-121). Springer New York.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871-884. DOI: 10.3758/APP.72.4.871

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive brain research*, *22*(1), 32-35. http://dx.doi.org/10.1016/j.cogbrainres.2004.07.003

Wallace, M. T. (2004). The Development of Multisensory Integration. In G. A. Calvert, C. Spence, & B. E. Stein (Eds), *The handbook of multisensory processes* (pp. 625-642). Cambridge, MA: MIT Press.

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105-123. DOI: 10.1016/j.neuropsychologia.2014.08.005

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale- Fourth Edition.* Bloomington, MN: Pearson.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence – Second Edition Manual.* Bloomington, MN: Pearson.

Welch, R. B. (1999). Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. *Advances in psychology*, *129,* 371-387.

Williams, L. E., Light, G. A., Braff, D. L., & Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia*, *48*(10), 3128-3136. DOI: 10.1016/j.neuropsychologia.2010.06.028.

Woodworth, R. S., & Schlosberg, H. (1962). *Experimental psychology*. Oxford, England: Holt, Rinehart and Winston.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research*, *152*(2), 198-210. doi: 10.1007/s00221-003-1536-z

Zmigrod, S., & Hommel, B. (2011). The relationship between feature binding and consciousness: Evidence from asynchronous multi-modal stimuli. *Consciousness and cognition*, *20*(3), 586-593. DOI: 10.1016/j.concog.2011.01.011

Zmigrod, L., & Zmigrod, S. (2016). On the Temporal Precision of Thought: Individual Differences in the Multisensory Temporal Binding Window Predict Performance on Verbal and Nonverbal Problem Solving Tasks. *Multisensory Research*, *29*(8), 679-701. DOI: 10.1163/22134808-00002532.

List of Appendices

York University

CONSENT TO PARTICIPATE IN RESEARCH

Sights and Sounds: How we integrate what we see and what we hear

You are asked to participate in a research study conducted by Melissa Ferland and Dr. James Bebko**,** from the **Department of Psychology at York University**. Melissa Ferland is a Master's student in Psychology at York University. The current research project will be conducted under the supervision of Dr. James Bebko. Results will be contributed towards Melissa's Master's thesis. This research has been reviewed and approved by the Human Participants Review Sub-Committee, York University's Ethics Review Board and conforms to the standards of the Canadian Tri-Council Research Ethics guidelines.

If you have any questions or concerns about the research, please feel free to contact:

Melissa Ferland OR Dr. James Bebko

PURPOSE OF THE STUDY

The current study will examine the audiovisual integration abilities of adults, in efforts to understand how people understand the information that they see and hear simultaneously.

PROCEDURES

If you volunteer to participate in this study, we would ask you to do the following things:

Computer Tasks: As part of this study you will be asked to do two different tasks in which you will view videos of someone talking, or someone playing the piano. In one task you will adjust the sound so that the sound and the video are playing at the same time. In the other task, you will say whether you think that the video is synced (the sound and picture are playing at the same time) or not. During the computer task your eye movements will be recording using an eye-tracking system that will be positioned on the desk in front of you. There is no harm associated with having your eye-movements recorded.

Measures & Questionnaires**:** As part of this study you will be administered two verbal problem-solving skills tests (vocabulary and how things are similar) and two non-verbal skill tests (problem solving). You will also complete a self-administered questionnaire about your own thoughts, feelings and behaviours.

POTENTIAL RISKS AND DISCOMFORTS

There are no foreseeable risks to participating in this study.

POTENTIAL BENEFITS TO PARTICIPANTS AND/OR TO SOCIETY
There are no direct benefits to participating in this study.
The proposed study has significant implications for understanding the processes by which audiovisual information is integrated and comprehended. The results of the current study will make significant contributions to the field's scientific understanding of audiovisual integration.

COURSE CREDIT FOR PARTICIPATION
If you are currently an undergraduate student at York University you are eligible for 2 course credits upon completion of this study.

CONFIDENTIALITY
Every effort will be made to ensure confidentiality of any identifying information that is obtained in connection with this study.

All questionnaire data will be stored in a locked cabinet at the Child Learning Projects Lab at York University, which is locked at all times and only accessible by project personnel. All other data will be stored on an external hard drive in an encrypted file that will be kept at the Child Learning Projects Lab at York University, which is locked at all times and only accessible by project personnel. All data will be retained for comparison in future studies until either it is no longer of use or until 15 years has passed. After 15 years have passed the data will at that point be destroyed.

PARTICIPATION AND WITHDRAWAL
You can choose whether to be in this study or not. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. You may exercise the option of removing your data from the study. You may also refuse to answer any questions you don't want to answer and still remain in the study. The investigator may withdraw you from this research if circumstances arise that warrant doing so.

RIGHTS OF RESEARCH PARTICIPANTS
You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. This study has been reviewed and received ethics clearance through the York University Research Ethics Board. If you have questions regarding your rights as a research participant, contact:

Sr. Manager & Policy Advisor for the Office of Research Ethics,

5th Floor, York Research Tower, York University

SIGNATURE OF RESEARCH PARTICIPANT

I have read the information provided for the study "Temporal Binding in Audiovisual Speech Integration**" as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.**

____

Name of Participant (please print)

______________________________________ ______________

Signature of Participant Date

SIGNATURE OF PRINCIPAL INVESTIGATOR

______________________________________ ______________

Melissa Ferland Date

For the following trials, press the **RIGHT ARROW** key until the sound and the picture are **out of synch**. Press the key repeatedly, but don't hold down.

➔

If you feel like you've gone "beyond" the point of **asynchrony** you can correct by adjusting with the LEFT arrow key.

When the audio and video are out of synch, press the **NEXT** key to proceed to the next trial.

For the following trials, press the **RIGHT ARROW** key until the sound and the picture are **in synch**. Press the key repeatedly, but don't hold down.

➔

If you feel like you've gone "beyond" the point of **synchrony** you can correct by adjusting with the LEFT arrow key.

When the audio and video are in synch, press the **NEXT** key to proceed to the next trial.