

# Age, not Autism, Influences Multisensory Integration of Speech Stimuli among Adults

Robert M. Jertberg III<sup>1</sup>, Sander Begeer<sup>1</sup>, Hilde Geurts<sup>2</sup>, Bhismadev Chakrabarti<sup>3</sup>, and Erik van der Burg<sup>2</sup>

<sup>1</sup>Vrije Universiteit Amsterdam

<sup>2</sup>University of Amsterdam

<sup>3</sup>University of Reading

November 20, 2023

## Abstract

Differences between autistic and non-autistic individuals in perception of the temporal relationships between sights and sounds are theorized to underlie difficulties in integrating relevant sensory information. These, in turn, are thought to contribute to problems with speech perception and higher level social behaviors. However, the literature establishing this connection often involves limited sample sizes and focuses almost entirely on children. To determine whether these differences persist into adulthood, we compared 469 autistic and 373 non-autistic adults (aged 17 to 75 years). Participants completed an online version of the McGurk/MacDonald paradigm, a multisensory illusion indicative of the ability to integrate audiovisual speech stimuli. Audiovisual asynchrony was manipulated, and participants responded both to the syllable they perceived (revealing their susceptibility to the illusion) and to whether or not the audio and video were synchronized (allowing insight into temporal processing). In contrast with prior research with smaller, younger samples, we detected no evidence of impaired temporal or multisensory processing in autistic adults. Instead, we found that in both groups, multisensory integration correlated strongly with age. This contradicts prior presumptions that differences in multisensory perception persist and even increase in magnitude over the lifespan of autistic individuals. It also suggests that the compensatory role multisensory integration may play as the individual senses decline with age is intact. These findings challenge existing theories and provide an optimistic perspective on autistic development. They also underline the importance of expanding autism research to better reflect the age range of the autistic population.

## Age, not Autism, Influences Multisensory Integration of Speech Stimuli among Adults

Robert M. Jertberg III<sup>1</sup>, Sander Begeer<sup>1</sup>, Hilde M. Geurts<sup>2,3</sup>, Bhismadev Chakrabarti<sup>4,5,6\*</sup>, and Erik Van der Burg<sup>2</sup>

<sup>1</sup>Department of Clinical and Developmental Psychology, Vrije Universiteit Amsterdam, The Netherlands and Amsterdam Public Health Research Institute, Amsterdam, Netherlands

<sup>2</sup>Dutch Autism and ADHD research Center (d'Arc), Brain & Cognition, Department of Psychology, Universiteit van Amsterdam, The Netherlands

<sup>3</sup>Leo Kannerhuis (Youz/Parnassiagroup), The Netherlands

<sup>4</sup>Centre for Autism, School of Psychology and Clinical Language Sciences, University of Reading, UK

<sup>5</sup>India Autism Center, Kolkata, India

<sup>6</sup>Department of Psychology, Ashoka University, India

\*Corresponding author: Robert Jertberg

r.m.jertberg@vu.nl

Keywords: autism, McGurk effect, rapid temporal recalibration, window of perceived synchrony, aging

## Abstract

Differences between autistic and non-autistic individuals in perception of the temporal relationships between sights and sounds are theorized to underlie difficulties in integrating relevant sensory information. These, in turn, are thought to contribute to problems with speech perception and higher level social behaviors. However, the literature establishing this connection often involves limited sample sizes and focuses almost entirely on children. To determine whether these differences persist into adulthood, we compared 469 autistic and 373 non-autistic adults (aged 17 to 75 years). Participants completed an online version of the McGurk/MacDonald paradigm, a multisensory illusion indicative of the ability to integrate audiovisual speech stimuli. Audiovisual asynchrony was manipulated, and participants responded both to the syllable they perceived (revealing their susceptibility to the illusion) and to whether or not the audio and video were synchronized (allowing insight into temporal processing). In contrast with prior research with smaller, younger samples, we detected no evidence of impaired temporal or multisensory processing in autistic adults. Instead, we found that in both groups, multisensory integration correlated strongly with age. This contradicts prior presumptions that differences in multisensory perception persist and even increase in magnitude over the lifespan of autistic individuals. It also suggests that the compensatory role multisensory integration may play as the individual senses decline with age is intact. These findings challenge existing theories and provide an optimistic perspective on autistic development. They also underline the importance of expanding autism research to better reflect the age range of the autistic population.

The brain is under a constant barrage of sensory signals. To most reliably and efficiently interpret these myriad inputs, it must make use of different types of congruence across modalities to tie relevant signals together in a process known as multisensory integration (MSI). MSI can facilitate faster detection of stimuli (Van der Burg et al., 2008), enhancement of perception (Sumbly & Pollack, 1954; Vroomen & de Gelder, 2000), and resolution of ambiguity (Green & Angelaki, 2010; Parise & Ernst, 2017; van Ee et al., 2009), as well as produce potent illusions (McGurk & Macdonald, 1976; Shams et al., 2000; Shipley, 1964). The types of relevant congruence range from basic stimulus features such as spatiotemporal alignment (Slutsky & Recanzone, 2001) to more complex features such as semantic (Iordanescu et al., 2008) and emotional overlap (Jertberg et al., 2019). However, temporal coincidence holds a special place in multisensory perception research, as it has been shown to produce powerful effects of MSI on its own (Van der Burg et al., 2008, 2011; Vroomen & de Gelder, 2004), and many other types of multisensory interactions depend on temporal proximity (Costantini et al., 2016; Munhall et al., 1996; Shams et al., 2000).

Of all the areas MSI affects our daily lives, speech perception may be the most obvious. The integration of visual signals with their auditory counterparts can greatly enhance our understanding of speech (Erber, 1969; Irwin & DiBlasi, 2017; Sumbly & Pollack, 1954; Woodhouse et al., 2008) and even produce powerful multisensory illusions (McGurk & Macdonald, 1976) when the stimuli are presented sufficiently close in time (Munhall et al., 1996). One can experience this influence by simply attempting to understand a speaker across a noisy room with or without one's eyes open. It is in such boisterous environments, in which the reliability of relevant auditory signals is compromised by competing inputs, that the most benefit is afforded by the integration of visual information (Erber, 1969; MacLeod & Summerfield, 1987; Sumbly & Pollack, 1954). It is notable, then, that it is precisely these circumstances in which autistic individuals struggle most with speech perception (Alcántara et al., 2004; Fadeev et al., 2023; Mamashli et al., 2017; Ruiz Callejo et al., n.d.). Autism is of particular interest to our understanding of this intersection of MSI, speech perception, and temporal processing because it appears to involve differences on some level in all three categories (Feldman et al., 2018; Kwok et al., 2015; Rapin & Dunn, 2003; Sperdin & Schaer, 2016; Zhou et al., 2018). Our understanding of these issues is, conversely, of particular significance to those with autism because of the manner in which they may contribute to broader social and communication differences.

Many individuals with autism demonstrate impairments in speech processing (Kwok et al., 2015; Rapin & Dunn, 2003; Sperdin & Schaer, 2016) as well as attenuated multisensory effects, particularly when young (Feldman et al., 2018). In light of the crucial role temporal dynamics have been shown to play in MSI, and MSI in turn on speech perception, differences in temporal processing may be underlying factors in both of these disparities. It is worth noting that even though auditory and visual information may originate from the same source, these signals never arrive perfectly simultaneously to the brain. Light travels faster than sound, but auditory stimuli have a lower signal transduction latency (Jain et al., 2015; Kemp, 1973), so the brain must be both tolerant and adaptable to varying degrees of asynchrony between sensory streams to allow integration of relevant stimuli. Tolerance to asynchrony is seen in the window of perceived synchrony (WPS), which is the range of stimulus onset asynchronies (SOAs) over which participants are still likely to perceive multisensory signals as simultaneous. Narrowing of this window, which can be seen as a refinement of temporal processing acuity, occurs during typical development (Hillock et al., 2011; Hillock-Dunn & Wallace, 2012; Lewkowicz & Flom, 2014), but is both delayed and diminished among those with autism (de Boer-Schellekens et al., 2013; Foss-Feig et al., 2010; Stevenson et al., 2014). However, recent research challenges the degree to which this applies to autistic adults (Weiland et al., 2022; Zhou et al., 2022).

Adaptability to asynchrony is seen in temporal recalibration (Fujisaki et al., 2004; Vroomen et al., 2004), an effect in which the point of subjective simultaneity (PSS), where participants are most likely to perceive audiovisual inputs as synchronized, shifts according to prior experience. For example, after hearing an auditory stimulus such as a beep leading a visual stimulus such as a flash, a participant will be more likely to perceive a similarly leading beep as simultaneous with a flash (Van der Burg et al., 2013). This effect also extends to more complex speech stimuli (Van der Burg & Goodbourn, 2015). Some studies have found that this rapid temporal recalibration effect is also diminished in those with autism (J. Noel et al., 2017; Turi et al., 2016), although that with the largest adult sample did not (Weiland et al., 2022), again raising questions about the persistence of temporal processing differences.

Together, these differences in MSI and temporal processing have given rise to theories that posit that basic sensory factors may contribute to the higher level social differences seen in autism via their influence on language and communication (Baum et al., 2015; Stevenson et al., 2018). Stevenson et al. (2018) found that autistic children’s WPS width correlates negatively with the degree of audiovisual integration they experience, which in turn correlates positively with recognition of speech in noise. They took this as evidence that MSI mediates an influence of temporal processing acuity on speech perception in autism. Given the evidence of these cascading effects, understanding the relationship between temporal processing differences, MSI, and speech perception is crucial to illuminating the broader autistic behavioral profile.

Most prominent among the paradigms used to investigate MSI with speech stimuli is the McGurk/MacDonald effect (McGurk & Macdonald, 1976). This effect occurs when participants are presented with conflicting phonemes (the smallest auditory components of speech) and visemes (their visual counterparts), leading to an illusion in which what is heard is influenced by what is seen. For example, the presentation of a /ba/ phoneme with a /ga/ viseme tends to lead participants to report hearing the phoneme /da/. This phenomenon, dubbed a fusion, is highly dependent upon temporal alignment (Munhall et al., 1996), and has been shown to correlate negatively with the width of the WPS (Stevenson et al., 2012), which is, again, wider on average among autistic individuals. As such, it is unsurprising that autistic individuals have shown attenuated susceptibility to the McGurk/MacDonald illusion, at least as children.

In a meta-analysis focusing on the McGurk/MacDonald illusion (Zhang et al., 2019), it was found not only that autistic individuals show less susceptibility to the effect, but also that the magnitude of this between group difference increases with age. This led the authors to conclude that while non-autistic people continue to develop in their ability to integrate audiovisual speech stimuli, autistic individuals’ progress may be hampered by heightened attention to local details and reduced orientation to social information. However, it bears noting that 8/9 studies included in their meta-analysis had child samples, and that the only adult study found no difference between groups in the strength of the McGurk/MacDonald effect (Saalasti et al., 2012). Additionally, two studies not included in the meta-analysis (Keane et al., 2010; Stevenson et al., 2018)

did not find a difference between groups in susceptibility to the illusion. Notably, this includes the study with the largest previous sample size (Stevenson et al., 2018) and one of the few with an adult sample (Keane et al., 2010). Such inconsistencies raise questions about the degree to which MSI findings with autistic children extend to adults. These are highlighted by findings that autistic children may catch up to their non-autistic peers in their ability to integrate audiovisual speech signals embedded in noise by early adolescence (Fuxe et al., 2015). Theories that posit that it is persistent MSI deficits that drive difficulties with speech perception and other higher order differences between autistic and non-autistic adults are challenged by these findings. Beyond theory, because multisensory training has been shown to be highly effective (Nava et al., 2020; O’Brien et al., 2023; Setti et al., 2014), understanding the ages at which these differences exist is essential to tailoring therapeutic interventions for autistic individuals.

In addition to age, a significant factor in the heterogeneity of findings may be sample size. In a review of McGurk/MacDonald studies, Magnotti & Beauchamp (2018) demonstrated that a publication bias towards significant results would produce a vast overestimation of real population differences given the small sample sizes conventional in this field of research. This led them to conclude that the published estimates of the differences between groups in MSI measured using the McGurk/MacDonald effect are inflated. They argued that to alleviate this effect size inflation and enhance replicability, sample sizes must be increased considerably.

In order to examine the degree to which findings from previous studies with children, limited in both scope and age range, extend to autistic adults, we recruited the largest sample to date for a study investigating differences between autistic and non-autistic adults in temporal processing and audiovisual integration of speech stimuli. We measured these using a version of the McGurk/MacDonald task involving manipulation of SOA and both syllable and simultaneity judgments. This allowed us to compare the rate at which the illusion occurs as well as the likelihood for participants to perceive stimuli as synchronized, their WPS, and the effects of rapid temporal recalibration. We predicted diminished susceptibility to the McGurk/MacDonald effect, blunted temporal acuity (i.e. a wider WPS), and an attenuated effect of temporal recalibration in autistic versus non-autistic participants.

## Methods

*Participants.* We recruited 666 autistic participants via the Netherlands Autism Register (NAR, <https://nar.vu.nl/>) and 517 non-autistic participants via the NAR as well as Prolific Academic. The autistic participants reported a formal diagnosis by an independent, qualified clinician. The non-autistic participants reported no diagnosis of autism. All 1183 participants were fluent in Dutch. NAR participants received €15 gift cards and Prolific Academic participant were paid £15. All participants were naïve to the purpose of the study and gave informed consent prior to the experiment. The experiment was approved by the ethical committee from the Vrije Universiteit Amsterdam (VCWE-2020-041R1) in accordance with all guidelines and regulations as specified in the Netherlands Code of Conduct for Research Integrity.

147 participants were excluded because either the demographic information, the AQ-28 data, or the ICAR data were missing. The data from another 167 participants were excluded from further analyses, since their performance on the task matched one or multiple exclusion criteria, as pre-registered in As Predicted (#102341). For precise information on the number of participants from each group excluded (and the reasons for their exclusion), refer to Supplementary Figure 1. The demographic information for the remaining 869 participants is depicted in Table 1, and the age distribution of the groups is shown in Figure 1. Note that the non-autistic sample used in this study comprises of the entirety of that in our earlier study (Jertberg et al., 2023), in addition to later recruits.

Table 1: demographic breakdown by group.

	Autism	Autism	No autism	No autism	Statistics
	(n = 469)	(n = 469)	(n = 373)	(n = 373)	
Sex					

	Autism	Autism	No autism	No autism	
females (%)	64.5		51.2		$t(867) = 4.294, p < .001$
males (%)	36.5		48.8		
Mean age (SD)	44.9	(14.0)	33.1	(13.0)	$t(867) = 12.669, p < .001$
Age Range	18-75		18-74		
Mean AQ-28 (SD)	83.3	(10.7)	59.9	(11.5)	$t(867) = 30.828, p < .001$
AQ-28 Range	46-109		30-96		
Mean ICAR (SD)	0.53	(0.21)	0.52	(0.19)	$t(867) = .396, p = .692$
ICAR Range	0-0.94	0.06-0.94	0.06-0.94		

SD: standard deviation

AQ: Autism Quotient

ICAR: International Cognitive Ability Resource (abbreviated intelligence quotient test)

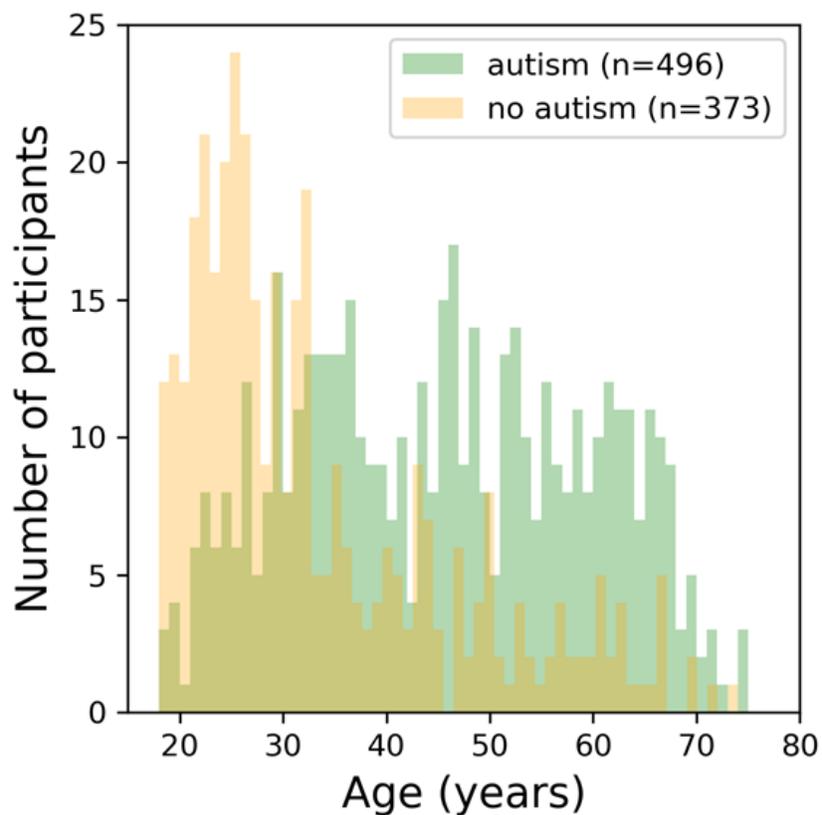


Figure : Age distribution of autistic and non-autistic participants.

*Apparatus and stimuli.* The experiment was programmed and conducted online using the Neurotask platform ([www.neurotask.com](http://www.neurotask.com)). Participants completed the task using their own hardware. The stimuli were taken from Hillock-Dun, Grantham, and Wallace (2016) and included videos of an actress saying the syllable /ga/ with either the corresponding audio (for congruent trials) or the audio of the same actress saying the syllable

/ba/ dubbed over the video (for incongruent trials). Stimulus onset asynchrony (SOA) was either -500, -260, 0, 260 or 500 ms. Here, negative signifies that audition was leading, and vice versa. All videos lasted 2000 ms.

*Procedure*. Trials began with a black fixation cross in the center of a white screen for 1000 ms. Subsequently, the video played for 2000 ms. After its conclusion, participants were prompted to make two self-paced responses. First, they were asked whether they heard /ba/, /da/, or /ga/ by pressing either the b, d, or g-key, respectively. Second, they were asked whether the video and audio were synchronized or not by pressing the 1 or 0-key, respectively. The following trial was initiated as soon as the synchrony judgment was given. The two congruency types and five SOAs produced a total of ten unique trial types. After reading the written instructions and completing ten practice trials (one of each stimulus type), participants completed ten repetitions of each trial type, for a total of 100 experimental trials in randomized order. The experiment lasted approximately 7 minutes and was part of a larger battery of online tasks (not reported here). Figure 2 illustrates an example trial sequence.

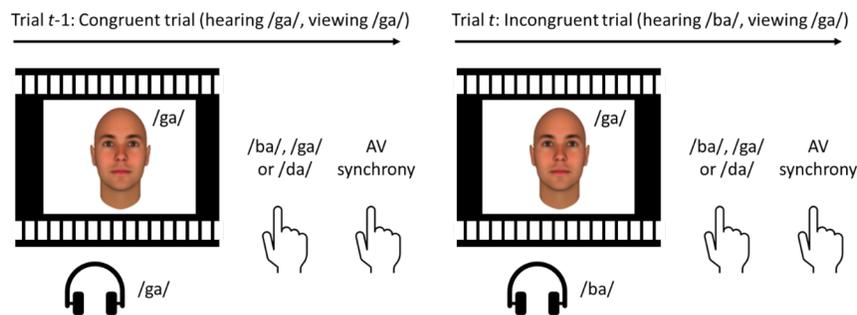


Figure : Two example trials used in the present study. The participants viewed a movie of an actress mouthing the syllable /ga/. On half the trials, audio corresponded to the video (i.e., congruent trials), whereas on the remaining trials, the audio of the syllable /ba/ was played (i.e., incongruent trials). The onset between the voice and the lip movement was manipulated, and the participants were instructed to make two judgments. First, participants reported whether they heard /ba/, /ga/, or /da/. Subsequently, they judged whether the voice was synchronized with the lip movements. Note: a computer generated face overlays the actress for privacy reasons.

## Results

*Syllable Responses.* Figure 2a illustrates the mean proportion of /da/ responses (the classical McGurk/Macdonald fusion) for each group as a function of SOA for congruent and incongruent trials. Figure 2b shows the mean proportion of /da/ responses on incongruent trials, collapsed across SOAs for each group, as a function of age (divided into bins of 10 years). This provides insight into the influence of age on the occurrence of the illusion, and allows a comparison between groups that is not confounded by their disparity in age (see Table 1 and Figure 1).

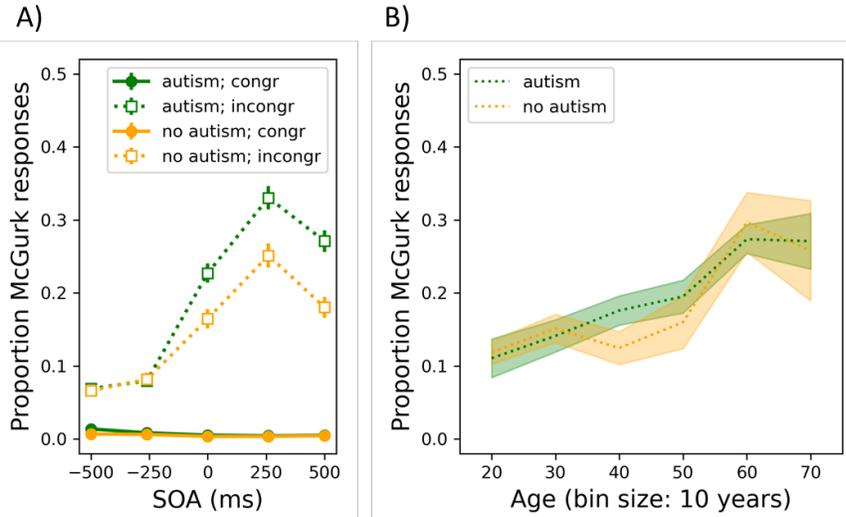


Figure 2. A) Proportion of /da/ responses per group as a function of SOA for congruent and incongruent trials. Here, negative SOAs indicate that the voice was leading the lip movements, and vice versa. B) Proportion of /da/ responses per group as a function of age-bin (bin size is 10 years) for incongruent trials (collapsed across SOAs). The error-bars reflect the standard error of the mean.

We conducted a repeated measures ANOVA on the mean proportion of /da/ responses with SOA and congruency as within subjects variables and group as a between subject factor using Just Another Statistical Program (JASP) (Love et al., 2019). Here, and elsewhere in the manuscript, alpha was set to .05, and  $p$  values were Hyunh-Feldt corrected to avoid sphericity violations. As seen in Figure 2a, the rates of /da/ responses varied significantly as a function of SOA ( $F(1,867) = 243.632, p < .001$ ) and congruency ( $F(1,867) = 529.059, p < .001$ ). What is more, there was a significant SOA  $\times$  congruency interaction ( $F(1,867) = 274.805, p < .001$ ), such that the McGurk/MacDonald illusion occurred most frequently in incongruent trials with a slight visual lead. However, follow-up two-tailed  $t$ -tests confirmed that there was a significant congruency effect (and, hence, McGurk/MacDonald effect) at all SOAs (minimum  $t(868) = 10.471$ , all  $p < .001$ ).

Between the groups, autistic participants experienced the illusion more frequently ( $F(1,867) = 11.448, p < .001$ ). There were also interactions between group  $\times$  congruency ( $F(867) = 9.344, p = .002$ ), group  $\times$  SOA ( $F(867) = 12.012, p < .001$ ), and group  $\times$  congruency  $\times$  SOA ( $F(867) = 14.662, p < .001$ ). However, due to the disparity in the age distributions between groups (see Table 1), we conducted an exploratory analysis with age as a covariate, and found that the group effect was no longer significant, nor were any of the previously significant interactions including group. Instead, we discovered a strong effect of age ( $F(1,866) = 54.194, p < .001$ ), wherein older participants from both groups tended to experience the illusion more frequently than younger ones (see Figure 2b). Age also interacted with SOA ( $F(1,866) = 37.815, p < .001$ ), congruency ( $F(1,866) = 50.695, p < .001$ ), and SOA  $\times$  congruency ( $F(1,866) = 38.530, p < .001$ ), augmenting all of their effects. Further evidence that age (rather than group) explains the pattern of results above is provided using age matched and Bayesian approaches in the Supplementary Materials (see Supplementary Tables 1 and 2).

Note that this section focused on the /da/ responses, as they reflect the classic McGurk/MacDonald illusion, and rates of visual capture (/ga/ responses in incongruent trials) were extremely low for both groups. For transparency, statistics and figures reflecting the rates of /ba/ and /ga/ responses can be found in the Supplementary Materials.

*Simultaneity judgments.* Figure 3a plots the mean proportion of simultaneity judgments as a function of SOA

and congruency for both groups. Figure 3b shows the mean proportion of simultaneity judgments collapsed across SOAs for both groups and congruency conditions as a function of age, divided into bins of 10 years.

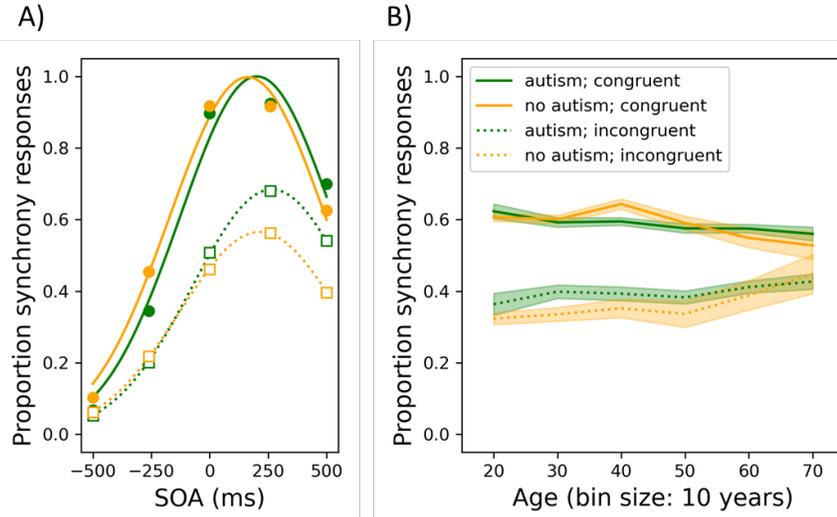


Figure 3. A) Proportion of synchrony judgments per group as a function of SOA for congruent and incongruent trials. Here, negative SOAs indicate that the voice was leading the lip movements, and vice versa. B) Proportion of synchrony judgments (collapsed across SOAs) as a function of age and congruency for each group (bins of 10 years). The error-bars reflect the standard error of the mean.

We conducted a repeated measures ANOVA on the mean proportion of synchrony responses with SOA and congruency as within subjects variables, group as a between subjects variable, and age as a continuous covariate. This yielded a significant main effect of SOA ( $F(1,866)=208.845, p < .001$ ). The rate of synchrony responses across the SOAs formed a typical Gaussian distribution with a slight visual leading offset (see Figure 3a). Additionally, the proportion of synchrony responses was much higher when the stimuli were congruent than incongruent ( $F(1,866)=377.547, p < .001$ ). Congruency also interacted with SOA, such that its effect was most pronounced when stimuli occurred simultaneously or with a slight visual lead ( $F(1,866)=48.803, p < .001$ ), and with group ( $F(1,866)=17.004, p < .001$ ). A follow up t-test comparing the difference between mean simultaneity judgment response rates for congruent and incongruent trials according to group revealed that the effect of congruency was greater for non-autistic participants than autistic ones ( $t(867)=6.76, p < .001$ ; see Figure 3a). SOA also interacted with group, with the differences between autistic and non-autistic participants emerging at the mid-range SOAs ( $F(1,866)=7.204, p < .001$ ), which is logical given that the longer SOAs were much more obvious to both groups. Both congruency ( $F(1,866)=31.101, p < .001$ ) and SOA ( $F(1,866)=55.233, p < .001$ ) also significantly interacted with age. Finally, we detected a significant three way interaction between congruency  $\times$  SOA  $\times$  age ( $F(1,866)=6.792, p < .001$ ). The difference between congruent and incongruent trials was greater at younger ages. Because SOA interacts with all significant factors due to the nature of simultaneity judgment tasks, these effects were not explored further.

*Window of Perceived Synchrony.* We fitted a Gaussian distribution to the synchrony distribution for each individual by using the `curve_fit` function from the `scipy` Python module to estimate a WPS, amplitude and PSS. Figure 3b illustrates the mean WPS as a function of age (bin size = 10 years) for participants with and without autism. Note that for one participant, the fitting procedure was not successful, resulting in exclusion from further analyses.

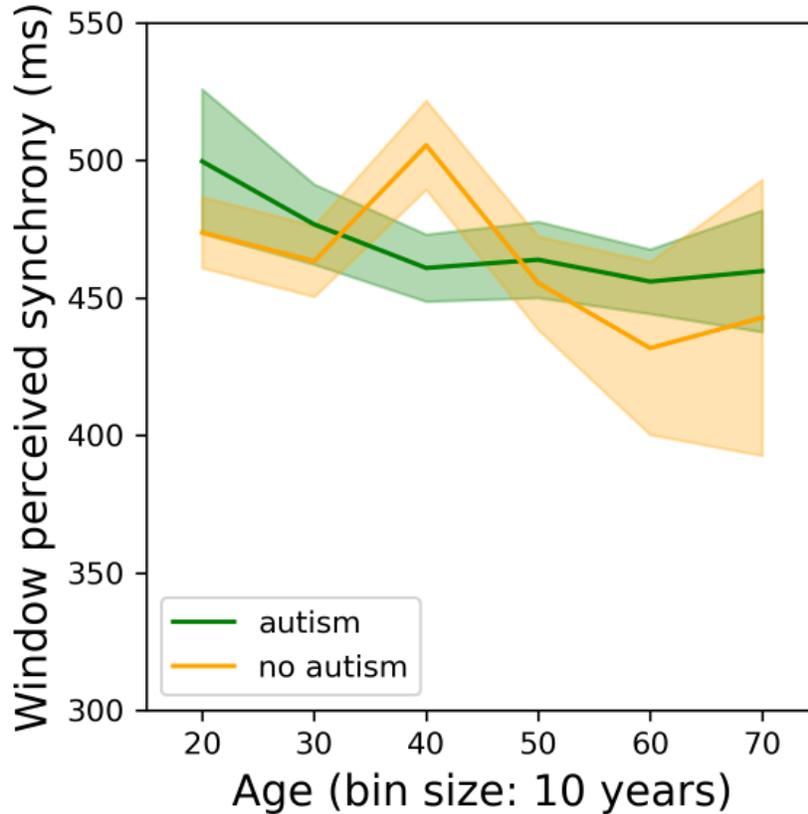


Figure 4: Mean WPS for autistic and non-autistic participants relative to age (in bins of 10 years).

We conducted an ANCOVA on the mean WPS with group as a between subjects variable and age as a covariate. The ANCOVA yielded no significant effect of group ( $F(1,865)=0.227, p=.634$ ) or age ( $F(1,865), p=.053$ ) on the WPS.

*Rapid temporal recalibration.* To measure rapid temporal recalibration, we excluded the first trial and split the rest into two categories: those following trials with either a -500 or -260 ms SOA (audition leads), and those following trials with a 260 or 500 ms SOA (vision leads). We then fit Gaussian functions (as described previously) to each modality order condition (see Figure 5a) and calculated the mean PSS by identifying the SOA at which each function reaches its peak. Rapid temporal recalibration was quantified as the difference in mean PSS between categories (i.e. PSS audition leads-PSS vision leads; see also (Van der Burg et al., 2013, 2018)). Note that one participant was excluded due to fitting issues.

Accordingly, Figure 5a reflects the mean proportion of synchrony responses as a function of SOA for each previous modality order and group (collapsed across congruency conditions). Figure 5b reflects the mean PSS derived from these synchrony distributions according to group and previous modality order. Figure 5c shows the  $\Delta$  PSS (i.e., rapid temporal recalibration) as a function of age (in bins of 10 years) for each group.

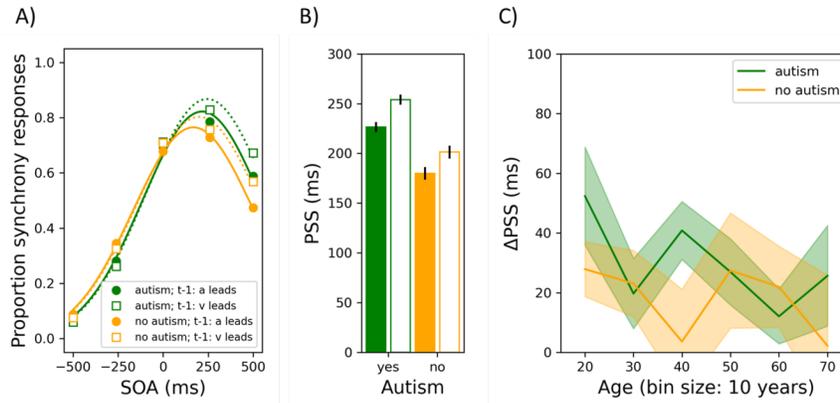


Figure 5. A) Proportion of synchrony judgments per group as a function of stimulus onset asynchrony (SOA), relative to the modality order of the preceding trial (collapsed across congruency conditions). Here, negative SOAs indicate that the voice was leading the lip movements, and vice versa. B) Point of subjective simultaneity (PSS) per group relative to the modality order on the preceding trial. C) Magnitude of recalibration effect (i.e., change in PSS between preceding modality orders) per group relative to age (in bins of 10 years). The error-bars reflect the standard error of the mean.

We conducted a repeated measures ANOVA on the mean PSS with previous modality order as a within subjects variable, group as a between subjects variable, and age as a covariate. We found a significant main effect of modality order ( $F(1,865)=13.823, p < .001$ ), such that the PSS was smaller when audition led in the previous trial than when vision led, as Figure 5b illustrates. Rapid temporal recalibration did not differ between groups, as the modality order  $\times$  group interaction failed to reach significance ( $F(1,865)=1.968, p = .161$ ). Autistic participants showed a larger average PSS (240 ms) than non-autistic participants (190 ms) overall ( $F(1,865)=13.866, p < .001$ ), reflecting a preference for a greater visual lead. Age did not significantly affect the magnitude of rapid temporal recalibration, as it did not interact with previous modality order ( $F(1,865)=2.624, p = .106$ ; see Figure 5c). However, older participants did have a higher overall PSS ( $F(1,865)=51.770, p < .001$ ). The mean PSS for participants above the median age was 249 ms, compared to a mean of 188 ms for those below the median age.

## Discussion

Based on studies primarily with children, it has been hypothesized that autistic individuals show attenuated MSI, particularly for speech stimuli. Our results provide compelling evidence that the presumption this difference extends into adulthood is unfounded. Once controlling for age, we found no significant difference between autistic and non-autistic individuals in susceptibility to the McGurk/MacDonald illusion. Because this ran against the grain of the findings of the largest meta-analysis on the topic (Zhang et al., 2019), we confirmed that group was not a significant factor in our results using both age matched and Bayesian follow up analyses (see Supplementary Materials). While Zhang et al. (2019) concluded that the difference between groups actually increases in magnitude with age, it only included one study with adults (Saalasti et al., 2012), which the original authors did not take as evidence for a difference in the strength of the McGurk/MacDonald effect. Moreover, some findings suggest that differences between autistic and non-autistic individuals in MSI may be resolved during adolescence (Foxe et al. 2015; Taylor, Isaac, & Milne, 2010). Our findings with adults are consistent with the trajectory of improvement these results imply.

Instead of a difference according to group, we found evidence that the degree of MSI increases with age (with the average rate of the illusion nearly tripling from the youngest to oldest participants) for both autistic and non-autistic individuals. While an increase in the rate of the McGurk/MacDonald effect between younger and older adults has been detected in non-autistic participants (Setti et al., 2013), this is the first study comparing them in both autistic and non-autistic samples. The near perfect overlap of the correlations

between age and MSI between groups serves as compelling evidence that while autistic children may not experience the development of visual influence on speech perception as early as their non-autistic peers, autistic adults do show comparable visual influence into their older years. These findings of similar age effects across adulthood resonate with recent longitudinal research suggesting similar cognitive aging profiles between autistic and non-autistic individuals (Torenvliet et al., 2023).

The reason for such a strong effect of age on the rate of the illusion could be reduced reliability of the auditory signal resulting from the progressive hearing loss common in aging, which often goes uncorrected (Walling & Dickson, 2012). The comparative reliability of auditory and visual inputs has been shown to affect the rate at which the McGurk/MacDonald effect occurs, and their respective influence shifts during development (Hirst et al., 2018). Additionally, MSI may also serve a compensatory role in speech perception as hearing declines. Both notions are supported by research showing an increase in MSI and visual dominance later in life (Diaconescu et al., 2013), as well as enhanced susceptibility to the McGurk/MacDonald effect associated with age-related hearing loss (Rosemann & Thiel, 2018; Stropahl & Debener, 2017). Cortical reorganization leading to increased functional connectivity between auditory and visual regions may facilitate these effects in those with age-related hearing loss (Puschmann & Thiel, 2017). It is encouraging that MSI appears to serve this compensatory role as effectively in autistic adults as non-autistic ones.

Another potential factor in differences between our findings and others is the threat of an attentional confound. Autistic children have been shown to demonstrate an atypical preference for non-social stimuli, viewing faces less frequently than their non-autistic peers (Gale et al., 2019; Vacas et al., 2021). Additionally, in two McGurk/MacDonald studies using eye-tracking, it was found that autistic children attended less to the pertinent areas of the face than non-autistic ones (Feng et al., 2021; J. R. Irwin et al., 2011), partially explaining differences in susceptibility to the illusion. Accordingly, studies that do not control for visual attention may overstate differences in MSI. A merit of our design is that while we do not directly measure eye movements, our simultaneity judgment task requires participants to attend to the mouth during trials. The performance of participants on this task, resembling a typical Gaussian distribution peaking near simultaneity, suggests that they were indeed attending to the faces. While the addition of eye-tracking would help to confirm this, in online experiments such as ours, where it is not possible (due to privacy reasons), the addition of an simultaneity judgment task provides an excellent means of reducing the risk of attentional differences being conflated with differences in MSI.

Beyond our findings with regard to the McGurk/MacDonald illusion, our results have illuminated much about the nuances of temporal processing and how they compare between autistic and non-autistic individuals. In many ways, our results remained consistent with standard findings in fundamental temporal processing research. Synchrony distributions followed a typical Gaussian shape, peaking with a slight visual lead, as is consistently found with audiovisual stimuli (Dixon & Spitz, 1980; Slutsky & Recanzone, 2001; Zampini et al., 2005). Incongruent stimuli were perceived as synchronous significantly less frequently than congruent ones, as was shown in other studies measuring simultaneity judgments for McGurk/MacDonald stimuli (Jertberg et al., 2023; Van Wassenhove et al., 2007). Rapid temporal recalibration was detected, with the PSS shifting according to the previous modality order (Van der Burg et al., 2013, 2015, 2018). However, our results also captured novel differences between groups.

Firstly, with regard to synchrony distributions, we found differences in the magnitude of the effect of congruency according to group. Both groups were less likely to perceive incongruent stimuli as synchronized, but this effect was particularly pronounced for the non-autistic ones. This was even true at 0 ms, when participants went from recognizing the physical simultaneity of the stimuli 91.8% to 46.2% of the time in the non-autistic group, and 89.7% to 50.8% of the time in the autistic group. This suggests a profound interference of phonetic incongruence on basic temporal processing, one that van Wassenhove et al. (2007) attributed to a weaker correlation between the facial kinematics (what is seen) and acoustic dynamic envelope (what is heard). But why does the magnitude of this difference vary between autistic and non-autistic individuals, when the disparity between these factors remains the same?

One interpretation might be that the autistic participants simply have a lower temporal resolution than

the non-autistic ones, and therefore less room for interference in temporal processing. However, we did not replicate findings that the WPS, the common measure of temporal acuity, differs between groups, so this interpretation is not supported by our results. Alternatively, these differences could be due to impoverished lip reading ability, which has been found to account for some or all of the disparity in susceptibility to the McGurk/MacDonald effect in autistic children (Iarocci et al., 2010; Smith & Bennetto, 2007). Impoverished lip reading ability may be viewed as a weaker association between a viseme and its associated phoneme. This may translate into a diminished incongruence effect, as the autistic participants would be less sensitive to the difference driving it. That being said, were this the case, one might also expect an attenuated visual influence of the visemes, and hence a lower rate of the McGurk/MacDonald effect, in the autistic participants. As such, further research into the lip-reading abilities of autistic adults and their potential influence on temporal processing of audiovisual speech stimuli is necessary.

Delving deeper into the temporal dynamics at play, we did not detect the differences between groups in the WPS or rapid temporal recalibration formerly reported. With regard to the WPS, the largest meta-analysis to date examining potential differences between autistic and non-autistic participants found a consistent enlargement of its width among those with autism (Zhou et al., 2018), suggesting blunted temporal acuity. However, there was again a limited number of studies germane to the topic (with only four studies investigating the audio-visual WPS), most had small samples (ranging from 32-64 participants), and all of them focused on children. More recent research involving adults paints a different picture. Two studies (Weiland et al., 2022; Zhou et al., 2022) with larger samples of adults found no difference between autistic and non-autistic participants in the width of the WPS, suggesting that autistic individuals may also catch up in the honing of temporal processing by the time they reach adulthood. A very similar pattern emerges with rapid temporal recalibration, where smaller studies with younger participants found differences between autistic and non-autistic individuals (J. Noel et al., 2017; Turi et al., 2016), but the largest adult study did not (Weiland et al., 2022). However, the research here is more limited, and Weiland et al. (2022) also recruited from the NAR, so their sample may partially overlap with ours. Accordingly, further examination of potential differences between autistic and non-autistic individuals in the WPS and rapid temporal recalibration (and the possibility of their resolution) is warranted.

We did, however, detect a difference between groups in the overall mean PSS value. Autistic participants showed a greater mean PSS, irrespective of stimulus type, suggesting a heightened sensory preference for visual lead. This finding may also explain the difference in the magnitude of the congruence effect between groups, at least in part, given that it was largest with a slight visual lead. The two most obvious potential explanations for the PSS difference would be either faster processing of auditory information or slower processing of visual information in autism. Unfortunately, there is little research into this topic, at least with regard to speech stimuli. Furthermore, as always, the studies that do exist tend to focus on children. We could not find studies investigating differences in auditory processing speed between autistic and non-autistic individuals. However, those that have looked into visual processing speed show faster visual processing in autism, if anything (Foss-Feig et al., 2013; Samson et al., 2011). Granted, this was in more basic sensory processing, such as recognition of motion (Foss-Feig et al., 2013), but it certainly does not provide evidence for a unisensory processing speed explanation. More research should be done into processing speed for the different sensory modalities in autism, both with simple and complex stimuli.

An alternative explanation falls more in line with our discussion of differences in representation of visual speech stimuli. If autistic individuals have differently developed representations of verbal lip movements (as suggested by their weaker lip-reading abilities) and weaker associations between them and the sounds of language, as suggested by van Wassenhove et al. (2007), it stands to reason that it might take them more time to interpret lip-movements and integrate them with their corresponding vocal sounds. This might translate into a greater sensory preference for visual lead when processing speech stimuli. However, given the dearth of evidence provided by the literature on the alternative sensory processing speed hypotheses, this interpretation is highly speculative, and further research should explore the factors contributing to differences in PSS between autistic and non-autistic individuals. An excellent starting point would be to see whether this difference is unique to speech stimuli (supporting the notion that it is driven by differences in

representation of verbal mouth movements) or whether it applies more broadly to simple audiovisual stimuli (suggesting a basic sensory processing speed explanation).

While the large size of our sample and sound experimental design are strengths of our study, it is, of course, not without its limitations. Firstly, this experiment was part of a larger online experimental battery, which placed constraints on the number of trials participants could complete. A larger number of trials and range of SOAs would have allowed more sophisticated analyses of temporal processing and higher resolution representation of participants' WPS and recalibration effects. This also would have allowed us to investigate potential effects of congruence on recalibration and, conversely, of recalibration on the likelihood for participants to perceive the illusion. A related shortcoming of this study is that the time limitation meant we were unable to include unisensory trial types. These allow a researcher to quantify participants' ability to identify visemes and phonemes on their own, which is important as autistic children have shown differences in their lip-reading abilities when compared to non-autistic ones (Fuxe et al., 2015; Iarocci et al., 2010; J. R. Irwin et al., 2011; Smith & Bennetto, 2007; Taylor et al., 2010). While we did not find a difference in audiovisual speech processing anyway, we are unable to speak to the influence of unisensory factors in our findings. Future research should assess the degree to which autistic and non-autistic adults may differ in their perception of visemes and phonemes exclusively as well as in combination to better isolate any potential differences in MSI.

Finally, it must be noted that well-educated adults with comparatively high IQs are overrepresented in the NAR sample (Scheeren et al., 2022). It could be argued that our sample is therefore less likely to capture the segments of the autistic population that may suffer from the most severe deficits in areas like MSI. In particular, those with intellectual disabilities are underrepresented. That being said, the parity in IQ (as estimated by the ICAR) between groups does mean that our results can speak directly to differences resulting from the sensory factors related to autism that are not confounded by cognitive ones related to intellectual impairment. If differences between groups in MSI were found to be driven by the individuals who are underrepresented here, it would be unclear whether they were due to autism or intellectual disability.

In conclusion, our study has confirmed several findings with regard to basic temporal and multisensory processing, as well as challenged the degree to which reported differences between autistic and non-autistic children in these areas extend to adulthood. Our findings that MSI, temporal processing acuity, and rapid temporal recalibration all seem to be intact among autistic adults is highly encouraging given the essential role MSI has in speech perception and compensation for the unisensory deterioration that is inevitable with aging. Additionally, our novel findings with regard to differences in the degree of interference in temporal processing posed by incongruent stimuli and in the mean PSS values between groups are intriguing, and demand further research to disentangle alternative explanations. Understanding these phenomena is of paramount importance given the relevance of temporal and multisensory processing to higher order social factors and the proven efficacy of multisensory training. Pinpointing the age at which related interventions may be of use is crucial to their proper timing, which our findings suggest is prior to adulthood. Finally, our results underline the importance of expanding sample sizes and age ranges in autism research. Restricting our focus to children does a disservice to the vast majority of individuals with autism and leads to a limited understanding of the broader trajectory of this developmental condition, which can only be broadened by giving autistic adults the attention they deserve.

### **Acknowledgements**

Special thanks to the volunteer participants at the Netherlands Autism Register, as well as Jacob Feldman and Mark Wallace for the original stimuli from Hillock-Dunn, Grantham, and Wallace (2016).

### **Author Contributions**

Design: RMJ, EVDB, HMG, BC, SB

Software development: RMJ, EVDB

Analysis: RMJ, EVDB

Writing: RMJ, EVDB

Editing: RMJ, EVDB, HMG, BC, SB

### Data Availability Statement

Our raw data includes sensitive personal information and identifiers not suitable for open access. However, upon request to the corresponding author (with a statement of research intent), access to an anonymized data set with all the information necessary to replicate analyses and the Python scripts we used for our analyses will be provided.

### Competing Interests Statement

The authors declare that they have no financial or non-financial competing interests that might compromise the objectivity or integrity of the research project.

### References

- Alcántara, J. I., Weisblatt, E., Moore, B., & Bolton, P. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* , 45 , 1107–1114. <https://doi.org/10.1111/j.1469-7610.2004.t01-1-00303.x>
- Baum, S. H., Stevenson, R. A., & Wallace, M. T. (2015). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Progress in Neurobiology* , 134 , 140–160. <https://doi.org/10.1016/j.pneurobio.2015.09.007>
- Colonus, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience* , 16 (6), 1000–1009. <https://doi.org/10.1162/0898929041502733>
- Costantini, M., Robinson, J., Migliorati, D., Donno, B., Ferri, F., & Northoff, G. (2016). Temporal limits on rubber hand illusion reflect individuals' temporal resolution in multisensory perception. *Cognition* , 157 , 39–48. <https://doi.org/10.1016/j.cognition.2016.08.010>
- 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. <https://doi.org/10.3389/fnint.2013.00008>
- Diaconescu, A. O., Hasher, L., & McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *NeuroImage* , 65 , 152–166. <https://doi.org/10.1016/j.neuroimage.2012.09.057>
- Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception* , 9 (6), 719–721. <https://doi.org/10.1068/p090719>
- Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research* , 12 (2), 423–425.
- Fadeev, K. A., Goyaeva, D. E., Obukhova, T. S., Ovsyannikova, T. M., Shvedovskiy, E. F., Nikolaeva, A. Y., Davydova, E. Y., Stroganova, T. A., & Orekhova, E. V. (2023). Difficulty with Speech Perception in the Background of Noise in Children with Autism Spectrum Disorders Is Not Related to Their Level of Intelligence. *Clinical Psychology and Special Education* , 12 (1), 180–212. <https://doi.org/10.17759/cpse.2023120108>
- Feldman, J. I., Dunham, K., Cassidy, M., Wallace, M. T., Liu, Y., & Woynaroski, T. G. (2018). Audiovisual multisensory integration in individuals with autism spectrum disorder: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews* , 95 , 220–234. [psych. https://doi.org/10.1016/j.neubiorev.2018.09.020](https://doi.org/10.1016/j.neubiorev.2018.09.020)
- Feng, S., Lu, H., Wang, Q., Li, T., Fang, J., Chen, L., & Yi, L. (2021). Face-viewing patterns predict audiovisual speech integration in autistic children. *Autism Research : Official Journal of the International Society for Autism Research* , 14 (12), 2592–2602. <https://doi.org/10.1002/aur.2598>

- 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. Embase. <https://doi.org/10.1007/s00221-010-2240-4>
- Foss-Feig, J. H., Tadin, D., Schauder, K. B., & Cascio, C. J. (2013). A Substantial and Unexpected Enhancement of Motion Perception in Autism. *Journal of Neuroscience* , 33 (19), 8243–8249. <https://doi.org/10.1523/JNEUROSCI.1608-12.2013>
- Foxe, J. J., Molholm, S., Del Bene, V. A., Frey, H.-P., Russo, N. N., Blanco, D., Saint-Amour, D., & Ross, L. A. (2015). Severe multisensory speech integration deficits in high-functioning school-aged children with Autism Spectrum Disorder (ASD) and their resolution during early adolescence. *Cerebral Cortex (New York, N.Y. : 1991)* ,25 (2), 298–312. <https://doi.org/10.1093/cercor/bht213>
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience* ,7 (7), 773–778. <https://doi.org/10.1038/nn1268>
- Gale, C. M., Eikeseth, S., & Klintwall, L. (2019). Children with Autism show Atypical Preference for Non-social Stimuli. *Scientific Reports* , 9 , 10355. <https://doi.org/10.1038/s41598-019-46705-8>
- Getz, L. M., & Toscano, J. C. (2021). Rethinking the McGurk effect as a perceptual illusion. *Attention, Perception, & Psychophysics* ,83 (6), 2583–2598. <https://doi.org/10.3758/s13414-021-02265-6>
- Green, A. M., & Angelaki, D. E. (2010). Multisensory integration: Resolving sensory ambiguities to build novel representations. *Current Opinion in Neurobiology* , 20 (3), 353–360. <https://doi.org/10.1016/j.conb.2010.04.009>
- 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. <https://doi.org/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. <https://doi.org/10.1111/j.1467-7687.2012.01171.x>
- Hirst, R. J., Stacey, J. E., Cragg, L., Stacey, P. C., & Allen, H. A. (2018). The threshold for the McGurk effect in audio-visual noise decreases with development. *Scientific Reports* , 8 (1), Article 1. <https://doi.org/10.1038/s41598-018-30798-8>
- Iarocci, G., Rombough, A., Yager, J., Weeks, D. J., & Chua, R. (2010). Visual influences on speech perception in children with autism. *Autism : The International Journal of Research and Practice* ,14 (4), 305–320. <https://doi.org/10.1177/1362361309353615>
- Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review* , 15 (3), 548–554.
- Irwin, J., & DiBlasi, L. (2017). Audiovisual speech perception: A new approach and implications for clinical populations. *Language and Linguistics Compass* , 11 (3), 77–91. psych.
- Irwin, J. R., Tornatore, L. A., Brancazio, L., & Whalen, D. H. (2011). Can children with autism spectrum disorders ‘hear’ a speaking face? *Child Development* , 82 (5), 1397–1403. psych. <https://doi.org/10.1111/j.1467-8624.2011.01619.x>
- Jain, A., Bansal, R., Kumar, A., & Singh, K. (2015). A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied and Basic Medical Research* ,5 (2), 124–127. <https://doi.org/10.4103/2229-516X.157168>
- Jertberg, R., Begeer, S., Geurts, H., Chakrabarti, B., & Van der Burg, E. (2023). *Perception of Temporal Synchrony Not a Prerequisite for Multisensory Integration* . <https://doi.org/10.21203/rs.3.rs-2890541/v1>

- Jertberg, R., Levitan, C. A., & Sherman, A. (2019). Multisensory processing of facial expressions in binocular rivalry. *Emotion (Washington, D.C.)* , 19 (7), 1214–1223. <https://doi.org/10.1037/emo0000522>
- 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. <https://doi.org/10.1016/j.rasd.2009.09.015>
- Kemp, B. J. (1973). Reaction time of young and elderly subjects in relation to perceptual deprivation and signal-on versus signal-off conditions. *Developmental Psychology* , 8 (2), 268–272. <https://doi.org/10.1037/h0034147>
- Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Oram Cardy, J. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders* , 9 , 202–222. <https://doi.org/10.1016/j.rasd.2014.10.008>
- Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child Development* , 85 (2), 685–694. <https://doi.org/10.1111/cdev.12142>
- MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology* , 21 (2), 131–141. <https://doi.org/10.3109/03005368709077786>
- Magnotti, J. F., & Beauchamp, M. S. (2018). Published estimates of group differences in multisensory integration are inflated. *PLOS ONE* , 13 (9), e0202908. <https://doi.org/10.1371/journal.pone.0202908>
- Magnotti, J. F., Smith, K. B., Salinas, M., Mays, J., Zhu, L. L., & Beauchamp, M. S. (2018). A causal inference explanation for enhancement of multisensory integration by co-articulation. *Scientific Reports* , 8 (1), 18032. <https://doi.org/10.1038/s41598-018-36772-8>
- Mamashli, F., Khan, S., Bharadwaj, H., Michmizos, K., Ganesan, S., Garel, K.-L. A., Hashmi, J. A., Herbert, M. R., Hamalainen, M., & Kenet, T. (2017). Auditory Processing in Noise Is Associated With Complex Patterns of Disrupted Functional Connectivity in Autism Spectrum Disorder. *Autism Research : Official Journal of the International Society for Autism Research* , 10 (4), 631–647. <https://doi.org/10.1002/aur.1714>
- Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature* , 264 (5588), 746–748. <https://doi.org/10.1038/264746a0>
- Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics* , 58 (3), 351–362. <https://doi.org/10.3758/bf03206811>
- Nava, E., Focker, J., & Gori, M. (2020). Children can optimally integrate multisensory information after a short action-like mini game training. *Developmental Science* , 23 (1), e12840. <https://doi.org/10.1111/desc.12840>
- Noel, J.P., De Nier, M. A., Stevenson, R., Alais, D., & Wallace, M. T. (2017). Atypical rapid audiovisual temporal recalibration in autism spectrum disorders. *Autism Research* , 10 (1), 121–129. <https://doi.org/10.1002/aur.1633>
- Noel, J.P., De Nier, M. A., Van der Burg, E., & Wallace, M. T. (2016). Audiovisual Simultaneity Judgment and Rapid Recalibration throughout the Lifespan. *PLOS ONE* , 11 (8), e0161698. <https://doi.org/10.1371/journal.pone.0161698>
- O'Brien, J., Mason, A., Chan, J., & Setti, A. (2023). Can We Train Multisensory Integration in Adults? A Systematic Review. *Multisensory Research* , 36 (2), 111–180. <https://doi.org/10.1163/22134808-bja10090>
- Parise, C. V., & Ernst, M. O. (2017). Noise, multisensory integration, and previous response in perceptual disambiguation. *PLOS Computational Biology* , 13 (7), e1005546. <https://doi.org/10.1371/journal.pcbi.1005546>

- Puschmann, S., & Thiel, C. M. (2017). Changed crossmodal functional connectivity in older adults with hearing loss. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* , 86 , 109–122. <https://doi.org/10.1016/j.cortex.2016.10.014>
- Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain and Development* , 25 (3), 166–172. [https://doi.org/10.1016/S0387-7604\(02\)00191-2](https://doi.org/10.1016/S0387-7604(02)00191-2)
- Rosemann, S., & Thiel, C. M. (2018). Audio-visual speech processing in age-related hearing loss: Stronger integration and increased frontal lobe recruitment. *NeuroImage* , 175 , 425–437. <https://doi.org/10.1016/j.neuroimage.2018.04.023>
- Ruiz Callejo, D., Wouters, J., & Boets, B. (n.d.). Speech-in-noise perception in autistic adolescents with and without early language delay. *Autism Research* , n/a (n/a). <https://doi.org/10.1002/aur.2966>
- Saalasti, S., Katsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., & Sams, M. (2012). Audio-visual speech perception and eye gaze behavior of adults with Asperger syndrome. *Journal of Autism and Developmental Disorders* , 42 (8), 1606–1615. psych. <https://doi.org/10.1007/s10803-011-1400-0>
- Samson, F., Mottron, L., Soulières, I., & Zeffiro, T. A. (2011). Enhanced visual functioning in autism: An ALE meta-analysis. *Human Brain Mapping* , 33 (7), 1553–1581. <https://doi.org/10.1002/hbm.21307>
- Sarko, D. K., Nidiffer, A. R., Powers III, A. R., Ghose, D., Hillock-Dunn, A., Fister, M. C., Krueger, J., & Wallace, M. T. (2012). Spatial and Temporal Features of Multisensory Processes: Bridging Animal and Human Studies. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* . CRC Press/Taylor & Francis. <http://www.ncbi.nlm.nih.gov/books/NBK92831/>
- Scheeren, A. M., Buil, J. M., Howlin, P., Bartels, M., & Begeer, S. (2022). Objective and subjective psychosocial outcomes in adults with autism spectrum disorder: A 6-year longitudinal study. *Autism: The International Journal of Research and Practice* , 26 (1), 243–255. <https://doi.org/10.1177/13623613211027673>
- Setti, A., Stapleton, J., Leahy, D., Walsh, C., Kenny, R. A., & Newell, F. N. (2014). Improving the efficiency of multisensory integration in older adults: Audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. *Neuropsychologia* , 61 , 259–268. <https://doi.org/10.1016/j.neuropsychologia.2014.06.027>
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature* , 408 (6814), 788–788. <https://doi.org/10.1038/35048669>
- Shipley, T. (1964). Auditory Flutter-Driving of Visual Flicker. *Science* , 145 (3638), 1328–1330. <https://doi.org/10.1126/science.145.3638.1328>
- Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport* , 12 (1), 7–10.
- Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry* , 48 (8), 813–821. psych. <https://doi.org/10.1111/j.1469-7610.2007.01766.x>
- Sperdin, H. F., & Schaer, M. (2016). Aberrant Development of Speech Processing in Young Children with Autism: New Insights from Neuroimaging Biomarkers. *Frontiers in Neuroscience* , 10 , 393. <https://doi.org/10.3389/fnins.2016.00393>
- Stevenson, R. A., Segers, M., Ncube, B. L., Black, K. R., Bebko, J. M., Ferber, S., & Barense, M. D. (2018). The cascading influence of multisensory processing on speech perception in autism. *Autism: The International Journal of Research and Practice* , 22 (5), 609–624. <https://doi.org/10.1177/1362361317704413>
- Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience* , 34 (3), 691–697. psych. <https://doi.org/10.1523/JNEUROSCI.3615-13.2014>

- Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research* , 227 (2), 249–261. <https://doi.org/10.1007/s00221-013-3507-3>
- 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. <https://doi.org/10.1037/a0027339>
- Stropahl, M., & Debener, S. (2017). Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. *NeuroImage. Clinical* , 16 , 514–523. <https://doi.org/10.1016/j.nicl.2017.09.001>
- Sumbly, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America* , 26 (2), 212–215. <https://doi.org/10.1121/1.1907309>
- 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. [psyh. https://doi.org/10.1007/s10803-010-1000-4](https://doi.org/10.1007/s10803-010-1000-4)
- Torenvliet, C., Groenman, A. P., Radhoe, T. A., Agelink van Rentergem, J. A., Van der Putten, W. J., & Geurts, H. M. (2023). A longitudinal study on cognitive aging in autism. *Psychiatry Research* , 321 , 115063. <https://doi.org/10.1016/j.psychres.2023.115063>
- Turi, M., Karaminis, T., Pellicano, E., & Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum. *Scientific Reports* , 6 , 21756. <https://doi.org/10.1038/srep21756>
- Vacas, J., Antoli, A., Sanchez-Raya, A., Perez-Duenas, C., & Cuadrado, F. (2021). Visual preference for social vs. Non-social images in young children with autism spectrum disorders. An eye tracking study. *PLoS ONE* , 16 (6), e0252795. <https://doi.org/10.1371/journal.pone.0252795>
- Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid Recalibration to Audiovisual Asynchrony. *Journal of Neuroscience* , 33 (37), 14633–14637. <https://doi.org/10.1523/JNEUROSCI.1182-13.2013>
- Van der Burg, E., Alais, D., & Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports* , 5 , 14526. <https://doi.org/10.1038/srep14526>
- Van der Burg, E., Alais, D., & Cass, J. (2018). Rapid recalibration to audiovisual asynchrony follows the physical—Not the perceived—Temporal order. *Attention, Perception, & Psychophysics* , 80 (8), 2060–2068. <https://doi.org/10.3758/s13414-018-1540-9>
- Van der Burg, E., & Goodbourn, P. T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech. *Proceedings of the Royal Society B: Biological Sciences* , 282 (1804), 20143083–20143083. <https://doi.org/10.1098/rspb.2014.3083>
- Van der Burg, E., L, N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance* , 34 (5), 1053–1065. <https://doi.org/10.1037/0096-1523.34.5.1053>
- Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage* , 55 (3), 1208–1218. <https://doi.org/10.1016/j.neuroimage.2010.12.068>
- van Ee, R., van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory Congruency as a Mechanism for Attentional Control over Perceptual Selection. *The Journal of Neuroscience* , 29 (37), 11641–11649. <https://doi.org/10.1523/JNEUROSCI.0873-09.2009>
- Van Wassenhove, V., Grant, K., & Poeppel, D. (2007). Temporal window of integration in auditory–visual speech perception. *Neuropsychologia* , 45 , 598–607. <https://doi.org/10.1016/j.neuropsychologia.2006.01.001>
- Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain Research* , 1111 (1), 134–142. <https://doi.org/10.1016/j.brainres.2006.05.078>

- Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance* ,26 (5), 1583–1590.
- Vroomen, J., & de Gelder, B. (2004). Temporal Ventriloquism: Sound Modulates the Flash-Lag Effect. *Journal of Experimental Psychology: Human Perception and Performance* , 30 (3), 513–518. <https://doi.org/10.1037/0096-1523.30.3.513>
- Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Research. Cognitive Brain Research* ,22 (1), 32–35. <https://doi.org/10.1016/j.cogbrainres.2004.07.003>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2014.08.005>
- Walling, A. D., & Dickson, G. M. (2012). Hearing Loss in Older Adults. *American Family Physician* , 85 (12), 1150–1156.
- Weiland, R. F., Polderman, T. J., Smit, D. J., Begeer, S., & Van der Burg, E. (2022). No differences between adults with and without autism in audiovisual synchrony perception. *Autism: The International Journal of Research and Practice* , 13623613221121414. <https://doi.org/10.1177/13623613221121414>
- Woodhouse, L., Hickson, L., & Dodd, B. (2008). Review of visual speech perception by hearing and hearing-impaired people: Clinical implications. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists* , 44 , 253–270. <https://doi.org/10.1080/13682820802090281>
- Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. *Perception & Psychophysics* ,67 (3), 531–544. <https://doi.org/10.3758/BF03193329>
- Zhang, J., Meng, Y., He, J., Xiang, Y., Wu, C., Wang, S., & Yuan, Z. (2019). McGurk effect by individuals with autism spectrum disorder and typically developing controls: A systematic review and meta-analysis. *Journal of Autism and Developmental Disorders* , 49 (1), 34–43. <https://doi.org/10.1007/s10803-018-3680-0>
- Zhou, H., Cai, X., Weigl, M., Bang, P., Cheung, E. F. C., & Chan, R. C. K. (2018). Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews* , 86 , 66–76. <https://doi.org/10.1016/j.neubiorev.2017.12.013>
- Zhou, H., Lai, I. Y. S., Hung, K. S. Y., Chan, M. K. M., Ho, Z. T. Y., Lam, J. P. H., Lui, S. S. Y., & Chan, R. C. K. (2022). Audiovisual temporal processing in adult patients with first-episode schizophrenia and high-functioning autism. *Schizophrenia* , 8 (1), Article 1. <https://doi.org/10.1038/s41537-022-00284-2>