RESEARCH ARTICLE

Differences in audiovisual temporal processing in autistic adults are specific to simultaneity judgments

Paula Regener[1,2] | Naomi Heffer[3,4] | Scott A. Love[5] | Karin Petrini[4,6] | Frank Pollick[2]

[1]Norwich Medical School, University of East Anglia, Norwich, UK

[2]School of Psychology and Neuroscience, University of Glasgow, Glasgow, UK

[3]School of Sciences, Bath Spa University, Bath, UK

[4]Department of Psychology, University of Bath, Bath, UK

[5]INRAE, CNRS, Université de Tours, PRC, Nouzilly, France

[6]The Centre for the Analysis of Motion, Entertainment Research and Applications (CAMERA), Bath, UK

Correspondence
Naomi Heffer, School of Sciences, Bath Spa University, Bath, UK.
Email: n.heffer@bathspa.ac.uk

Abstract

Research has shown that children on the autism spectrum and adults with high levels of autistic traits are less sensitive to audiovisual asynchrony compared to their neurotypical peers. However, this evidence has been limited to simultaneity judgments (SJ) which require participants to consider the timing of two cues together. Given evidence of partly divergent perceptual and neural mechanisms involved in making temporal order judgments (TOJ) and SJ, and given that SJ require a more global type of processing which may be impaired in autistic individuals, here we ask whether the observed differences in audiovisual temporal processing are task and stimulus specific. We examined the ability to detect audiovisual asynchrony in a group of 26 autistic adult males and a group of age and IQ-matched neurotypical males. Participants were presented with beep-flash, point-light drumming, and face-voice displays with varying degrees of asynchrony and asked to make SJ and TOJ.

The results indicated that autistic participants were less able to detect audiovisual asynchrony compared to the control group, but this effect was specific to SJ and more complex social stimuli (e.g., face-voice) with stronger semantic correspondence between the cues, requiring a more global type of processing. This indicates that audiovisual temporal processing is not generally different in autistic individuals and that a similar level of performance could be achieved by using a more local type of processing, thus informing multisensory integration theory as well as multisensory training aimed to aid perceptual abilities in this population.

Lay Summary

Detecting whether two events (e.g., the sound of a beep and a flash of light) occur at the same time is an important process for the brain when combining information from the different senses to create a perceptual whole. Previous research has suggested that there are differences between autistic and neurotypical people in

the way that they process the relative timing of auditory and visual information. Our research suggests that in autistic adults differences in temporal processing might be limited to specific tasks and types of social information that require a focus more on the overall perception, rather than focusing on individual sensory cues.

INTRODUCTION

Autism Spectrum Disorders (ASD) are a set of neurodevelopmental conditions characterized by difficulties with social communication and interaction, as well as repetitive patterns of behavior, interests and activities (APA, 2013). Prevalence estimates suggest that 1 in 36 children in the USA (CDC, 2023) and approximately 1% of the UK population (NHS England, 2020) are on the autism spectrum, revealing the pressing need to better understand ASD. Clinical research has repeatedly described differences in sensory processing between autistic individuals and neurotypical individuals (Lane et al., 2010; Robertson & Simmons, 2013; Szelag et al., 2004). Consequently, sensory processing differences have been adopted as diagnostic criteria for ASD in the DSM-V (APA, 2013).

As well as there being established differences with unisensory processing in ASD, there is accumulating evidence that autistic individuals may also differ in terms of multisensory processing. For example, there is evidence that autistic people perceive audiovisual illusions such as the McGurk effect (Mcgurk & Macdonald, 1976) less than neurotypical controls (Gelder et al., 1991; Irwin et al., 2011; Mongillo et al., 2008), benefit less from information provided by an additional sensory modality (Feldman et al., 2018; Smith & Bennetto, 2007) and show less effective neural integration during audiovisual tasks (Brandwein et al., 2013, 2015).

Studies employing a variety of age groups, stimuli and analysis techniques have found that autistic children and adolescents are less sensitive to audiovisual asynchrony than neurotypical controls (i.e., they perceive auditory and visual cues as synchronous for larger temporal lags; Bebko et al., 2006; de Boer-Schellekens et al., 2013; Foss-Feig et al., 2010; Grossman et al., 2009; Kwakye et al., 2011; Stevenson, Siemann, et al., 2014). More recent studies have also shown that adults with high levels of autistic traits show lower abilities to detect audiovisual asynchrony than individuals with lower levels of autistic traits in terms of communication, speech, and attention switching processes (Van Laarhoven et al., 2019; Yaguchi & Hidaka, 2018). However, such studies have relied predominantly on a single type of task, the Simultaneity Judgment task (SJ).

The SJ is one of the two most common tasks used to investigate perception of audiovisual asynchrony, the other being the Temporal Order Judgment task (TOJ). In SJ, participants are asked to judge the synchrony between the auditory and visual information, whereas in TOJ they are asked to determine whether the auditory or visual information was presented first. Research has increasingly shown that there are important differences in the neural and perceptual mechanisms underlying how temporal judgments are made in SJ and TOJ tasks (e.g., Binder, 2015; Love et al., 2013, 2018). For example, simultaneity judgments require estimation of the temporal correspondence of the audio and visual cue and thus depend on a more global level of processing (considering the stimulus as a whole), whereas temporal order judgments could in principle be performed by focusing on only one sensory cue to detect whether it came first or not, thus depending on more local level processing (e.g., attending to the one cue that arrives first without the need to wait for the arrival of the other; Love et al., 2013).

Given these differences between SJ and TOJ, the temporal binding hypothesis of ASD (Brock et al., 2002) would predict differing performance in these tasks for autistic people, on the basis that ASD is associated with weak central coherence (Frith, 2003), which means that autistic people may preferentially employ a perceptual processing style which focuses mostly on local rather than global aspects of information. If a lower ability to detect audiovisual asynchrony among autistic individuals is due to difficulties in processing of global information (i.e., difficulties in assessing the temporal co-occurrence of the auditory and visual cues together) then one would expect to see a more pronounced difference between autistic and non-autistic individuals in SJ compared to TOJ. Therefore, the lack of existing research directly comparing the performance of autistic people and neurotypical controls on both SJ and TOJ represents an important gap in the literature, as it remains unclear whether the observed difference points to a general lower ability of autistic individuals to process audiovisual asynchrony, or whether these differences are task and perhaps stimulus specific.

Previous research has shown that audiovisual processing of different stimulus types is also based on distinct perceptual mechanisms (Love et al., 2013; Petrini et al., 2020) and that the complexity of a stimulus influences temporal binding of the component cues

(e.g., Arrighi et al., 2006; Petrini et al., 2009; Vatakis & Spence, 2006a, 2006b). Consistent with this, the extent to which differences in temporal processing have been shown to be evident in autistic people compared to neurotypical controls appears to be dependent on the complexity and social salience of the stimuli, even within studies using the same SJ task, with processing differences being most evident for complex, social stimuli such as speech stimuli (Feldman et al., 2018; Stevenson, Segers, et al., 2014, Stevenson, Siemann, et al., 2014). This again could relate to the temporal binding hypothesis of ASD, as processing more complex cues with stronger semantic correspondence (such as a male face talking with a male voice or a drummer's movement producing a drumming sound) would likely rely more on a global type of processing due to the effect of the unity assumption, which describes the situation where semantic matching of information in different modalities strengthens the perception that the two cues belong to the same event and source (Chen & Spence, 2017). This highlights the importance of not only investigating differences in the performance of autistic relative to neurotypical individuals for different types of temporal judgment tasks, but also comparing these groups on different stimulus conditions across different tasks.

An important further consideration is that the majority of the existing research examining differences in audiovisual temporal processing among autistic individuals has been conducted with children and adolescents (e.g., de Boer-Schellekens et al., 2013; Stevenson, Siemann, et al., 2014). Recent evidence suggests that multisensory processing differences in autistic individuals may diminish during adolescence and later development (Ainsworth & Bertone, 2023; Foxe et al., 2015), and so while investigating changes associated with development is not an aim of this research project, investigating audiovisual temporal processing in a group of autistic adults may provide some evidence toward indicating if reduced sensitivity to audiovisual asynchrony associated with ASD persists later in adulthood.

In this study we examined the ability to detect audiovisual asynchrony in a group of autistic adults when asked to report SJ and TOJ for different audiovisual stimuli, ranging from flashes and beeps to complex human actions and speech, in order to more fully characterize differences in audiovisual temporal processing between autistic and neurotypical adults. To better understand why certain processes may be different in autistic and neurotypical adults we used an Independent-Channels Model (Alcalá-Quintana & García-Pérez, 2013) to derive and compare specific perceptual and decisional parameters underlying overall task performance for each group. In line with the temporal binding hypothesis of ASD (Brock et al., 2002) we did not expect any difference in unimodal sensory processing between the two groups, rather we expected to see differences in decisional mechanisms depending on multisensory temporal resolution.

This research should allow for better understanding of the wider implications of audiovisual temporal processing differences in autism for social processing and perception more generally.

METHODS

Participants

Participants in the study were 26 autistic adult males and 26 age-, sex- and IQ-matched neurotypical participants (Table 1). Given established differences in male and female cognitive profiles in this population (Hull et al., 2017) and the higher prevalence of ASD among males (Fombonne, 2009), we chose to focus exclusively on a male sample. All participants in the ASD group reported a diagnosis of an ASD based on DSM-IV criteria from a qualified clinician (APA, 2000). All were native English speakers, had normal or corrected-to-normal vision and reported no hearing difficulties. The Autism Quotient (Baron-Cohen et al., 2001), a 50 item self-report scale designed to measure autistic traits, was administered to participants and supported the diagnostic status of the ASD group ($M = 36.64$, $SD = 8.80$), based on a cut-off score of 28 for a diagnosable ASD, and confirmed the assumption that individuals in the neurotypical (NT) group were unlikely to have an ASD ($M = 12.57$, $SD = 3.70$). The participants were matched pairwise on age ($t (50) = 0.45$, $p = 0.656$) and group-wise on full scale IQ (FSIQ) ($t (50) = -0.56$, $p = 0.580$) as measured using the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999).

The experimental procedures were approved by the School of Psychology at the University of Glasgow and also the Greater Glasgow and Clyde National Health Service ethics board. All participants provided informed written consent prior to participating in the study.

STIMULI

Three stimulus types were used: beep-flash (BF), point-light drumming (PLD) and face-voice (FV), which are shown in Figure 1. These three stimulus types have been used previously to study audiovisual perception in neurotypical individuals (Love et al., 2013) and were used in

TABLE 1 Participant demographic information.

Group	Age		AQ		FSIQ	
	Mean	SD	Mean	SD	Mean	SD
ASD	26.62	7.01	36.64	8.80	117.54	11.14
NT	25.81	5.93	12.57	3.70	119.08	8.63

Abbreviations: AQ, Autism Quotient; ASD, autistic group; FSIQ, Full scale IQ; NT, neurotypical group.
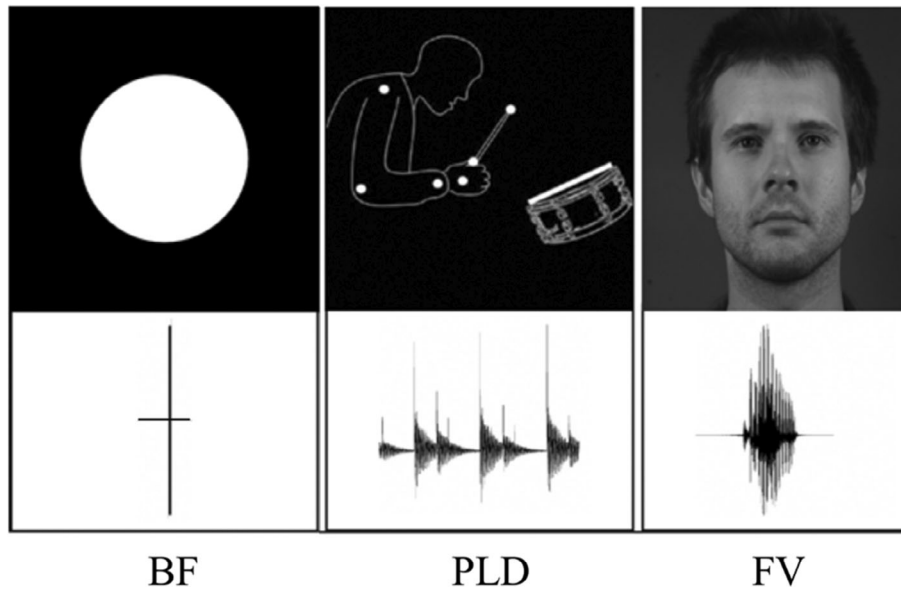
FIGURE 1 Stimulus types used in the Simultaneity (SJ) and Temporal Order Judgment (TOJ) tasks. *Note*: The top panel shows the visual information that participants were presented with for each stimulus type. The bottom panel shows the auditory waveform for each type of stimulus. The beep-flash (BF) stimuli consisted of a white flash on a black background and a beep sound. For the point-light drumming (PLD) stimuli, the figure shows one frame from the video clips and the waveform of the corresponding drumbeat. The outlines of the drum and drummer are for illustrative purposes only, as participants would have seen only the point-lights. For the face-voice (FV) stimuli, the figure shows one frame from the video clips and the waveform represents the spoken word "tomorrow". Please note that the images are not to scale and that the stimuli were standardized such that the size of the white flash for the BF stimuli (which subtended a visual angle of 4.4°) approximated the area of the drummer's arm and the speaker's mouth in the PLD and FV displays, respectively.

the current study because they varied in complexity and social nature, with BF stimuli being the least complex and the FV stimuli being the most complex in terms of social and visual contextual information. For a more detailed consideration of the relative complexity of the different stimulus types, see the supplementary material. For each stimulus type, the auditory and visual cues were separated in time to create 11 Stimulus Onset Asynchrony (SOA) levels: five audio-leading, five video-leading and one synchronous. For the BF and PLD stimuli the SOA levels used were 333, 267, 200, 133, 67 and 0 ms. A wider range of SOAs was used for the FV displays in line with previous research (e.g., Stevenson et al., 2010; Van Wassenhove et al., 2007): 400, 320, 240, 160, 80 and 0 ms. Further information about the stimuli can be found in the supplementary material.

Procedure

The experiments took place in a quiet and dimly lit room and participants were seated such that the viewing distance from the monitor displaying the stimuli was approximately 90 cm. The experiment was run separately for each stimulus type and the order of stimulus types was randomized for every participant. For each stimulus type, the experiment consisted of 24 blocks: half of the blocks were Simultaneity Judgment (SJ) blocks and the other half were Temporal Order Judgment (TOJ) blocks,

and these were presented in a randomized order. At the start of the experiment the participants read through the instructions and for each stimulus type they had the chance to complete three practice trials of each of the two tasks (SJ and TOJ) and then to ask any questions to clarify the experiment. The experimenter then left the room and the participants started the experiment by pressing a key.

At the start of every block of the task, instructions appeared on screen for 4 s to indicate whether the block that followed would be an SJ or TOJ block. Within each block there were 11 trials: one presentation of each SOA level for the given stimulus type. Participants could only make a response once they had watched the entire stimulus. After each stimulus the current task question and possible responses were displayed on screen until the participant responded, which triggered the next trial. During SJ blocks participants were asked to press "1" on the keyboard if they believed the audio and visual cues were presented synchronously and "2" if they perceived them as being asynchronous. During blocks of the TOJ task, they were asked to press "1" if they perceived the video as being presented first and "2" if they perceived that the audio was presented first. After completing the experiment for each stimulus type, participants completed a debrief questionnaire which asked them to rate the difficulty of the two tasks on a five-point Likert scale ranging from "easy" to "very difficult" (Love et al., 2013). If participants gave the two tasks the same difficulty rating the

questionnaire also included a forced choice question: "Which task did you find more difficult?". Analysis of the difficulty ratings for each task and stimulus combination showed that participants found the TOJ more difficult than the SJ across all conditions, although there were no significant group differences in difficulty ratings evident for any condition. These results are reported in full in the supplementary material.

Participants completed 12 blocks of the SJ and TOJ for each stimulus type which meant that they had data for 12 trials per SOA level for each combination of task and stimulus type. This is a similar number of trials to what has been used in previous research (Vatakis & Spence, 2006a, 2006b) and Petrini et al. (2010) showed that results are comparable regardless of whether 10 or 20 trials are used per SOA level. The participants were encouraged to take breaks between completing the tasks for each stimulus type, and overall the experiment took approximately 1.25 h to complete.

Data analysis

In line with the methods used by Love et al. (2013), for SJ the proportion of synchronous responses at each SOA level was fitted to a Gaussian probability density function, and for TOJ the proportion of video first responses was fitted to a Gaussian cumulative distribution function separately for each participant for each stimulus type. These fits derived two parameters of interest: the point of subjective simultaneity (PSS) and the audiovisual synchrony window (ASW). The PSS represents the level of SOA that participants perceived as most synchronous; it was taken as the maximum of the best-fitting SJ curve and the 50% point from the TOJ curve. The ASW represents the range of SOAs, centered on the PSS, within which participants could not reliably perceive asynchrony or cue order, and this was defined by the standard deviation of each best-fitting Gaussian.

R^2 values (which represent the goodness-of-fit between the data and fitted function) were calculated for estimates of the ASW and PSS, and values of below 0.5 were regarded as indicating that participants were unable to achieve a task/stimulus combination, and so these cases were excluded from the analysis (Love et al., 2013). Cases were also excluded from the analysis if estimates of the ASW or PSS lay outside the range of SOAs. These exclusion criteria were the same as were used in Love et al. (2013). We opted to run the analyses separately for each stimulus type as different cases had to be excluded for each stimulus type and we wanted to retain as much of the data as possible for each stimulus type. This decision was also theoretically driven, as previous research has shown that audiovisual processing of different stimulus types is based on distinct perceptual mechanisms (Love et al., 2013; Petrini et al., 2020).

Following the main analysis of ASW and PSS estimates, we planned to use the Independent-Channels Model (ICM; Alcalá-Quintana & García-Pérez, 2013) to follow up on any significant group differences to allow us to infer about the potential mechanisms underlying group differences in behavioral responses. In contrast to the traditional psychometric models, which explain just the shape of the distribution of behavioral responses in temporal discrimination tasks, the ICM is a generative model that represents the underlying sensory and decisional processes that lead to the pattern of responses. Further information about the ICM model and parameters may be found in the supplementary material.

Independent samples t-tests were predominantly used to compare PSS and ASW values and estimates of ICM parameters between participants in the ASD and NT groups separately for each task and stimulus condition. The parametric assumptions of an independent samples t-test were checked by using Levene's test to check for the assumption of homogeneity of variance and the Shapiro–Wilk test to confirm whether the data were likely to be normally distributed. Where the Shapiro–Wilk test indicated that the assumption of normality was potentially violated, the distribution of the data was visualized in a histogram to determine whether the distribution of the data still approximated the normal distribution. In cases where these checks suggested that either of the assumptions were violated, a Mann–Whitney U test was used rather than an independent samples t-test. Throughout the results section, where 95% confidence intervals are reported these are the confidence intervals for the mean difference between groups.

RESULTS

PSS and ASW

Beep-Flash stimuli

Based on the exclusion criteria outlined above, two cases from the ASD group were excluded for the TOJ task; no cases were excluded for the SJ task. For the SJ task the PSS, $t(50) = -0.24$, $p = 0.811$, 95% CI[$-26.56, 20.87$], and ASW, $t(50) = 1.72$, $p = 0.092$, 95% CI[$-3.65, 46.34$], did not significantly differ between ASD and NT groups. For the TOJ task the analyses also revealed that neither PSS, $t(48) = -0.65$, $p = 0.517$, 95% CI[$-52.44, 26.73$], nor ASW, $U = 265$, $p = 0.361$, 95% CI[$-6.93, 76.27$], differed significantly between groups. The results are presented in Figures 2 and 3 for PSS and ASW estimates respectively. Mean R^2 values for each group and task are reported in Table 2.

Point-Light drumming stimuli

For the SJ task no cases were excluded. The PSS did not differ significantly between groups, $t(50) = 0.41$,
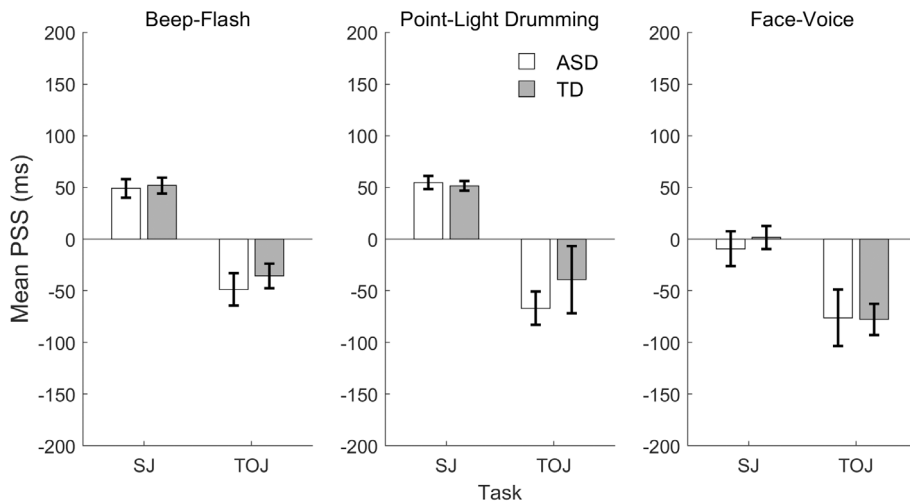
FIGURE 2 Mean estimates of PSS. *Note*: Error bars represent the standard error of the mean.
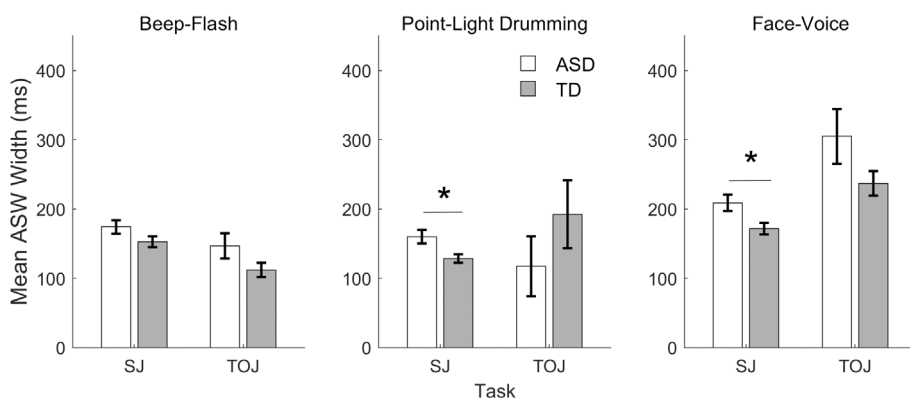

FIGURE 3 Mean estimates of ASW width. *Note*: Error bars represent the standard error of the mean.

TABLE 2 Mean (SD) R^2 values for the different stimulus and task conditions.

	Group and task			
	ASD		NT	
	SJ	TOJ	SJ	TOJ
Beep-Flash	0.89	0.90	0.92	0.94
	(0.06)	(0.09)	(0.05)	(0.06)
Point-Light Drumming	0.88	0.73	0.91	0.78
	(0.06)	(0.14)	(0.05)	(0.15)
Face-Voice	0.87	0.82	0.91	0.87
	(0.08)	(0.13)	(0.07)	(0.09)

significantly differed between groups. Additional within-subjects comparisons comparing the ASW in the SJ and TOJ tasks within the ASD group were carried out to follow up on these main findings, and are reported and discussed in the supplementary material.

Face-Voice stimuli

Based on the exclusion criteria outlined in the method, one case was excluded for the ASD group in the SJ task, and eight cases for the ASD group and two cases for the NT group were excluded for the TOJ task. For the SJ task the PSS did not differ significantly between groups, t (41.37) $= -0.54$, $p = 0.590$, 95% CI[-52.29, 30.11], but the ASW was significantly larger in the ASD group compared to the NT group, $U = 198$, $p = 0.017$, 95% CI[7.98, 66.73], $d = 0.72$. For the TOJ task, the analyses revealed that neither PSS, t (25.74) $= 0.04$, $p = 0.967$, 95% CI[-68.47, 71.34], nor ASW, $U = 176$, $p = 0.309$, 95% CI[-29.71, 165.67], significantly differed between groups. Additional within-subjects comparisons comparing the ASW in the SJ and TOJ tasks within the ASD group were carried out to follow up on these main findings, and are reported and discussed in the supplementary material.

$p = 0.682$, 95% CI[-12.57, 19.05], but the ASW was significantly larger in the ASD group compared to the NT group, t (50) $= 2.80$, $p = 0.007$, 95% CI[8.96, 54.38], $d = 0.78$. For the TOJ data, 17 participants had to excluded from the ASD group and 18 from the NT group, resulting in just nine and eight participants in each group, respectively. The analyses revealed that neither PSS, $U = 28$, $p = 0.441$, 95% CI[-109.83, 54.49], nor ASW, $U = 25$, $p = 0.290$, 95% CI[-225.61, 75.46],

TABLE 3 Estimates of ICM parameters for the point-light drumming SJ and TOJ tasks.

| | Group and task | | | |
| | ASD | | NT | |
	SJ	TOJ	SJ	TOJ
N	26	9	26	8
λ_a (ms) [s.e.m]	0.34 [0.07]	0.45 [0.13]	0.44 [0.08]	0.36 [0.07]
λ_v (ms) [s.e.m]	0.24 [0.06]	0.24 [0.09]	0.29 [0.06]	0.55 [0.13]
τ (ms) [s.e.m]	−40.85 [8.06]	−51.58 [24.31]	−41.48 [6.91]	29.68 [34.49]
δ (ms) [s.e.m]	186.92 [9.05]	173.66 [11.21]	169.16 [7.68]	125.68 [13.87]
ξ (ms) [s.e.m]	–	0.85 [0.06]	–	0.62 [0.14]
εTF (ms) [s.e.m]	0.07 [0.02]	0.03 [0.01]	0.01 [<0.01]	0.13 [0.05]
εRF (ms) [s.e.m]	0.13 [0.04]	0.47 [0.05]	0.06 [0.02]	0.33 [0.07]
εSJ-S (ms) [s.e.m]	0.04 [0.01]	–	0.03 [0.01]	–

Note: λa and λv are sensory parameters which describe the rate of processing for auditory and visual information, respectively. τ represents the latency difference at which the two cues arrive at the central mechanism. δ is a decisional parameter which refers to the smallest time difference that the central mechanism can resolve. ξ is specific to TOJ and describes the bias of observers to more often respond "audio first" or "video first". εTF = error term for auditory-leading trials; εRF = error term for visual-leading trials; εSJ-S = error term for simultaneous trials on the SJ task.

TABLE 4 Estimates of ICM parameters for the face and voice SJ and TOJ tasks.

| | Group and task | | | |
| | ASD | | NT | |
	SJ	TOJ	SJ	TOJ
N	25	18	26	24
λ_a (ms) [s.e.m]	0.32 [0.07]	0.35 [0.08]	0.31 [0.08]	0.34 [0.08]
λ_v (ms) [s.e.m]	0.12 [0.05]	0.30 [0.07]	0.10 [0.04]	0.30 [0.07]
τ (ms) [s.e.m]	33.10 [15.11]	22.23 [24.77]	11.26 [13.70]	16.26 [22.82]
δ (ms) [s.e.m]	226.93 [9.28]	156.01 [21.25]	211.37 [12.05]	172.69 [8.65]
ξ (ms) [s.e.m]	–	0.50 [0.06]	–	0.62 [0.03]
εTF (ms) [s.e.m]	0.15 [0.05]	0.17 [0.06]	0.02 [0.01]	0.06 [0.02]
εRF (ms) [s.e.m]	0.10 [0.02]	0.08 [0.03]	0.02 [0.01]	0.07 [0.01]
εSJ-S (ms) [s.e.m]	0.11 [0.02]	–	0.09 [0.02]	–

Note: λa and λv are sensory parameters which describe the rate of processing for auditory and visual information, respectively. τ represents the latency difference at which the two cues arrive at the central mechanism. δ is a decisional parameter which refers to the smallest time difference that the central mechanism can resolve. ξ is specific to TOJ and describes the bias of observers to more often respond "audio first" or "video first". εTF = error term for auditory-leading trials; εRF = error term for visual-leading trials; εSJ-S = error term for simultaneous trials on the SJ task.

ICM parameters

ICM parameters were compared between groups only for the stimulus types where we observed a significant difference between groups for estimates of the ASW, as the purpose of this further analysis was to understand the potential mechanisms underlying observed group differences in temporal processing. For these stimulus types we analyzed the ICM parameters for both tasks, to try to better understand if these task-related group differences could be explained by different underlying mechanisms. For the sake of brevity, only the results for the ICM parameters showing significant differences between groups are reported here. The full ICM results are reported in the supplementary material.

Point-Light drumming stimuli

Mean values for the ICM parameters in the SJ and TOJ tasks for PLD stimuli are reported in Table 3. For auditory-leading trials in the SJ task, the results showed that the ASD group was more likely to make errors than the NT group, $U = 243$, $p = 0.044$, 95% CI[0.01, 0.11],

$d = 0.67$. That is, the ASD group was more likely to erroneously judge asynchronous cues as being synchronous for auditory-leading trials. In addition, for the TOJ task, τ parameter estimates (which represent the latency difference at which the two cues arrive at the central mechanism) were significantly higher in the NT group compared to the ASD group, $U = 15$, $p = 0.043$, 95% CI [−169.61, 7.09], $d = −0.95$, and δ parameter estimates (which indicate the smallest time difference that the central mechanism can resolve) were significantly higher in the ASD compared to the NT group, $U = 12$, $p = 0.021$, 95% CI[10.33, 85.64], $d = 1.32$. There were no significant group differences in estimates of any of the other ICM parameters for either task.

Face-Voice stimuli

Mean values for the ICM parameters in the SJ and TOJ tasks for FV stimuli are reported in Table 4. For auditory-leading trials of the SJ task, the results showed that the ASD group was more likely to make errors than the NT group, $U = 200$, $p = 0.006$, 95% CI[0.03, 0.23],

$d = 0.76$. Similarly, for visual-leading trials, the ASD group was also more likely to make errors than the NT group, $U = 208$, $p = 0.019$, 95% CI[0.03, 0.13], $d = 0.93$. That is, the ASD group was more likely to erroneously judge asynchronous cues as being synchronous for both visual- and auditory-leading trials. There were no significant group differences in estimates of any of the other ICM parameters for either task.

DISCUSSION

To investigate the underlying processes of reduced sensitivity to audiovisual asynchrony observed among autistic individuals (e.g., de Boer-Schellekens et al., 2013; Stevenson, Siemann, et al., 2014), performance on SJ and TOJ was compared between groups of autistic and neurotypical adults using stimuli varying in complexity and social relevance. For SJ, the autistic group had a significantly wider ASW compared to the control group, and this effect was specific to judgments about the more complex, social stimulus types (face and voice and point-light drumming) and did not apply to the simple beep and flash stimuli. By contrast, across the different stimulus types, ASW estimates were comparable between the two groups for TOJ, and PSS estimates were comparable between groups for both the SJ and TOJ tasks. Together, these results add new insight on the previously found wider ASWs for autistic children and adults with high levels of autistic traits (e.g., de Boer-Schellekens et al., 2013; Stevenson, Siemann, et al., 2014; Van Laarhoven et al., 2019; Yaguchi & Hidaka, 2018), by showing that differences in audiovisual temporal processing in autistic adults are specific to simultaneity judgments involving complex, social stimuli.

The discrepancy in the results for the SJ and TOJ tasks for the two groups is consistent with the predictions of the temporal binding hypothesis (Brock et al., 2002) that autistic participants would be more likely to achieve neurotypical levels of performance on TOJ. SJ require the observer to estimate the temporal correspondence of the auditory and visual cues and thus may depend on a more global level of processing. This contrasts with TOJ which could in principle be performed by focusing on only one sensory cue to detect whether or not it came first, thus depending on more local level processing. This suggests that difficulties with audiovisual integration in autistic individuals may be linked to difficulties at a global level of information processing, in line with theories of a central coherence deficit (Frith, 2003) and temporal binding deficit in ASD (Brock et al., 2002).

This explanation linking task-dependent differences in temporal processing to difficulties with global processing in autistic individuals is further reinforced by the finding of wider ASWs in the ASD group compared to the NT group specifically for SJ for complex stimuli such as point-light drumming and face-voice stimuli. This result is in line with Stevenson, Siemann, et al. (2014), who observed wider ASW estimates for autistic children compared to neurotypical controls, but only for the more complex face and voice stimuli and not for simpler non-social stimuli. This again could relate to the temporal binding hypothesis of ASD, as processing more complex cues with stronger semantic correspondence likely involves stronger involvement of global processes, due to semantic congruence strengthening the assumption that the two cues in these complex stimuli pertain to the same event and source (Chen & Spence, 2017), making these semantically matching cues more difficult to consider separately, especially for the SJ task. Together, these results support the conclusion that differences in the precision of audiovisual temporal processing in autistic individuals become more evident with increasing stimulus complexity and social relevance, and so may have a particular impact on social functioning in everyday contexts.

The finding that there were no group differences in ASW width for TOJ contrasts with previous results showing that autistic adolescents demonstrate a wider ASW in TOJ tasks compared to neurotypical controls (de Boer-Schellekens et al., 2013). This suggests that while more general difficulties with both TOJ and SJ may be evident for autistic individuals earlier in development, differences in audiovisual temporal processing that are specific to simultaneity judgments may be most likely to persist into adulthood. Autistic adults have previously been shown to develop compensatory strategies in tasks that autistic children typically have difficulties with (McKay et al., 2012). Therefore, we could argue that in the case of TOJs it may be easier for autistic adults to develop compensatory strategies over time by depending more on local level processing of individual sensory cues, which is not possible for SJs which require the stimulus to be considered as a whole. However, investigating changes associated with development was not an aim of this research project, so future research should seek to directly compare a younger population of autistic individuals with autistic adults before any conclusions about developmental changes in audiovisual temporal processing in this population can be made.

The fact that we only observed group differences in performance for the SJ tasks, and not the TOJ tasks, supports the general finding in the literature of different mechanisms supporting simultaneity and temporal order judgments (Love et al., 2013, 2018; Van Eijk et al., 2008), and we have suggested that differences in the requirement of global versus local processing in the two tasks may be responsible for the differing performance of the ASD group relative to the neurotypical control group in the SJ task. However, it is important to acknowledge that there are other potential differences between the two tasks that could affect participant performance, for example, the TOJ task (which is generally rated as more difficult than the SJ, e.g., Love et al., 2013; Petrini et al., 2020, as was also the case in our study), may entail an additional level

of complexity because participants may engage in a two-step decision process in order to decide which out of the auditory or visual cues was presented first (i.e., step 1 – are the cues synchronous? Step 2 – which came first?).

While this difference in the complexity of the two tasks provides another explanation for why participants' performance may differ between the two tasks, it doesn't help us to explain our findings that performance in the ASD group relative to the neurotypical control group was specifically impaired for the SJ task, since according to this explanation, the added level of complexity is for the TOJ task for which we did not find any difference in the ASW between the two groups. Our hypothesis that performance is specifically impaired for the ASD group for the SJ task due to its potential reliance on global processing abilities fits with what we know about difficulties with global versus local processing among autistic individuals, and so provides one potential explanation for the observed results. However, before any concrete conclusions can be made about why we observed differences between the two groups specific to the SJ task, it will be necessary to conduct further research investigating the specific executive strategies that are used by participants in the two tasks.

To follow up on the significant group differences in ASW estimates and allow us to infer about the potential mechanisms underlying these differences, we also fitted an Independent-Channels Model to the data (Alcalá-Quintana & García-Pérez, 2013). Examining the estimated parameters describing unisensory and decisional factors in the SJ and TOJ tasks for the face-voice and point-light drumming stimuli, revealed, as expected, that there were no significant group differences in any of the unisensory parameters for either task. The finding that audiovisual temporal processing differences between the two groups could not be explained by unisensory processing parameters supports previous findings by Stevenson, Siemann, et al. (2014) who showed that there were no group differences in performance for temporal judgments involving only audio or visual cues.

However, group differences were found for the ICM-derived error parameters, indicating that autistic individuals were more likely to make errors on auditory- and visual-leading trials of the SJ task for the face-voice stimuli, and were more likely to make errors on auditory-leading trials of the SJ task for the point-light drumming stimuli, erroneously judging asynchronous cues as being "synchronous" more often than the neurotypical group. This is of interest because it has been shown that individuals are generally better at detecting audio-leading asynchronies (Dixon & Spitz, 1980; Love et al., 2013; Van Eijk et al., 2008), which has been attributed to the fact that in natural situations, auditory cues generally lag visual cues, and so audio-leading asynchronies are more noticeably different based on our everyday experience and natural heuristics (e.g., Chandrasekaran et al., 2009). In line with previous results (e.g., Love et al., 2013;

Petrini et al., 2020), the current findings show that neurotypical individuals find auditory-leading asynchronies easier to detect, although this does not seem to be the case for autistic individuals, who demonstrated a significantly higher error rate for auditory-leading trials compared to neurotypical controls. This discrepancy in error rate between groups occurred in spite of the finding that there were no significant differences in self-reported difficulty ratings between groups for any task or stimulus type. This suggests that this difference cannot simply be explained by perceived task difficulty but seems to be related to a lower ability of autistic adults to use prior experience and natural heuristics to improve their audio-visual temporal perception.

These results are encouraging for potential interventions to improve sensory processing for autistic individuals, as they suggest that observed differences in audiovisual temporal processing may be due to lower ability to integrate existing heuristics into temporal processing, which is something that could be improved using training. It has been shown that ASW width becomes smaller through training (Che et al., 2022; Powers et al., 2009; Stevenson et al., 2013), and that those with the widest ASWs improve the most after training.

LIMITATIONS

One limitation of the study is that for the TOJ point-light drumming task a large number of cases had to be excluded due to estimates of the PSS and ASW lying outside of the range of SOAs, indicating that participants were unable to achieve the task/stimulus combination. We used a non-parametric test to account for the limited sample size, but these analyses were still likely underpowered and so the results should be interpreted with caution. The reasons for the high number of exclusions for this task and stimulus type are discussed in more detail in the supplementary material.

A more direct way to address our research question concerning the role of stimulus complexity in manifesting differences in audiovisual temporal processing between autistic and neurotypical participants would have been to run a 2×(3) ANOVA with participant group (ASD, TD) and stimulus type (BF, FV, PLD) as the two factors. However, in this study we opted to run the analyses separately for each stimulus type as different cases had to be excluded for each stimulus type (in particular, there was a high number of exclusions in the TOJ PLD task, as described above) and we wanted to retain as much of the data as possible for each stimulus type. This decision was also theoretically driven, as previous research has shown that audiovisual processing of different stimulus types is based on distinct perceptual mechanisms (Love et al., 2013; Petrini et al., 2020).

Another potential limitation of the current study was the decision to use an ICM model with a high number of

error parameters, as this increases the risk of over-parameterization. However, we felt that it was important to include all of the possible error parameters, as errors and biases have too often been unaccounted for in psychophysics research, and previous developmental research has demonstrated important individual differences in these parameters (Chen et al., 2016; Petrini et al., 2020).

A limitation to the generalizability of the findings is the fact that this study focused only on adult males. Sex and gender differences in the etiology and symptoms of ASD have been the subject of much study in recent years (Lai et al., 2015), which has identified important differences in core autism spectrum condition traits between males and females with ASD (Hull et al., 2017). Given established differences in male and female cognitive profiles in this population and the higher prevalence of ASD among males (Fombonne, 2009), we chose to focus exclusively on a male sample. This has allowed us to better characterize audiovisual temporal processing in adults males with autism, but the extent to which these findings generalize to other genders cannot be determined because, as of yet, there are few studies exploring sensory differences between males and females with autism (Gould, 2017).

CONCLUSION

This study investigated audiovisual integration in autistic adult men using SJ and TOJ tasks and showed that some of the differences in audiovisual temporal binding which have been observed in autistic individuals earlier in development may persist into adulthood. Furthermore, it was found that differences in audiovisual temporal binding in autistic adults were specific to SJs involving complex, social stimuli. This suggests that difficulties with audiovisual integration in autistic individuals may be linked to difficulties with global level information processing and are likely to impact particularly on social functioning.

CONFLICT OF INTEREST STATEMENT
The authors declare no competing interests.

DATA AVAILABILITY STATEMENT
Data are available upon written request to the corresponding author. Data have not been shared publicly in a repository because participants did not give written consent to public data sharing at the time of data collection.

ORCID
Naomi Heffer https://orcid.org/0000-0002-6597-0134

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

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods*, *45*, 972–998. https://doi.org/10.3758/s13428-013-0325-2

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). American Psychiatric Pub.

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

Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, *6*(3), 266–268. https://doi.org/10.1167/6.3.6

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

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

Binder, M. (2015). Neural correlates of audiovisual temporal processing – Comparison of temporal order and simultaneity judgments. *Neuroscience*, *300*, 432–447. https://doi.org/10.1016/j.neuroscience.2015.05.011

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

Brock, J. O. N., Brown, C. C., Boucher, J., & Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Development and Psychopathology*, *14*(2), 209–224. https://doi.org/10.1017/S0954579402002018

Centers for Disease Control and Prevention. (2023). Data and statistics on ASD. https://www.cdc.gov/ncbddd/autism/data.html

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*(7), e1000436. https://doi.org/10.1371/journal.pcbi.1000436

Che, Y., Jicol, C., Ashwin, C., & Petrini, K. (2022). An RCT study showing few weeks of music lessons enhance audio-visual temporal processing. *Scientific Reports*, *12*(1), 20087. https://doi.org/10.1038/s41598-022-23340-4

Chen, L., Zhang, M., Ai, F., Xie, W., & Meng, X. (2016). Crossmodal synesthetic congruency improves visual timing in dyslexic children.

Research in Developmental Disabilities, *55*, 14–26. https://doi.org/10.1016/j.ridd.2016.03.010

Chen, Y. C., & Spence, C. (2017). Assessing the role of the 'unity assumption' on multisensory integration: A review. *Frontiers in Psychology*, *8*, 445. https://doi.org/10.3389/fpsyg.2017.00445

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

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

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 & Biobehavioral Reviews*, *95*, 220–234. https://doi.org/10.1016/j.neubiorev.2018.09.020

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203

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

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*, *25*(2), 298–312. https://doi.org/10.1093/cercor/bht213

Frith, U. (2003). *Autism: Explaining the enigma*. Blackwell Publishing.

Gelder, B. d., Vroomen, J., & Van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology*, *3*(1), 69–86. https://doi.org/10.1080/09541449108406220

Gould, J. (2017). Towards understanding the under-recognition of girls and women on the autism spectrum. *Autism*, *21*(6), 703–705. https://doi.org/10.1177/1362361317706174

Grossman, R. B., Schneps, M. H., & Tager-Flusberg, H. (2009). Slipped lips: Onset asynchrony detection of auditory-visual language in autism. *Journal of Child Psychology and Psychiatry*, *50*(4), 491–497. https://doi.org/10.1111/j.1469-7610.2008.02002.x

Hull, L., Mandy, W., & Petrides, K. V. (2017). Behavioural and cognitive sex/gender differences in autism spectrum condition and typically developing males and females. *Autism*, *21*(6), 706–727. https://doi.org/10.1177/1362361316669087

Irwin, J. R., Tornatore, L. A., Brancazio, L., & Whalen, D. (2011). Can children with autism spectrum disorders "hear" a speaking face? *Child Development*, *82*(5), 1397–1403. https://doi.org/10.1111/j.1467-8624.2011.01619.x

Kwakye, L., Foss-Feig, J., Cascio, C., Stone, W., & Wallace, M. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in Integrative Neuroscience*, *4*, 129. https://doi.org/10.3389/fnint.2010.00129

Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/gender differences and autism: Setting the scene for future research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(1), 11–24. https://doi.org/10.1016/j.jaac.2014.10.003

Lane, A. E., Young, R. L., Baker, A. E., & Angley, M. T. (2010). Sensory processing subtypes in autism: Association with adaptive behavior. *Journal of Autism and Developmental Disorders*, *40*, 112–122. https://doi.org/10.1007/s10803-009-0840-2

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS One*, *8*(1), e54798. https://doi.org/10.1371/journal.pone.0054798

Love, S. A., Petrini, K., Pernet, C. R., Latinus, M., & Pollick, F. E. (2018). Overlapping but divergent neural correlates underpinning audiovisual synchrony and temporal order judgments. *Frontiers in Human Neuroscience*, *12*, 274. https://doi.org/10.3389/fnhum.2018.00274

Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

McKay, L. S., Simmons, D. R., McAleer, P., Marjoram, D., Piggot, J., & Pollick, F. E. (2012). Do distinct atypical cortical networks process biological motion information in adults with autism spectrum disorders? *NeuroImage*, *59*(2), 1524–1533. https://doi.org/10.1016/j.neuroimage.2011.08.033

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

NHS England. (2020). People with a learning disability, autism or both (Pub. No. 000948). https://www.england.nhs.uk/wp-content/uploads/2020/01/Learning-disability-and-autism.pdf

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F., Puce, A., & Pollick, F. E. (2009). Multisensory integration of drumming actions: Musical expertise affects perceived audiovisual asynchrony. *Experimental Brain Research*, *198*(2), 339–352. https://doi.org/10.1007/s00221-009-1817-2

Petrini, K., Denis, G., Love, S. A., & Nardini, M. (2020). Combining the senses: The role of experience-and task-dependent mechanisms in the development of audiovisual simultaneity perception. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(10), 1105–1117. https://doi.org/10.1037/xhp0000827

Petrini, K., Holt, S. P., & Pollick, F. (2010). Expertise with multisensory events eliminates the effect of biological motion rotation on audiovisual synchrony perception. *Journal of Vision*, *10*(5), 2. https://doi.org/10.1167/10.5.2

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

Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental Disorders*, *43*, 775–784. https://doi.org/10.1007/s10803-012-1608-7

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry*, *48*(8), 813–821. https://doi.org/10.1111/j.1469-7610.2007.01766.x

Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., & James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *NeuroImage*, *49*(4), 3308–3318. https://doi.org/10.1016/j.neuroimage.2009.12.001

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., 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. https://doi.org/10.1523/JNEUROSCI.3615-13.2014

Stevenson, R. A., Wilson, M. M., Powers, A. R., & Wallace, M. T. (2013). The effects of visual training on multisensory temporal processing. *Experimental Brain Research*, *225*, 479–489. https://doi.org/10.1007/s00221-012-3387-y

Szelag, E., Kowalska, J., Galkowski, T., & Pöppel, E. (2004). Temporal processing deficits in high-functioning children with autism. *British Journal of Psychology*, *95*(3), 269–282. https://doi.org/10.1348/0007126041528167

Van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of

experimental method and stimulus type. *Perception & Psychophysics*, *70*, 955–968. https://doi.org/10.3758/PP.70.6.955

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), 1–11. https://doi.org/10.1038/s41598-019-46084-0

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

Vatakis, A., & Spence, C. (2006a). 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

Vatakis, A., & Spence, C. (2006b). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neuroscience Letters*, *393*(1), 40–44. https://doi.org/10.1016/j.neulet.2005.09.032

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. Psychological Corporation.

Yaguchi, A., & Hidaka, S. (2018). Distinct autistic traits are differentially associated with the width of the multisensory temporal binding window. *Multisensory Research*, *31*(6), 523–536. https://doi.org/10.1163/22134808-00002612

SUPPORTING INFORMATION

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

How to cite this article: Regener, P., Heffer, N., Love, S. A., Petrini, K., & Pollick, F. (2024). Differences in audiovisual temporal processing in autistic adults are specific to simultaneity judgments. *Autism Research*, *17*(5), 1041–1052. https://doi.org/10.1002/aur.3134