Age, not autism, influences multisensory integration of speech stimuli among adults in a McGurk/MacDonald paradigm

Article

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/ejn.16319

Publisher: Wiley

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH REPORT

Age, not autism, influences multisensory integration of speech stimuli among adults in a McGurk/MacDonald paradigm

Robert M. Jertberg[1] | Sander Begeer[1] | Hilde M. Geurts[2,3] |
Bhismadev Chakrabarti[4,5,6] | Erik Van der Burg[2]

[1]Department of Clinical and Developmental Psychology, Vrije Universiteit Amsterdam, The Netherlands and Amsterdam Public Health Research Institute, Amsterdam, Netherlands

[2]Dutch Autism and ADHD Research Center (d'Arc), Brain & Cognition, Department of Psychology, Universiteit van Amsterdam, Amsterdam, The Netherlands

[3]Leo Kannerhuis (Youz/Parnassiagroup), Den Haag, The Netherlands

[4]Centre for Autism, School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

[5]India Autism Center, Kolkata, India

[6]Department of Psychology, Ashoka University, Sonipat, India

Correspondence
Robert Jertberg, Department of Clinical and Developmental Psychology, Vrije Universiteit Amsterdam, The Netherlands and Amsterdam Public Health Research Institute, Amsterdam, Netherlands.
Email: jertbergiii@gmail.com

Funding information
ZonMw, Grant/Award Number: 60-63600-98-834

Edited by: John Foxe

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 behaviour. 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 496 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.

KEYWORDS
ageing, autism, McGurk effect, rapid temporal recalibration, window of perceived synchrony

1 | INTRODUCTION

The brain is immersed in a rich world 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). The most important and widely studied avenue for MSI is temporal proximity, upon which all multisensory interactions depend to some degree (Chen & Vroomen, 2013; Costantini et al., 2016; Meredith et al., 1987; Munhall et al., 1996; Shams et al., 2000). Although there are many different ways in which MSI can enhance perception (Green & Angelaki, 2010; Parise & Ernst, 2017; Sumby & Pollack, 1954; Van der Burg et al., 2008; van Ee et al., 2009; Vroomen & de Gelder, 2000), speech perception may be the most significant to daily interactions. The integration of visual signals with their auditory counterparts can greatly enhance our understanding of speech (Erber, 1969; Irwin & DiBlasi, 2017; Sumby & Pollack, 1954; Woodhouse et al., 2008) and even produce 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 environments, in which the reliability of relevant auditory signals is compromised by competing inputs, that the most benefit is generally afforded by the integration of visual information (Erber, 1969; MacLeod & Summerfield, 1987; Sumby & Pollack, 1954). It is notable, then, that it is precisely under these circumstances that autistic individuals struggle most with speech perception (Ruiz Callejo et al., 2023; Alcántara et al., 2004; Fadeev et al., 2023; Mamashli et al., 2017). 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 each area (Feldman et al., 2018; Kwok et al., 2015; Rapin & Dunn, 2003; Sperdin & Schaer, 2016; van Laarhoven et al., 2019; 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[1] demonstrate impairments in speech processing (Kwok et al., 2015; Rapin & Dunn, 2003; Sperdin & Schaer, 2016) as well as attenuated multisensory effects, particularly when young

---

[1]We alternate between person-first and identity-first language to address the varying preferences of the international autism community, as recommended in Buijsman et al. (2023).

(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 in the environment, 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 can be observed 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, Siemann, et al., 2014). However, recent research challenges the degree to which this applies to autistic adults (Ainsworth & Bertone, 2023; 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 report 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 (Noel et al., 2017; Turi et al., 2016), although the one with the largest adult sample did not (Weiland et al., 2022), again raising questions about the existence of temporal processing differences in adults.

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; Donohue et al., 2012; Stevenson et al., 2018; Stevenson, Segers, et al., 2014). 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 behavioural 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 non-autistic people continue to develop in their ability to integrate audiovisual speech stimuli, whereas 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 (Foxe et al., 2015). In fact, it

has been found with several paradigms that the differences in MSI seen between autistic and non-autistic children are no longer apparent by adulthood (Beker et al., 2018; Crosse et al., 2022). 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 and 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 of 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.

2 | METHODS

2.1 | 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 participants 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.

A total of 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 because their performance on the task matched one or multiple exclusion criteria, as pre-registered in As Predicted (#102341, see https://aspredicted.org/g93a3.pdf). For precise information on the number of participants from each group excluded (and the reasons for their exclusion), refer to Figure S1. 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 the entirety of that in our earlier study (Jertberg et al., 2023), in addition to later recruits.

2.2 | Apparatus and stimuli

The experiment was programmed and conducted online using the Neurotask platform (www.neurotask.com). Participants completed the task using their own hardware and the Google Chrome browser. The stimuli were taken from Hillock-Dunn et al. (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). 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.

2.3 | Procedure

Trials began with a black fixation cross in the centre 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
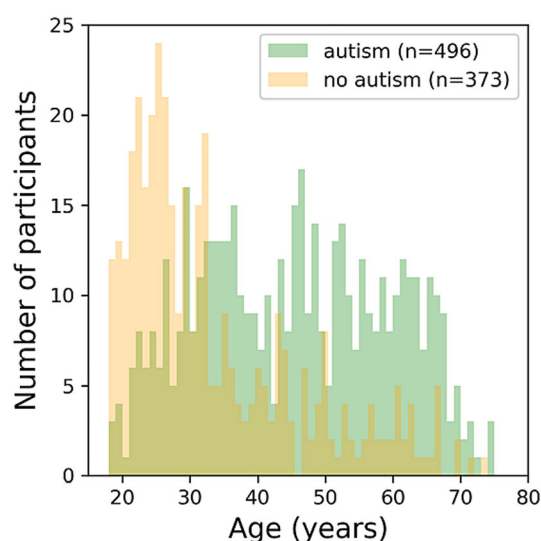

FIGURE 1 Age distribution of autistic and non-autistic participants. Darker sections reflect overlap between groups.

TABLE 1 Demographic breakdown by group.

	Autism (n = 496)		No autism (n = 373)		Statistics
Sex					
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		

Abbreviations: AQ, Autism Quotient; ICAR, International Cognitive Ability Resource (abbreviated intelligence quotient test); SD, standard deviation.
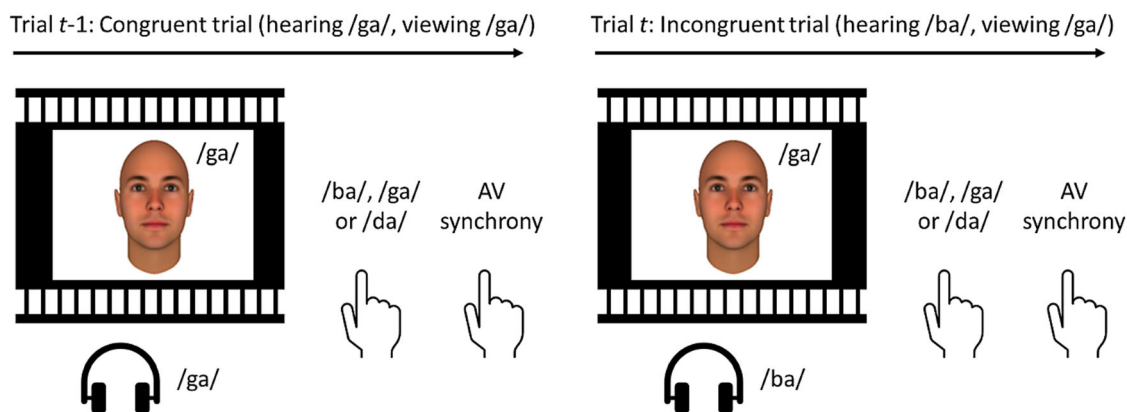
WILEY

Trial *t*-1: Congruent trial (hearing /ga/, viewing /ga/)

/ga/

/ga/

/ba/, /ga/ or /da/

AV synchrony

/ga/

Trial *t*: Incongruent trial (hearing /ba/, viewing /ga/)

/ga/

/ga/

/ba/, /ga/ or /da/

AV synchrony

/ba/

FIGURE 2 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.

FIGURE 3 (a) Proportion of /da/ responses per group as a function of stimulus onset asynchrony (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.

(a)

autism; congr
autism; incongr
no autism; congr
no autism; incongr

(b)

autism
no autism
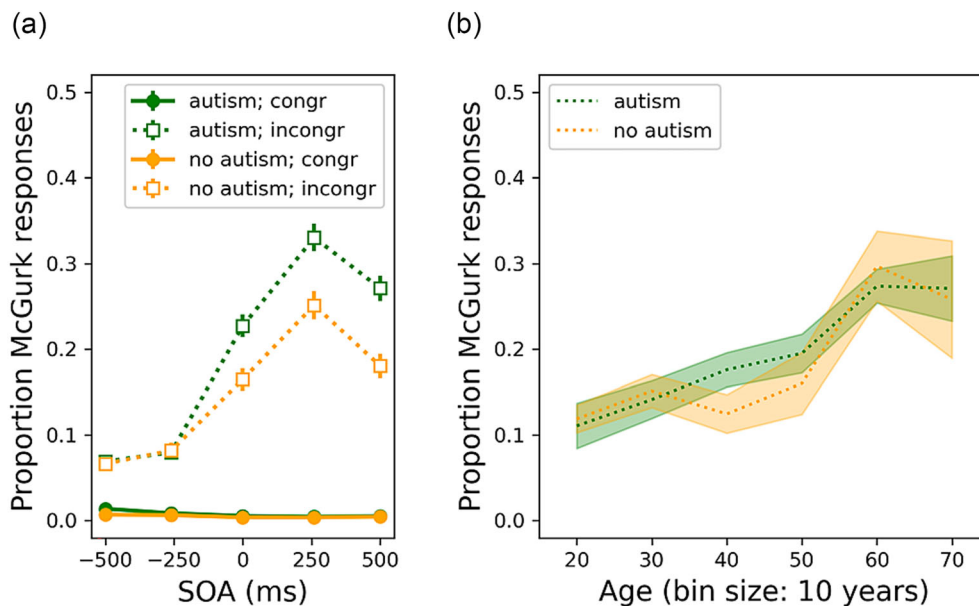
were asked whether they heard /ba/, /da/, or /ga/ by pressing either the b, d, or g key, respectively. Second, they were asked whether the video and audio were synchronized or not by pressing the 1 or 0 key, respectively. The following trial was initiated as soon as the synchrony judgement was given. The two congruency types and five SOAs produced a total of 10 unique trial types. After reading the written instructions and completing 10 practice trials (one of each stimulus type), participants completed 10 repetitions of each trial type, for a total of 100 experimental trials in randomized order. The experiment lasted approximately 7 min and was part of a larger battery of online tasks (not reported here). Figure 2 illustrates an example trial sequence.

3 | RESULTS

3.1 | Syllable responses

Figure 3a 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 3b 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).

We conducted a repeated measures analysis of variance (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 3a, the rates of /da/ responses varied significantly as a function of SOA (F[1867] = 243.632, p < .001) and congruency (F[1867] = 529.059, p < .001). What is more, there was a significant SOA × congruency interaction (F[1867] = 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[1867] = 11.448, p < .001). There were also interactions between group × congruency (F[867] = 9.344, p = .002), group × SOA (F[867] = 12.012, p < .001) and group × congruency × SOA (F[867] = 14.662, p < .001). However, because of 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[1866] = 54.194, p < .001), wherein older participants from both groups tended to experience the illusion more frequently than younger ones (see Figure 3b). Age also interacted

with SOA (F[1866] = 37.815, p < .001), congruency (F[1866] = 50.695, p < .001), and SOA × congruency (F[1866] = 38.530, p < .001), augmenting all of their effects.

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 (approximately 2% of trials). For transparency, information regarding the rates of /ba/ and /ga/ responses can be found in Tables S3–S6 and Figure S2.

3.2 | Simultaneity judgments

Figure 4a plots the mean proportion of simultaneity judgments as a function of SOA and congruency for both groups. Figure 4b 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.

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[1866] = 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 4a). Additionally, the proportion of synchrony responses was much higher when the stimuli were congruent than incongruent (F[1866] = 377.547, p < .001). Congruency also interacted with SOA, such that its effect

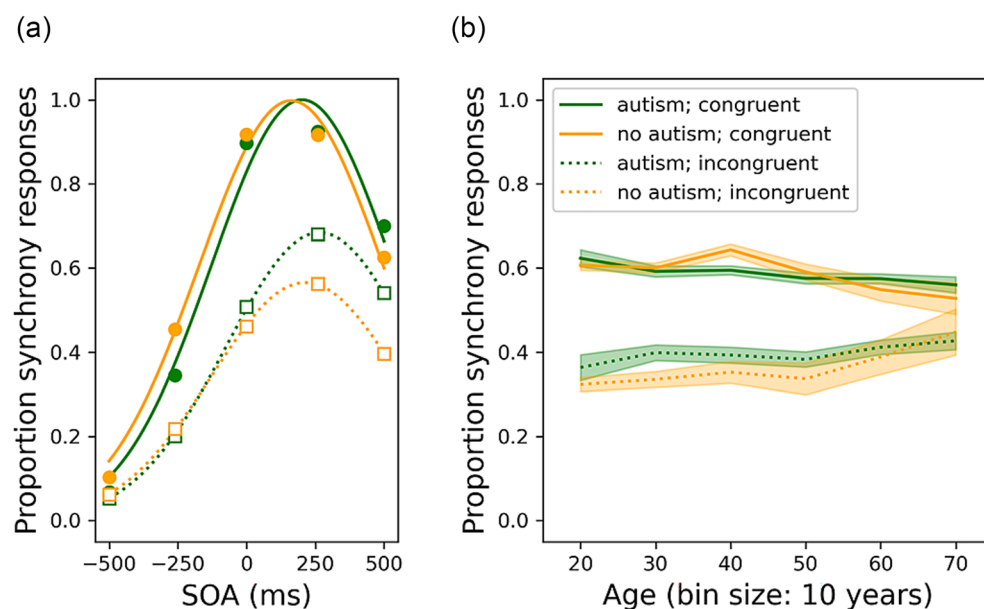FIGURE 4 (a) Proportion of synchrony judgments per group as a function of stimulus onset asynchrony (SOA) for congruent and incongruent trials. Here, negative SOAs indicate that the voice was leading the lip movements, and vice versa. (b) Proportion of synchrony judgments (collapsed across SOAs) as a function of age and congruency for each group (bins of 10 years). The error bars reflect the standard error of the mean.

was most pronounced when stimuli occurred simultaneously or with a slight visual lead ($F[1866] = 48.803$, $p < .001$), and with group ($F[1866] = 17.004$, $p < .001$). A follow-up t-test comparing the difference between mean simultaneity judgement 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 4a). SOA also interacted with group, with the differences between autistic and non-autistic participants emerging at the mid-range SOAs ($F[1866] = 7.204$, $p < .001$), which is logical given that the longer SOAs were much more obvious to both groups. Both congruency ($F[1866] = 31.101$, $p < .001$) and SOA ($F[1866] = 55.233$, $p < .001$) also significantly interacted with age. Finally, we detected a significant three-way interaction between congruency × SOA × age ($F[1866] = 6.792$, $p < .001$). 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 judgement tasks, these effects were not explored further.

3.3 | 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 5 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
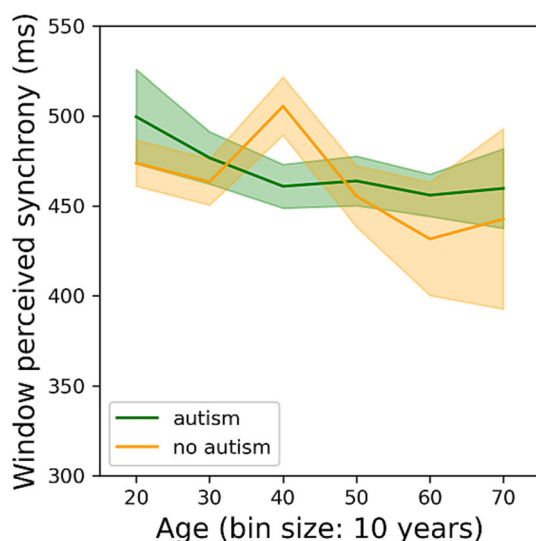

FIGURE 5 Mean window of perceived synchrony for autistic and non-autistic participants relative to age (in bins of 10 years). The error bars reflect the standard error of the mean.

fitting procedure was not successful, resulting in exclusion from further analyses.

We conducted an analysis of covariance (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 ($F[1865] = 0.227$, $p = .634$) or age ($F[1865]$, $p = .053$) on the WPS.

3.4 | 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 6a) 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 because of fitting issues.

Accordingly, Figure 6a reflects the mean proportion of synchrony responses as a function of SOA for each previous modality order and group (collapsed across congruency conditions). Figure 6b reflects the mean PSS derived from these synchrony distributions according to group and previous modality order. Figure 6c shows the Δ PSS (i.e., rapid temporal recalibration) as a function of age (in bins of 10 years) for each group.

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 ($F[1865] = 13.823$, $p < .001$), such that the PSS was smaller when audition led (206 ms) in the previous trial than when vision led (231 ms), as Figure 6b illustrates. Rapid temporal recalibration did not differ between groups, as the modality order × group interaction failed to reach significance ($F[1865] = 1.968$, $p = .161$). Autistic participants showed a larger average PSS (240 ms) than non-autistic participants (190 ms) overall ($F[1865] = 13.866$, $p < .001$), 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 the previous modality order ($F[1865] = 2.624$, $p = .106$; see Figure 6c). However, older participants did have a higher overall PSS ($F[1865] = 51.770$, $p < .001$). The mean PSS for participants above the median age was 249 ms, compared with a mean of 188 ms for those below the median age.
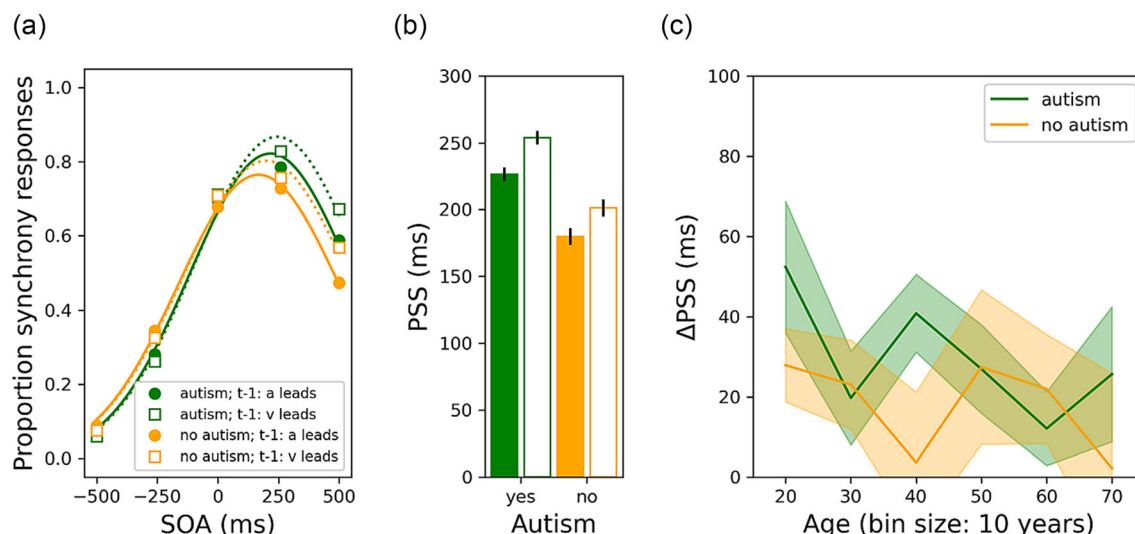
FIGURE 6 (a) Proportion of synchrony judgments per group as a function of stimulus onset asynchrony (SOA), relative to the modality order of the preceding trial (collapsed across congruency conditions). Here, negative SOAs indicate that the voice was leading the lip movements, and vice versa. (b) Point 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.

3.5 | Age matched /da/ response analysis

Because of the significant effect of age on /da/ response rates and the skewness of the age distributions (see Figure 1), we conducted an exploratory follow-up analysis in which 10 participants were selected at random from each age bin depicted in Figure 3b for both groups. This allowed us to conduct another repeated measures ANOVA on the mean proportion of /da/ responses including SOA and congruency as within-subjects variables, group as a between-subjects variable, and age as a covariate, this time with a sub-group matched on age. The mean ages were 45.7 years for autistic participants and 45.5 years for non-autistic participants ($t[1118] = 0.066$, $p = .948$), and each group comprised 60 participants. The results of this analysis confirmed all of those discussed in the main results with the full sample (see Table S1).

3.6 | Bayesian analysis

To evaluate the evidence that group did *not* drive differences in susceptibility to the illusion, we conducted an ANCOVA focusing on the rate of /da/ responses in incongruent trials (collapsed across SOAs) with group as a between-subjects factor and age as a covariate. Here, we found that age ($F[1866] = 53.104$, $p < .001$) but not group ($F[1866] = 0.034$, $p = .854$) had a significant influence on /da/ response rates. We then repeated this ANCOVA using JASP's Bayesian statistics module. We

used uniform model priors (assuming equal likelihood of the alternative models including age, age + group, group and the null hypothesis) and default priors on coefficients (r scale prior width $= 0.5$ for fixed effects and 0.354 for covariates). We found that the best model was provided by age, compared with which the BF01 of the model including age and group was 11.928, the BF01 of the model including group alone was 1.117×10^{11} and the BF01 of the null model was 1.450×10^{12}. The full results of the Bayesian analysis can be seen in Table S2.

3.7 | Gender analyses

Because of the imbalance in gender between groups, we conducted exploratory follow-up analyses with gender included as a variable. Although gender did interact significantly with some factors in our /da/ response and simultaneity judgement analyses (and women had a higher mean PSS than men), it did not change the significance of any of the aforementioned main effects or interactions. For transparency, the results of these analyses can be found in Tables S7–S9 and Figures S3–S5. Gender had no significant influence on the WPS analysis.

4 | DISCUSSION

Based on studies primarily with children and adolescents, it has been hypothesized that autistic individuals show

attenuated MSI, particularly for speech stimuli. Our results provide compelling evidence that some differences found in children may not persist into adulthood. We found no significant difference between autistic and non-autistic individuals in susceptibility to the McGurk/MacDonald illusion once we accounted for age differences in our sample. Because this ran contrary to 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 the Supporting Information). 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 et al., 2010). Our findings with adults are consistent with the trajectory of improvement these results imply.

Instead of a difference between groups, 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. Although an increase in the rate of the McGurk/MacDonald effect between younger and older adults has been detected in non-autistic participants (Mcgurk & Macdonald, 1976; Sekiyama et al., 2014; 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 although 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 ageing 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 a reduced reliability of the auditory signal resulting from the progressive hearing loss common in ageing, 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 possibility 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 (although Foxe et al., 2015, found little influence of looking behaviour). Accordingly, studies that do not control for visual attention may overstate differences in MSI. A merit of our design is that although we do not directly measure eye movements, our simultaneity judgement 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. Although 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 a simultaneity judgement 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 spoken to 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 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; Vroomen & Keetels, 2010). 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.

First, 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 sample. This was even true at 0 ms, when participants dropped from recognizing the physical simultaneity of the stimuli on 91.8% to 46.2% of trials in the non-autistic group and on 89.7% to 50.8% of trials in the autistic group. This suggests a profound interference of phonetic incongruence on basic temporal processing. Van Wassenhove et al. (2007) attributed a similar finding 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; E. G. 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. An alternative explanation is that autistic participants may be less subject to a cognitive bias to judge the incongruent stimuli asynchronous because they seem unnatural, as described in Vroomen and Keetels (2010), and thereby better able to perform the task. However, as is discussed in Jertberg et al. (2023), the manner in which the effect scales with SOA (and occurs regardless of whether participants experience the illusion) suggests that a cognitive bias may not be the full explanation. As such, further research into the lip reading abilities of autistic adults and their potential influence on the 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 audiovisual WPS), most had small samples (ranging from 32 to 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 (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. It is also worth noting that performance on the SJ task was quite poor overall, possibly due to the difficulty of recognizing the timing of a velar consonant like /ga/. This may have adversely affected the sensitivity of our WPS analysis to subtle temporal differences. 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. Research investigating responses to simple tones and disks suggests similarly protracted reaction times for auditory and visual stimuli among autistic individuals, as well as attenuated multisensory benefits and abnormal electrophysiological responses as early as 100 ms after stimulus presentation (Brandwein et al., 2013, 2015). Research into visual motion recognition, on the other hand, shows faster reaction times for autistic participants (Foss-Feig et al., 2013). It is possible that the processing speed for speech stimuli also differs,

as they are often found to have unique temporal processing profiles (Stevenson & Wallace, 2013). Although research here is more limited, there is some evidence of faster recognition of voices among autistic individuals (Lin et al., 2016). To fully evaluate the possibility of a processing speed explanation for the PSS differences observed in this study, more research should be done to compare auditory and visual reaction times for speech stimuli among autistic and non-autistic individuals.

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 preference for greater visual lead 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).

Although 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 large online experimental battery, which limited our control over the hardware/settings participants used during the experiment and placed constraints on the number of trials they could complete. Regarding the former, we were unable to control the volume/size and quality of the audio and video input to the extent that would have been possible in the laboratory. Although findings are mixed with regard to whether the McGurk/MacDonald effect differs in laboratory versus online experiments (Getz & Toscano, 2021; Magnotti et al., 2018), it is possible that differences in the choice of hardware and settings could have influenced the reliability of sensory information and therefore the strength of the illusion. Although we have no reason to believe that meaningful systematic differences in these choices existed between groups, it is possible that older participants may have been more likely to lack headphones or have older devices. Regarding the latter, 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 the 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 with non-autistic ones (Foxe et al., 2015; Iarocci et al., 2010; J. R. Irwin et al., 2011; E. G. Smith & Bennetto, 2007; Taylor et al., 2010). Although we did not find a difference in audiovisual speech processing between groups, we are unable to speak to the influence of unisensory factors in our findings because of the lack of audio and video-only trials. 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.

Additionally, although the McGurk/MacDonald paradigm is the most widely used tool for studying audiovisual speech integration, it is not without its shortcomings. High variability in the rate of the illusion and its somewhat contrived nature have raised concerns about its stability and ecological validity (Alsius et al., 2018; Getz & Toscano, 2021; Van Engen et al., 2022). Speech in noise experiments, in which the addition of visual information facilitates speech perception, rather than distorting it, offers a more naturalistic alternative. However, speech in noise performance does not always correlate with McGurk/MacDonald susceptibility (Stevenson et al., 2018; Van Engen et al., 2017). Accordingly, our findings cannot serve as conclusive evidence on their own that autistic adults do not face difficulties with audiovisual speech perception in their daily lives, particularly given that time constraints limited our stimulus set to only two syllable pairings. Still, they do capture a form of audiovisual speech integration that has been shown to correlate positively with communication skills among autistic and non-autistic individuals (Feldman et al., 2022), and they resonate with the speech in noise results of Foxe et al. (2015). As such, they remain encouraging findings with regard to the development of MSI of speech stimuli in autism. Future research should continue to evaluate the degree to which laboratory paradigms predict real-world outcomes for autistic individuals.

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 suggests 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 only found among the individuals with lower IQs (who are underrepresented in our sample), it would be unclear whether they were due to autism or intellectual impairment. Our sample is also notable in that women are overrepresented in our autism group, likely due to the fact that they tend to be more likely to participate in online surveys (Becker, 2022), which make up the majority of NAR projects. However, as seen in the Supporting Information, although gender did interact with certain factors in our analyses, it did not alter the significance of any other main effects or interactions. So this peculiarity of our sample is unlikely to limit the generalizability of our findings.

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 are highly encouraging given the essential role MSI has in speech perception and compensation for the unisensory deterioration that is inevitable with ageing. 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 leads to a limited understanding of the broader trajectory of this developmental condition, which can only be extended by giving autistic adults the attention they deserve.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

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

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ejn.16319.

DATA AVAILABILITY STATEMENT

Our raw data include sensitive personal information and identifiers not suitable for open access. However, upon request to the Netherlands Autism Register (info@nederlandsautismeregister.nl) 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.

ORCID

Erik Van der Burg https://orcid.org/0000-0003-2522-7925

REFERENCES

Ainsworth, K., & Bertone, A. (2023). Audiovisual temporal binding window narrows with age in autistic individuals. *Autism Research*, 16(2), 355–363. https://doi.org/10.1002/aur.2860

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

Alsius, A., Paré, M., & Munhall, K. G. (2018). Forty years after hearing lips and seeing voices: The McGurk effect revisited.

Multisensory Research, *31*(1–2), 111–144. https://doi.org/10.1163/22134808-00002565

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

Becker, R. (2022). Gender and Survey Participation: An Event History Analysis of the Gender Effects of Survey Participation in a Probability-based Multi-wave Panel Study with a Sequential Mixed-mode Design. *Methods, Data, Analyses*, *16*(1), Article 1. https://doi.org/10.12758/mda.2021.08

Beker, S., Foxe, J. J., & Molholm, S. (2018). Ripe for solution: Delayed development of multisensory processing in autism and its remediation. *Neuroscience and Biobehavioral Reviews*, *84*, 182–192. https://doi.org/10.1016/j.neubiorev.2017.11.008

Brandwein, A. B., Foxe, J. J., Butler, J. S., Frey, H. P., Bates, J. C., Shulman, L., & Molholm, S. (2015). Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *Journal of Autism and Developmental Disorders*, *45*(1), 230–244. https://doi.org/10.1007/s10803-014-2212-9

Brandwein, A. B., Foxe, J. J., Butler, J. S., Russo, N. N., Altschuler, T. S., Gomes, H., & Molholm, S. (2013). The development of multisensory integration in high-functioning autism: High-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cerebral Cortex*, *23*(6), 1329–1341. https://doi.org/10.1093/cercor/bhs109

Buijsman, R., Begeer, S., & Scheeren, A. M. (2023). 'Autistic person' or 'person with autism'? Person-first language preference in Dutch adults with autism and parents. *Autism: The International Journal of Research and Practice*, *27*(3), 788–795. https://doi.org/10.1177/13623613221117914

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790–811. https://doi.org/10.3758/s13414-013-0475-4

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

Crosse, M. J., Foxe, J. J., Tarrit, K., Freedman, E. G., & Molholm, S. (2022). Resolution of impaired multisensory processing in autism and the cost of switching sensory modality. *Communications Biology*, *5*(1), 601. https://doi.org/10.1038/s42003-022-03519-1

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

Donohue, S., Darling, E., & Mitroff, S. (2012). Links between multisensory processing and autism. *Experimental Brain Research*, *222*(4), 377–387. https://doi.org/10.1007/s00221-012-3223-4

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. https://doi.org/10.1044/jshr.1202.423

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., Conrad, J. G., Kuang, W., Tu, A., Liu, Y., Simon, D. M., Wallace, M. T., & Woynaroski, T. G. (2022). Relations between the McGurk effect, social and communication skill, and autistic features in children with and without autism. *Journal of Autism and Developmental Disorders*, *52*(5), 1920–1928. https://doi.org/10.1007/s10803-021-05074-w

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. psyh. 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*, 381–389. 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., Grantham, D. W., & Wallace, M. T. (2016). The temporal binding window for audiovisual speech: Children are like little adults. *Neuropsychologia*, 88, 74–82. https://doi.org/10.1016/j.neuropsychologia.2016.02.017

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), 12372. 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

Irwin, J., & DiBlasi, L. (2017). Audiovisual speech perception: A new approach and implications for clinical populations. *Lang & Ling Compass*, 11(3), 77–91. psyh. https://doi.org/10.1111/lnc3.12237

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. psyh. 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 & 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

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

Lin, I.-F., Agus, T. R., Suied, C., Pressnitzer, D., Yamada, T., Komine, Y., Kato, N., & Kashino, M. (2016). Fast response to human voices in autism. *Scientific Reports*, 6(1), 26336. https://doi.org/10.1038/srep26336

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E.-J. (2019). JASP: Graphical Statistical Software for Common Statistical Designs. *Journal of Statistical Software*, 88(2), 1–17. https://doi.org/10.18637/jss.v088.i02

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., Hämäläinen, 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

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*, 7(10), 3215–3229.

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., Föcker, 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., De Niear, M. A., Stevenson, R., Alais, D., & Wallace, M. T. (2017). Atypical rapid audio-visual temporal recalibration in autism spectrum disorders. *Autism Research*, 10(1), 121–129. psyh. https://doi.org/10.1002/aur.1633

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

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. (2023). Speech-in-noise perception in autistic adolescents with and without early language delay. *Autism Research*, *16*(9), 1719–1727. https://doi.org/10.1002/aur.2966

Saalasti, S., Kätsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., & Sams, M. (2012). Audiovisual speech perception and eye gaze behavior of adults with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *42*(8), 1606–1615. psyh. https://doi.org/10.1007/s10803-011-1400-0

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

Sekiyama, K., Soshi, T., & Sakamoto, S. (2014). Enhanced audiovisual integration with aging in speech perception: A heightened McGurk effect in older adults. *Frontiers in Psychology*, *5*, 323. https://doi.org/10.3389/fpsyg.2014.00323

Setti, A., Burke, K., Kenny, R. A., & Newell, F. (2013). Susceptibility to a multisensory speech illusion in older persons is driven by perceptual processes. *Frontiers in Psychology*, *4*, 575. https://doi.org/10.3389/fpsyg.2013.00575

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

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, *12*(1), 7–10. https://doi.org/10.1097/00001756-200101220-00009

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry*, *48*(8), 813–821. psyh. 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., Ferber, S., Barense, M. D., & Wallace, M. T. (2014). The impact of multisensory integration deficits on speech perception in children with autism spectrum disorders. *Frontiers in Psychology*, *5*, 379. https://doi.org/10.3389/fpsyg.2014.00379

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. psyh. 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

Sumby, 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

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., Antolí, A., Sánchez-Raya, A., Pérez-Dueñas, 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. https://doi.org/10.1098/rspb.2014.3083

Van der Burg, E., Olivers, C. N. L., 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 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 Engen, K. J., Dey, A., Sommers, M. S., & Peelle, J. E. (2022). Audiovisual speech perception: Moving beyond McGurk. *The Journal of the Acoustical Society of America*, *152*(6), 3216–3225. https://doi.org/10.1121/10.0015262

Van Engen, K. J., Xie, Z., & Chandrasekaran, B. (2017). Audiovisual sentence recognition not predicted by susceptibility to the McGurk effect. *Attention, Perception & Psychophysics*, *79*(2), 396–403. https://doi.org/10.3758/s13414-016-1238-9

van Laarhoven, T., Stekelenburg, J. J., & Vroomen, J. (2019). Increased sub-clinical levels of autistic traits are associated with reduced multisensory integration of audiovisual speech. *Scientific Reports*, *9*(1), 9535. https://doi.org/10.1038/s41598-019-46084-0

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

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., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871–884. https://doi.org/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. *Brain Research. Cognitive Brain Research*, *22*(1), 32–35. https://doi.org/10.1016/j.cogbrainres.2004.07.003

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*, *27*(4), 927–937. 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. psyh. 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), 75. https://doi.org/10.1038/s41537-022-00284-2

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.