Autistic Communication and Imagination Sub-Traits Are Related to Audiovisual Temporal Integration in the Stream–Bounce Illusion

Ayako Yaguchi1,2,3,* and Souta Hidaka1

1 Department of Psychology, Rikkyo University, 1-2-26, Kitano, Niiza-shi, Saitama, 352-8558, Japan
2 Department of Rehabilitation for Brain Functions, Research Institute of National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Saitama, 359-8555, Japan
3 Japan Society for the Promotion of Science, Tokyo, 102-0083, Japan

Received 30 September 2019; accepted 16 November 2020

Abstract

Autism spectrum disorder (ASD) is characterized by atypical social communication and restricted and repetitive behaviors; such traits are continuously distributed across nonclinical and clinical populations. Recently, relationships between ASD traits and low-level multisensory processing have been investigated, because atypical sensory reactivity has been regarded as a diagnostic criterion of ASD. Studies regarding an audiovisual illusion (the double-flash illusion) reported that social communication difficulties are related to temporal aspects of audiovisual integration. This study investigated whether similar relationships exist in another audiovisual illusion (the stream–bounce effect). In this illusion, two visual objects move toward each other, coincide, and pass each other, and the presentation of a transient sound at their coincidence induces a dominant perception that they bounce away from each other. Typically developing adults were recruited to perform experimental trials involving the stream–bounce effect. We measured their ASD traits using the Autism-Spectrum Quotient. The total quotient score was not related to any behavioral measurements of the effect. In contrast, for participants with higher difficulty in communication, the greatest magnitude of the stream–bounce effect occurred when the presentation timing of the sound tended to follow the visual coincidence. Participants with higher difficulty in imagination also showed the greatest magnitude of the effect when the presentation timing of the sound preceded that of the visual coincidence. Our findings regarding the stream–bounce effect, along with previous findings regarding the double-flash illusion, suggest that atypical temporal audiovisual integration is uniquely related to ASD sub-traits, especially in social communication.

* To whom correspondence should be addressed. E-mail: yaguchi@rikkyo.ac.jp

© Yaguchi and Hidaka, 2020

DOI:10.1163/22134808-bja10042
This is an open access article distributed under the terms of the CC BY-NC 4.0 license.
Keywords
autism spectrum disorder, audiovisual integration, stream–bounce effect, social communication, Autism-Spectrum Quotient

1. Introduction

Individuals with autism spectrum disorder (ASD) can experience difficulties relating to higher cognitive function, such as difficulties regarding social communication and social interaction, as well as restricted interests and repetitive behaviors (American Psychiatric Association, 2013). In addition, atypical sensory processing (e.g., hyper- and hypo-reactivity) is also frequently observed in such individuals (Tomchek and Dunn, 2007), and it was added as a diagnostic criterion of ASD (American Psychiatric Association, 2013). Notably, studies have shown that, in ASD, such atypical sensory characteristics are related to the characteristics of social communication or restricted and repetitive behaviors (Baum et al., 2015; Jeste and Nelson, 2009; Leekam et al., 2011; Stevenson et al., 2014a).

In daily life, we establish coherent and robust perceptions of our environments by automatically integrating various inputs from multiple sensory modalities (Ernst and Bulthoff, 2004). Multisensory processing requires accurate temporal processing, because temporal correspondence is a key factor for integration (Dixon and Spitz, 1980; Spence and Squire, 2003). Previous crossmodal studies have focused on characteristics of audiovisual temporal processing to understand the atypical sensory processes in individuals who have been diagnosed with ASD (Feldman et al., 2018; Zhou et al., 2018). Such studies have focused on both low-level stimuli, such as flashes and beeps (de Boer-Schellekens et al., 2013; Noel et al., 2017; Stevenson et al., 2014a), as well as complex, higher-order stimuli such as faces and voices (e.g., Stevenson et al., 2014a; see also Beker et al. 2018 for a review). For example, the manner of perceptual learning for audiovisual simultaneity is reported to be different between an individual with ASD and a typically developing (TD) person (Turi et al., 2016). Further, some studies have examined the size of individuals’ temporal binding window, a temporal interval in which different sensory inputs can be integrated as a unified object. A narrow temporal binding window is assumed to reflect more accurate temporal processing (Conrey and Pisoni, 2006; Stevenson et al., 2012). For instance, Foss-Feig et al. (2010) used the double-flash illusion as a tool for measuring the temporal binding window regarding audiovisual integration (this illusion comprises the concurrent presentation of a single flash and multiple auditory beeps, which creates the illusory effect of the single flash being perceived as multiple flashes; Shams et al. 2002), and reported that individuals with a diagnosis of ASD
were shown to have a wider temporal binding window compared to TD individuals.

ASD characteristics can vary among individuals, and can be regarded as having a continuous distribution across the population in terms of their magnitude (Baron-Cohen et al., 2001; Woodbury-Smith et al., 2005). Based on this viewpoint, the Autism-Spectrum Quotient (AQ) was developed (Baron-Cohen et al., 2001) to measure ASD traits for individuals regardless of whether they had been diagnosed with ASD. In addition to providing total scores that reflect overall tendencies of ASD, the AQ enables us to measure five sub-traits underlying ASD: social skill, communication, attention switching, attention to detail, and imagination. Some unique relationships have been reported between ASD sub-traits and audiovisual temporal processing. In a simultaneity judgment (SJ) task, where participants are asked to judge whether the onsets of visual and auditory stimuli are simultaneous or not, as the total AQ and subscales increased, subjective simultaneity shifted toward the timing where the sound was presented first, except for the attention to detail scores (Donohue et al., 2012). Kawakami et al. (2018) reported that people with higher AQ scores in the social skills sub-trait had a narrower temporal binding window in the double-flash illusion. Further, Yaguchi and Hidaka (2018) demonstrated that people who have higher scores in the social skill and communication AQ sub-traits had narrower and wider temporal binding windows, respectively. These investigations of TD individuals suggest that the ASD sub-traits for social communication are particularly related to temporal processing of audiovisual integration.

Whereas the investigations on the temporal characteristics of audiovisual perpetual illusions contribute to understand the manner of perceptual binding processes of audiovisual inputs, previous findings showing the relationship between ASD and its sub-traits and the audiovisual temporal characteristics have been sourced from trials involving single low-level audiovisual illusions (e.g., the double-flash illusion). Thus, it remains unclear whether these relationships are generally observed for the audiovisual integration process. Considering this, the aim of the current study was to investigate whether the relationships between ASD traits and the temporal aspects of audiovisual integration also exist for another type of audiovisual illusion. Here, we chose to focus on the ‘stream–bounce effect’. This effect involves two identical visual objects moving towards each other, coinciding, and then moving apart in a two-dimensional display; individuals who observe this display perceive one of two motion trajectories: (1) that the objects ‘stream’ past each other or (2) that the objects ‘bounce off’ each other (Fig. 1). During the display, a transient auditory tone presented at or near the point when the two visual objects coincide commonly induces a
perceived bounce. This effect predominantly occurs when the sound is presented at the visual coincidence or 150 ms before the visual coincidence (Sekuler et al., 1997). We also measured the magnitude of the non-clinical participants’ ASD traits using the Japanese version of the AQ (Wakabayashi et al., 2004). This enabled us to investigate the relationships unique to each ASD sub-trait with a relatively large dataset (Kawakami et al., 2018; Yaguchi and Hidaka, 2018). If the relationships between the ASD sub-traits related to social communication and the temporal characteristic of the audiovisual integration observed in the double-flash illusion (Kawakami et al., 2018; Yaguchi and Hidaka, 2018) reflect a general aspect of ASD-specific sensory processing, a similar tendency would also be observed for the stream–bounce effect.

Figure 1. Schematic illustrations of the experimental design for the stream–bounce effect. White disks moved horizontally toward each other, coincided at the midpoint, and proceeded to move away from each other. The stimulus onset asynchrony between the visual coincidence and the beep was ±0–500 ms. The participants were asked to report their perceived movement of the disks (streaming or bouncing).
2. Methods

2.1. Participants

Ninety-one university students were recruited as participants for this study. Of these, 23 were excluded from our analyses because their data were not well-fitted to a Gaussian function in our analysis (determination coefficient $R^2 < 0.25$), or because they were regarded as outliers, as their data exceeded two standard deviations of the mean (SDs; see the data analysis section for details). All participants reported that they had normal or corrected-to-normal vision, and that they had normal hearing. All were naïve to the purpose of the experiment. The experimental procedures in this study were approved by the local ethics committee of Rikkyo University. Informed consent was obtained from each participant before conducting the experiments.

2.2. Apparatus

For the experiments, we presented the visual stimuli on a CRT display (FlexScan T776, 19-inch, EIZO, Ishikawa, Japan) and the auditory stimuli via an audio interface (EDIROL UA-25 EX, Roland, Shizuoka, Japan) and headphones (HDA200, Sennheiser, Wedemark, Germany). The participants were seated in front of the CRT display at a distance of 57.3 cm and their heads were maintained in a stable position using a chin rest. We confirmed synchronization of the visual and auditory stimuli using a digital oscilloscope (PDS5022TFT, OWON, Fu Jian, China). MATLAB (MathWorks, Inc., Natick, MA, USA), with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), was used to control the experiment. The participants responded using a keyboard. The experiments were conducted in a dark room. The Japanese version of the AQ (Wakabayashi et al., 2004) was used to measure participants’ ASD traits.

2.3. Stimuli

A red fixation cross (14 cd/m²) was initially presented on a black background (0.08 cd/m²). Then, two white disks (1 deg in diameter; 71.04 cd/m²) were presented above the fixation cross, at a distance of 12 deg. The white disks moved horizontally toward each other at 11.61 deg/s, coincided at the midpoint, and then continued moving away from each other. A beep (1850 Hz; 4 ms) was also presented; the stimulus onset asynchrony (SOA) between the visual coincidence and the beep was ±0, 25, 50, 75, 100, 150, 200, 300, 400, or 500 ms. Positive values denote that the beep was presented after the visual coincidence, and negative values denote that the beep was presented before the visual coincidence.
2.4. Procedure

At 1200–1500 ms (randomly assigned in each trial) after presentation of the fixation cross, the motion display comprising the two disks was presented. During the motion display, the beep was presented at one of the SOAs (Fig. 1). We also included a control condition, in which no beep was presented (no-beep condition). After the presentation of the motion display, the participants were asked to report, with a key press, whether the disks were perceived as streaming past one another or bouncing off each other. Each condition (the 19 SOAs and no-beep) was repeated 10 times, meaning the participants completed 200 trials in total. The presentation order of each SOA and the no-beep condition was randomized for each participant and counterbalanced among the participants.

After the experimental session, we measured the participants’ ASD traits using the Japanese version of the AQ (Fig. 2A). The AQ is a self-report questionnaire that contains 50 items describing ASD traits. Participants were asked to rate the degree to which the content of each item applied to them, using a four-point Likert scale (‘definitely agree’, ‘slightly agree’, ‘slightly disagree’, and ‘definitely disagree’). The results of the Japanese version of the AQ have previously been confirmed to be similar to those of the UK version, which suggests that the reliability and validity of the Japanese questionnaire is sufficient (Wakabayashi et al., 2004). As in previous studies (Davis et al., 2017;
Hoekstra et al., 2008; Palmer et al., 2015), we adapted Likert scoring (3-2-1-0) instead of binary scoring (1-1-0-0) to efficiently estimate individual variability. The scores for the 50 items were summed to calculate the total AQ score; we also calculated the scores of the five subscales (social skill, attention switching, communication, attention to detail, and imagination) using Likert scoring. Following the practice of Baron-Cohen et al. (2001), the 50 items were divided into the five subscales, each comprising 10 items. Scores ranged from 0–150 for the total AQ score and 0–30 for each of the five subscales; higher scores indicated a greater magnitude of ASD traits.

2.5. Data Analysis

For the stream–bounce judgments, the percentage of responses in which each participant reported perceiving a bounce was plotted against each SOA, and the respective data were fitted to a Gaussian function (Fig. 2B) using the curve-fitting toolbox in MATLAB. Then, in order to define the temporal binding window, we estimated the width (SD) of the function as in previous studies (Donohue et al., 2012; Yaguchi and Hidaka, 2018; Zampini et al., 2005). Higher width (SD) values indicate a wider temporal binding window (we refer to this index as ‘temporal binding window’ in the following). The SOA values at the peak of the function (peak SOA) were also estimated and used to represent an index of the temporal characteristics of the stream–bounce effect. The positive/negative values denote that the peak of the stream–bounce effect had shifted toward the timing where the presentation of the beep followed/preceded that of the visual coincidence. We also estimated the peak values of the function as indicating the magnitude of the stream–bounce effect; this was because previous studies involving the double-flash illusion have reported differences between people with ASD and TD individuals regarding the magnitude of audiovisual integration (Foss-Feig et al., 2010; Stevenson et al., 2014b). Higher values indicate that participants reported the stream–bounce effect more frequently. For the no-beep condition, we calculated the percentage of participants who reported perceiving a bouncing perception. Fifteen participants’ data were excluded from the analyses because their data did not fit the Gaussian function well (determination coefficient \(R^2 < 0.25\)); we also excluded eight participants’ data as outliers because one or more indices exceeded the range of two SDs from the mean. The mean AQ score of these excluded participants was 70.17 (SD = 13.20), ranging from 53 to 98.

We performed correlation analysis between the total AQ score and the estimated indices of the stream–bounce effect with an alpha level \([p = 0.017 (0.05/3)]\) corrected using the Bonferroni method. Each AQ subscale score was associated with the single component (the AQ total score), and there were some significant correlations between the AQ subscale scores, using a Bonferroni-corrected alpha level \([p = 0.005 (0.05/10)]\) (supplementary
Table S1 and supplementary Fig. S1). To detect relationships between the AQ subscale scores and the estimated indices of the stream–bounce effect while controlling these potential correlations, we performed multiple regression analyses (step down method), including all five AQ subscale scores as explanatory variables and each index of the stream–bounce effect as a target variable, with a Bonferroni-corrected alpha level [\( p = 0.017 \) (0.05/3)]. Correlations between the AQ subscale scores and each index of the stream–bounce effect are also reported for reference. In addition, we performed correlation and regression analysis for the relationships between the percentage of bouncing perceptions in the no-beep condition and the total AQ score and AQ subscales, with a 5% alpha level.

3. Results

The participants’ mean (SD) total AQ scores and scores for the five subscales based on the Likert scoring system were as follows (Fig. 2A): total AQ score: 65.42 (11.33); social skill: 13.52 (5.36); communication: 11.81 (3.89); imagination: 10.18 (3.98); attention switching: 15.87 (3.83); and attention to detail: 14.04 (4.58; see also supplementary Fig. S2 for the distribution of participants’ total AQ scores based on the binary scoring system). The participants’ mean (SD) values of the indices of the stream–bounce effect were as follows; temporal binding window: 107.10 (73.23); peak SOA: 11.22 (54.04); peak: 83.75 (12.66); and percentage of bouncing perception in the no-beep condition: 30.59 (22.78).

We conducted correlation analyses with a Bonferroni-corrected alpha level \( [p = 0.017 \) (0.0167) (0.05/3)], and no significant correlations were found between the total AQ score and the temporal binding window \( [r(66) = 0.13, p = 0.304] \), peak SOA \( [r(66) = 0.11, p = 0.392] \), or peak \( [r(66) = 0.29, p = 0.017 \) (0.0169); Fig. 3]. The correlations between each AQ subscale score and each index of the stream–bounce effect were also not significant with correction for multiple comparisons \( [p = 0.003 \) (0.05/15); supplementary Table S2 and supplementary Fig. S3]. However, our multiple regression analyses \{a Bonferroni-corrected alpha level \( [p = 0.017 \) (0.05/3)]\} and controlling the correlations among the AQ subscale scores (supplementary Table S1 and supplementary Fig. S1) revealed a significant equation between the AQ subscale scores and the peak SOA \( [F(3,64) = 4.01, R^2_{\text{adj}} = 0.12, p = 0.011] \); this showed a positive relationship for the communication score \( (\beta = 0.31, p = 0.02) \) and a negative relationship for the imagination score \( (\beta = -0.31, p = 0.02; \text{Fig. 4 and supplementary Table S3}) \). We found no significant multiple regression equations for the other indices; peak: \( [F(1,66) = 4.63, R^2_{\text{adj}} = 0.05, p = 0.04] \); temporal binding window: \( [F(1,66) = 2.00, R^2_{\text{adj}} = 0.02, p = 0.162 \) (supplementary Tables S4 and S5). For the fraction of bouncing perceptions in
Figure 3. Correlation between total AQ score and peak value (left), peak-stimulus onset asynchrony (SOA) (center), temporal binding window (right). No significant correlation was found.

Figure 4. Results of a multiple regression analysis. Significant positive and negative relationships for the communication and imagination scores, respectively, were observed when compared to the peak-stimulus onset asynchrony (SOA).

the no-beep condition, the correlations for the total AQ score \( r(66) = -0.05, p = 0.675 \) and the multiple regression equation for the AQ subscales scores \( F(1,66) = 2.02, R^2_{adj} = 0.02, p = 0.16 \) were not significant.

4. Discussion

The current study investigated the relationships between ASD traits and the perceptual characteristics of an audiovisual illusion (the stream–bounce effect) among TD people. We identified some relationships between the AQ sub-trait scores and the perceptual indices of the stream–bounce effect; in contrast, there were no significant relationships for the control condition (the no-beep condition). This indicated that our results reflected the manner of audiovisual integration, instead of responses or decisional aspects, in terms of ASD traits.

We found a positive relationship between the score for the communication subscale and the peak SOA: the participants with higher communication difficulty showed the greatest magnitude of the stream–bounce effect when the presentation of the sound followed the onset timing of the visual coincidence more. This tendency can be interpreted as atypical audiovisual integration related to ASD, because the maximum stream–bounce effect is
typically observed when the presentation timing of the sound is consistent with or prior to that of the visual coincidence (Sekuler et al., 1997). Atypical audiovisual integration of ASD has also been reported for speech stimuli. For instance, Stevenson et al. (2014a) showed that people with a diagnosis of ASD have wider temporal binding windows and a lower probability of audiovisual (speech sounds and mouse movements) integration. In addition, individuals with a diagnosis of ASD, having a wider temporal binding window, were reported to show a lower magnitude of audiovisual integration, and these characteristics can be considered to contribute to poor speech perception (Stevenson et al., 2018). These findings indicate that, for people with ASD, difficulties in communication may be due to atypical temporal processing of audiovisual integration.

It should be noted that we found a relationship between AQ subscale scores and peak SOA instead of a relationship between the former and the temporal window. This contrasts with the findings of previous studies involving the double-flash illusion, which reported a relationship with the temporal binding window (Kawakami et al., 2018; Yaguchi and Hidaka, 2018). This discrepancy may simply be explained by considering the differences in the basic characteristics of these phenomena. The stream–bounce effect is typically observed in situations during which the sound is presented at or before the visual coincidence (Sekuler et al., 1997), while the double-flash illusion occurs when sounds are presented before, simultaneously, or after the flash (Shams et al., 2002). We could speculate that the presentation timing of the sound is critical for the stream–bounce effect, while the temporal interval between the sound and visual onsets is a modulatory factor for the occurrence of the double-flash illusion. An audiovisual SJ task is also known to have a time-locked response characteristic (Zampini, et al., 2005), and the total AQ score and the AQ subscale scores were related to the points of subjective simultaneity, but not to the temporal window of the audiovisual simultaneity judgment (Donohue et al., 2012). This suggests that the relationships between ASD sub-traits and audiovisual temporal binding may depend on perceptual tasks/responses.

We should note that our results did not show a relationship between the social skill subscale and all indices of the stream–bounce effect, although the communication and social skill scores were positively correlated. These results may be due to covariance of the communication and social subscales. The current study adopted multiple regression analysis in order to take correlations between the AQ subscales into consideration, and this could reveal a unique relationship between the communication sub-trait score and an index of the stream–bounce effect. Furthermore, the items of the social skill subscale seem to be related to difficulty adjusting to novel social situations, whereas those of the communication subscale are regarded as reflecting difficulties in
communication. This suggests that the social skill and communication subscales capture different dimensions of social communication. Our earlier study found an opposite tendency between the social skill and communication subscales for the relationship with the double-flash illusion (Yaguchi and Hidaka, 2018). The double-flash illusion may involve a relatively lower form of sensory processing, because it contains only simple beeps and flashes. We could speculate that audiovisual interactions at a lower processing stage may be related to difficulties in both aspects. On the other hand, the stream–bounce effect may involve higher perceptual demands, such as motion perception and prediction. The current study demonstrated that the peak SOA of the stream–bounce effect was positively related only to the communication score. Of note, the McGurk illusion, where the presentation of mismatched visual (lip movements) and auditory (phonemes) speech cues induces an illusory phoneme perception (McGurk and MacDonald, 1976), needs the integration of visual motion information and sounds. This illusion was reported to occur frequently when the presentation of the sound followed that of the visual stimuli (Munhall et al., 1996; Van Wassenhove et al., 2007). The stream–bounce effect was also shown to occur frequently when the sound is presented simultaneously with or precedes the visual coincidence (Sekuler et al., 1997). We may assume that the temporal order of the audiovisual inputs is a key factor for audiovisual integration with higher perceptual demands. We may also note that a lower probability of the McGurk illusion was associated with poor speech perception in ASD (Stevenson et al., 2018). We could speculate that, as in the McGurk illusion, the atypical characteristics of audiovisual temporal integration in the stream–bounce effect, which involves higher perceptual demands such as motion perception and prediction, is specifically related to the AQ communication subscale score, which is assessed by items related to verbal communications (e.g., ‘Other people frequently tell me that what I’ve said is impolite, even though I think it is polite’). This possibility should be examined in future studies.

We also found a negative relationship between the imagination subscale score and peak SOA, which was opposite to the result for the communication score: the participants with greater imagination difficulty showed the greatest magnitude of the stream–bounce effect when the presentation of the beep preceded the onset timing of the visual coincidence more. The items in the imagination subscale are assumed to reflect difficulties in developing imagination (e.g., ‘I find it difficult to imagine what it would be like to be someone else’ and ‘When I’m reading a story, I find it difficult to work out the characters’ intentions’), whereas the communication subscale reflects difficulties in communication. We could also note that there was no significant correlation between the imagination and communication scores (supplementary Fig. S1).
We speculated that the imagination and communication subscales are associated with different dimensions of ASD characteristics, and these dimensions could be uniquely linked with audiovisual integration. We assumed that the greater negative peak SOA in the stream–bounce effect would indicate that, for temporally integrating sounds and visual motion, the sounds were required to precede the visual coincidence more. Our findings may indicate that more time may be necessary to integrate information for people with greater difficulties in using their imagination.

Our results further showed that the indices of the stream–bounce effect were not related to AQ subscales regarding attention (attention switching and attention to detail). It has been reported that a higher magnitude of audiovisual perceptual learning for motion is associated with higher AQ attention switching subscale scores (Hidaka and Yaguchi, 2018) and higher scores for attention to detail were related to a lower magnitude of audiovisual temporal adaptation (Stevenson et al., 2017). We can speculate that the attentional aspects of ASD may be mainly related to acquiring novel audiovisual relationships, so that these aspects have no relationship to the indices of the stream–bounce effect.

We found no significant relationships between the total AQ score and the indices of the stream–bounce effect, although participants with greater ASD traits tended to report a greater magnitude of the effect \( p = 0.0169 \) against the Bonferroni-corrected alpha level \( p = 0.0167 \) and this seemed to be consistent with the findings of a previous study in which people with a diagnosis of ASD showed a greater magnitude of the double-flash illusion (Foss-Feig et al., 2010). There are inconsistencies among previous studies comparing the manner of audiovisual interactions between people with and without a diagnosis of ASD (Foss-Feig et al., 2010; Keane et al., 2010; Marco et al., 2011; Stevenson et al., 2014a; van der Smagt et al., 2007). We consider that one plausible reason for these discrepancies is that these studies did not consider the fact that ASD characteristics comprise various aspects. In fact, studies have reported unique relationships between AQ subscales and audiovisual interactions (e.g., Donohue et al., 2012; Kawakami et al., 2018; Yaguchi and Hidaka, 2018). Based on these, we could assume that the current study did not find significant relationships between the total AQ score and the indices of audiovisual interactions because of the unique relationships that were observed between the AQ subscales and the indices.

It has been highlighted that, for people with a diagnosis of ASD, atypical audiovisual characteristics are related to social communication difficulties (Beker et al., 2018; Stevenson et al., 2014a). While our findings are limited to a TD population, our results regarding the stream–bounce effect suggest that, in accordance with previous findings relating to the double-flash illusion.
(Kawakami et al., 2018; Yaguchi and Hidaka, 2018), atypical audiovisual perception in a temporal domain may underlie the altered features of social communication. Researchers have proposed audiovisual integration as a potential target for interventions to improve outcomes regarding high-order social communication in children with ASD (Cascio et al., 2016). There is some evidence that audiovisual integration is malleable with training (Feldman et al., 2018; Powers et al., 2009, 2016). Our findings suggest that it may be important to focus on separate relationships, especially those between ASD-related difficulties in communication and temporal aspects of audiovisual integration, in such interventions.

**Acknowledgements**

This research was supported by a Grant-in-Aid for Scientific Research (C) (No. 17K00214) and a Grant-in-Aid for JSPS Fellows (No.18J22022) from the Japan Society for the Promotion of Science.

**Supplementary Material**

Supplementary material is available online at: https://doi.org/10.6084/m9.figshare.13280600

**References**

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Association, Arlington, VA, USA.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. and Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, male and females, scientists and mathematicians, *J. Autism Dev. Disord.* 31, 5–17. DOI:10.1023/a:1005653411471.

Baum, S. H., Stevenson, R. A. and Wallace, M. T. (2015). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder, *Prog. Neurobiol.* 134, 140–160. DOI:10.1016/j.pneurobio.2015.09.007.

Beker, S., Foxe, J. J. and Molholm, S. (2018). Ripe for solution: delayed development of multisensory processing in autism and its remediation, *Neurosci. Biobehav. Rev.* 84, 182–192. DOI:10.1016/j.neubiorev.2017.11.008.

Brainard, D. H. (1997). The psychophysics toolbox, *Spat. Vis.* 10, 433–436.

Cascio, C. J., Woynaroski, T., Baranek, G. T. and Wallace, M. T. (2016). Toward an interdisciplinary approach to understanding sensory function in autism spectrum disorder, *Autism Res.* 9, 920–925. DOI:10.1002/aur.1612.

Conrey, B. and Pisoni, D. B. (2006). Auditory–visual speech perception and synchrony detection for speech and nonspeech signals, *J. Acoust. Soc. Am.* 119, 4065–4073. DOI:10.1121/1.2195091.
Davis, J., McKone, E., Zirnsak, M., Moore, T., O’Kearney, R., Apthorp, D. and Palermo, R. (2017). Social and attention-to-detail subclusters of autistic traits differentially predict looking at eyes and face identity recognition ability, *Br. J. Psychol.* **108**, 191–219. DOI:10.1111/bjop.12188.

de Boer-Schellekens, L., Eussen, M. and Vroomen, J. (2013). Diminished sensitivity of audio-visual temporal order in autism spectrum disorder, *Front. Integr. Neurosci.* **7**, 8. DOI:10.3389/fnint.2013.00008.

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

Donohue, S. E., Darling, E. F. and Mitroff, S. R. (2012). Links between multisensory processing and autism, *Exp. Brain Res.* **222**, 377–387. DOI:10.1007/s00221-012-3223-4.

Ernst, M. O. and Bülthoff, H. H. (2004). Merging the senses into a robust percept, *Trends Cogn. Sci.* **8**, 162–169. DOI:10.1016/j.tics.2004.02.002.

Feldman, J. I., Dunham, K., Cassidy, M., Wallace, M. T., Liu, Y. and Woynaroski, T. G. (2018). Audiovisual multisensory integration in individuals with autism spectrum disorder: a systematic review and meta-analysis, *Neurosci. Biobehav. Rev.* **95**, 220–234. DOI:10.1016/j.neubiorev.2018.09.020.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L. and Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders, *Exp. Brain Res.* **203**, 381–389. DOI:10.1007/s00221-010-2240-4.

Hidaka, S. and Yaguchi, A. (2018). An investigation of the relationships between autistic traits and crossmodal correspondences in typically developing adults, *Multisens. Res.* **31**, 729–751. DOI:10.1163/22134808-20181304.

Hoekstra, R. A., Bartels, M., Cath, D. C. and Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups, *J. Autism Dev. Disord.* **38**, 1555–1566. DOI:10.1007/s10803-008-0538-x.

Jeste, S. S. and Nelson, C. A. 3rd (2009). Event related potentials in the understanding of autism spectrum disorders: an analytical review, *J. Autism Dev. Disord.* **39**, 495–510. DOI:10.1007/s10803-008-0652-9.

Kawakami, S., Uono, S., Otsuka, S., Zhao, S. and Toichi, M. (2018). Everything has its time: narrow temporal windows are associated with high levels of autistic traits via weaknesses in multisensory integration, *J. Autism Dev. Disord.* **50**, 1–11. DOI:10.1007/s10803-018-3762-z.

Keane, B. P., Rosenthal, O., Chun, N. H. and Shams, L. (2010). Audiovisual integration in high functioning adults with autism, *Res. Autism Spectr. Disord.* **4**, 276–289. DOI:10.1016/j.rasd.2009.09.015.

Leekam, S. R., Prior, M. R. and Ulijarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders: a review of research in the last decade, *Psychol. Bull.* **137**, 562–593. DOI:10.1037/a0023341.

Marco, E. J., Hinkley, L. B. N., Hill, S. S. and Nagarajan, S. S. (2011). Sensory processing in autism: a review of neurophysiologic findings, *Pediatr. Res.* **69**, 48R–54R. DOI:10.1203/PDR.0b013e3182130c54.

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices, *Nature* **264**, 746–748. DOI:10.1038/264746a0.
Munhall, K. G., Gribble, P., Sacco, L. and Ward, M. (1996). Temporal constraints on the McGurk effect, *Percept. Psychophys.* 58, 351–362. DOI:10.3758/bf03206811.

Noel, J.-P., De Niear, M. A., Stevenson, R., Alais, D. and Wallace, M. T. (2017). Atypical rapid audio-visual temporal recalibration in autism spectrum disorders, *Autism Res.* 10, 121–129. DOI:10.1002/aur.1633.

Palmer, C. J., Paton, B., Enticott, P. G. and Hohwy, J. (2015). ‘Subtypes’ in the presentation of autistic traits in the general adult population, *J. Autism Dev. Disord.* 45, 1291–1301. DOI:10.1007/s10803-014-2289-1.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies, *Spat. Vis.* 10, 437–442.

Powers, A. R. 3rd, Hillock, A. R. and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding, *J. Neurosci.* 29, 12265–12274. DOI:10.1523/JNEUROSCI.3501-09.2009.

Powers, A. R. III, Hillock-Dunn, A. and Wallace, M. T. (2016). Generalization of multisensory perceptual learning, *Sci. Rep.* 6, 23374. DOI:10.1038/srep23374.

Sekuler, R., Sekuler, A. B. and Lau, R. (1997). Sound alters visual motion perception, *Nature* 385, 308. DOI:10.1038/385308a0.

Shams, L., Kamitani, Y. and Shimojo, S. (2002). Visual illusion induced by sound, *Cogn. Brain Res.* 14, 147–152. DOI:10.1016/s0926-6410(02)00069-1.

Spence, C. and Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony, *Curr. Biol.* 13, R519–R521. DOI:10.1016/s0960-9822(03)00445-7.

Stevenson, R. A., Zemtsov, R. K. and Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions, *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1517–1529. DOI:10.1037/a0027339.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M. and Wallace, M. T. (2014a). Multisensory temporal integration in autism spectrum disorders, *J. Neurosci.* 34, 691–697. DOI:10.1523/JNEUROSCI.3615-13.2014.

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M. and Wallace, M. T. (2014b). Evidence for diminished multisensory integration in autism spectrum disorders, *J. Autism Dev. Disord.* 44, 3161–3167. DOI:10.1007/s10803-014-2179-6.

Stevenson, R. A., Toulmin, J. K., Youm, A., Besney, R. M. A., Schulz, S. E., Barense, M. D. and Ferber, S. (2017). Increases in the autistic trait of attention to detail are associated with decreased multisensory temporal adaptation, *Sci. Rep.* 7, 14354. DOI:10.1038/s41598-017-14632-1.

Stevenson, R. A., Segers, M., Ncube, B. L., Black, K. R., Bebko, J. M., Ferber, S. and Barense, M. D. (2018). The cascading influence of multisensory processing on speech perception in autism, *Autism* 22, 609–624. DOI:10.1177/1362361317704413.

Tomchek, S. D. and Dunn, W. (2007). Sensory processing in children with and without autism: a comparative study using the short sensory profile, *Am. J. Occup. Ther.* 61, 190–200. DOI:10.5014/ajot.61.2.190.

Turi, M., Karaminis, T., Pellicano, E. and Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum, *Sci. Rep.* 6, 21756. DOI:10.1038/srep21756.
van der Smagt, M. J., van Engeland, H. and Kemner, C. (2007). Brief report: can you see what is not there? Low-level auditory-visual integration in autism spectrum disorder, *J. Autism Dev. Disord.* 37, 2014–2019. DOI:10.1007/s10803-006-0346-0.

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

Wakabayashi, A., Tojo, Y., Baron-Cohen, S. and Wheelwright, S. (2004). (The Autism–Spectrum Quotient (AQ) Japanese version: evidence from high-functioning clinical group and normal adults), *Shinrigaku Kenkyu* (Jpn. J. Psychol.) 75, 78–84. DOI:10.4992/jjpsy.75.78.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S. and Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: a preliminary study of its diagnostic validity in clinical practice, *J. Autism Dev. Disord.* 35, 331–335. DOI:10.1007/s10803-005-3300-7.

Yaguchi, A. and Hidaka, S. (2018). Distinct autistic traits are differentially associated with the width of the multisensory temporal binding window, *Multisens. Res.* 31, 523–536. DOI:10.1163/22134808-00002612.

Zampini, M., Guest, S., Shore, D. I. and Spence, C. (2005). Audio-visual simultaneity judgments, *Percept. Psychophys.* 67, 531–544. DOI:10.3758/bf03193329.

Zhou, H.-Y., Cai, X.-L., Weigl, M., Bang, P., Cheung, E. F. C. and Chan, R. C. K. (2018). Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: a systematic review and meta-analysis, *Neurosci. Biobehav. Rev.* 86, 66–76. DOI:10.1016/j.neubiorev.2017.12.013.