Received: 14 January 2022

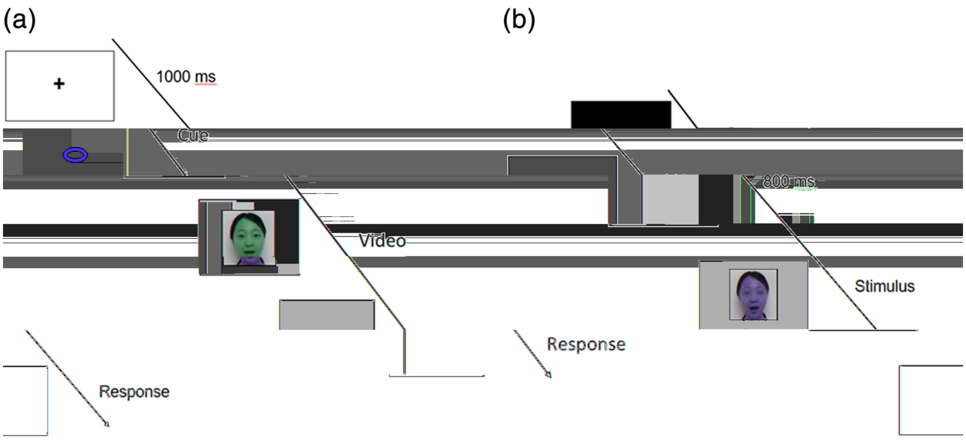


FIGURE 1 Procedures of a sample trial in Experiment 1 (a) and Experiment 2 (b). *Note*. (a) This figure shows the procedure of a trial in Experiment 1. First, a fixation was presented at the center of the screen for 1000 ms. Then, a black screen was displayed for 800 ms. Subsequently, the stimulus was presented. Finally, a black screen was shown, and the children responded. (b) This figure shows the procedure of a sample trial in the cue-to-mouth condition in Experiment 2. First, a black screen with an oval at the position where the speaker's mouth appeared was presented. Then, the stimulus was presented once the children kept fixating on the oval area for 500 ms. Finally, a black screen was displayed until the children responded

The stimuli were displayed at the center of the screen using MATLAB (The MathWorks, Natick, MA) and Psychotoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Sounds were presented through two speakers located on the two sides of the screen. Children were required to perform the McGurk task by reporting what the speaker said, and their eye movements were recorded using a Tobii X 120 eye tracker (sampling rate: 120 Hz).

Children performed a practice session to familiarize themselves with the McGurk task before the formal experiment. At the beginning of the formal experiment, the children's eye movements were calibrated using Tobii's five-point calibration method. The calibration was accepted only when all five points showed a good fit, with error vectors smaller than 0.5 degree of the visual angle. As mentioned above, the experiment consisted of a clear-eyes condition and a blurred-eyes condition. Each condition included four trials of congruent "ba," four trials of congruent "ga," and 12 trials of incongruent "AbVg" (auditory "ba" + visual "ga"). Each trial began with a black fixation at the center of the screen for 1000 ms, and children were asked to look at it. Then, a black screen was displayed for 800 ms. Subsequently, the stimulus was presented. Finally, a black screen was displayed until the children responded. Children's responses were recorded by the experimenter, that is, by pressing "b," "d," and "g" on the keyboard for responses of "ba," "da," and "ga" respectively. For a sample trial procedure, please refer to Figure 1a. The 20 trials in each condition were presented in random order, and the order of the two conditions was counterbalanced among children. Children took rest between the conditions. The experiment lasted for approximately 25 min.

2.1.4 | Data analysis

Eye movement data analysis. We defined five areas of interest (AOIs) for the speaker's face: the whole face, the eyes (left eye and right eye), the mouth, the nose, and the other area (the area on the face excluding eyes, nose, and mouth; see Figure 2). We extracted fixations from the raw gaze data, as specified by Tobii (I-VT fixation filter; Olsen, 2012). In particular, the minimum fixation duration was set at 60 ms within a velocity of 30 deg/s. Then, we obtained the fixation data, which included the onset, the offset, and the position (x-coordinate, y-coordinate in pixels) of each fixation. For each trial, we extracted the fixation data during the time the video was displayed on the screen

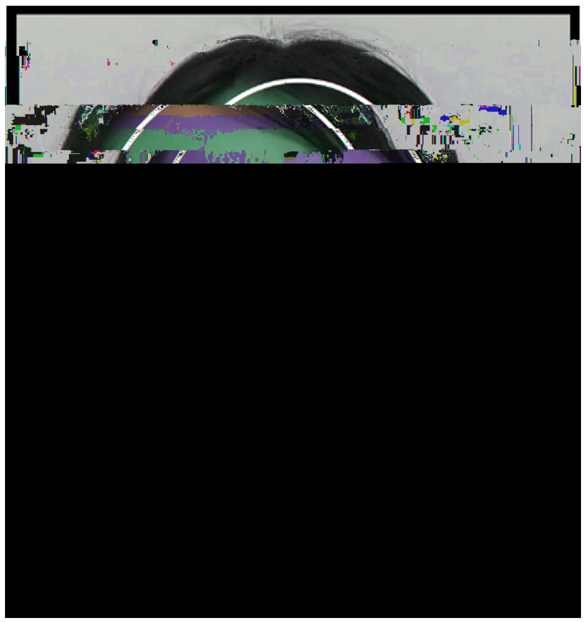


FIGURE 2 Sample AOIs used in the eye movement data analysis. *Note*. This figure shows the five AOIs in the eye-movement data analysis. The five AOIs included the whole face, eyes (left eye and right eye), mouth, nose, and other areas (the area on the face excluding eyes, nose, and mouth)

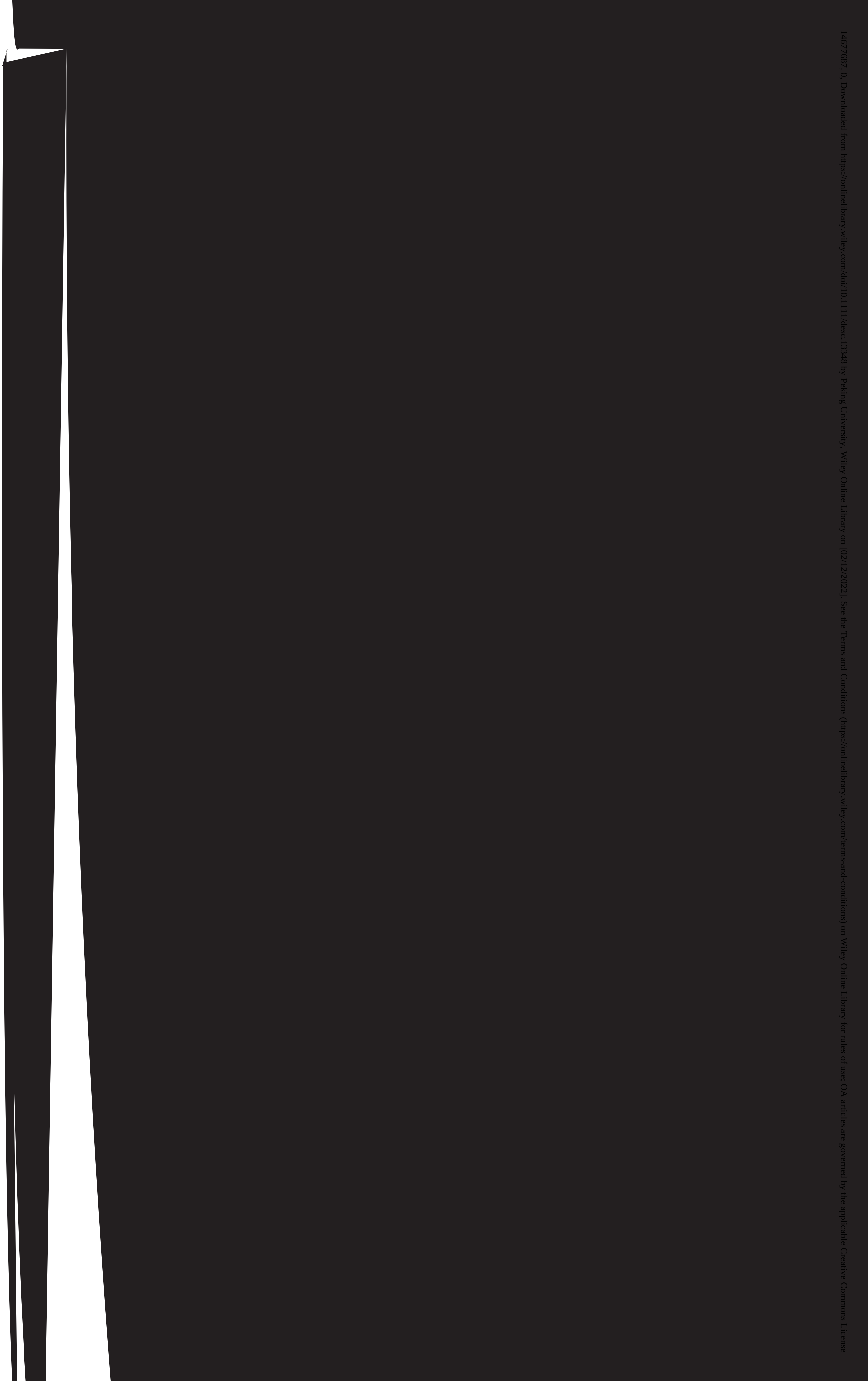
(i.e., from the time point that the video appeared on the screen to the time point that the video disappeared on the screen) and calculated the duration of each fixation by using the offset to minus the onset. After that, we selected fixations within each AOI and summed their durations separately, obtaining the total looking time on each AOI. Finally, we calculated the average total looking time on each AOI for each participant and for each group. We chose looking time as the dependent variable by referring to Gurler et al. (2015). In this study, Gurler et al. used looking time as the dependent variable and found that mouth-looking time was positively correlated with McGurk effect. Moreover, looking time was widely used in previous studies to reflect participants' attention to a specific AOI (e.g., Chawarska & Shic, 2009; Tsang et al., 2022).

Behavioral data analysis. We analyzed the incongruent trials and excluded congruent trials as they were used as filler trials. For the incongruent trials, children made three types of responses: auditory responses "ba," visual responses "ga," and fused responses "da" (McGurk response). By referring to Stevenson et al. (2014), we took the fused response "da" as the McGurk percept. We computed children's percentages of each type of response in both conditions. We conducted the following analyses using non-parametric statistical analyses (i.e., repeated measures permutation ANOVA, Wilcoxon signed ranks tests, Mann-Whitney *U*-test) as the data violated the normal distribution.

2.2 | Results

2.2.1 | Blurring eyes decreased eyes-looking time and increased mouth-looking time

To explore whether blurring eyes could change looking time in the two groups on the speaker's eyes and mouth, we conducted a $2 \times 2 \times 2$ repeated measures ANOVA on looking time with Condition (clear-eyes vs. blurred-eyes) and Region (eyes vs. mouth) as the within-subject factors, and Group (AC vs. NAC) as the between-subject factor using the R package "bruceR." We found a significant main effect of Condition, $F(1, 58) = 4.23$, $p = 0.04$, $\eta_p^2 = 0.07$, a significant main effect of Region, $F(1, 58) = 24.07$, $p < 0.001$, $\eta_p^2 = 0.29$, and a significant main effect of Group, $F(1, 58) = 13.13$, $p = 0.001$, $\eta_p^2 = 0.19$. It also showed a significant Region × 175.758273.969Tm[(=)]TJegion



until the children kept viewing the black screen for at least 500 ms. Trials in each condition were presented in random order, and the order of the three conditions was counterbalanced among children. Children took rest after completing a condition. The experiment lasted for approximately 35 min.

3.1.3 | Data analysis

As for the eye movement data, we also computed children's *total looking time* on each AOI by using identical procedures to Experiment 1. As for the behavioral data, children also made three types of responses in Experiment 1. We also first computed children's percentages of each kind of response in each condition. We then performed non-parametric analyses as the data did not conform to the normal distribution.

3.2 | Results

3.2.1 | Cuing to mouth increased the mouth-looking time and cuing to eyes increased the eyes-looking time

To explore whether cuing to the mouth or eyes could change looking time in the two groups on the eyes and the mouth, we conducted a 3 × 2 × 2 repeated measures ANOVA on looking time with Condition (cue-to-mouth vs. cue-to-eyes vs. free-viewing) and Region (eyes vs. mouth) as the within-subject factors, and Group (AC vs. NAC) as the between-subject factor using the R package "bruceR." Results showed a significant main effect of Condition, $F(2, 160) = 3.63$, $p = 0.03$, $\eta_p{}^2 = 0.04$, a significant main effect of Region, $F(1, 80) = 40.93$, $p < 0.001$, $\eta_p{}^2 = 0.34$, and a significant main effect of Group, $F(1, 80) = 4.83$, $p = 0.03$, $\eta_p{}^2 = 0.06$. Results also showed a significant Condition × Region interaction, $F(2, 160) = 38.91$, $p < 0.001$, $\eta_p{}^2 = 0.33$. None of the Condition × Group interaction, $F(2, 160) = 1.71$, $p = 0.18$, $\eta_p{}^2 = 0.02$, the Region × Group interaction, $F(1, 80) = 0.43$, $p = 0.51$, $\eta_p{}^2 = 0.005$, or the Condition × Region × Group interaction, $F(2, 160) = 0.94$, $p = 0.39$, $\eta_p{}^2 = 0.01$, was significant.

For the significant Condition × Region interaction, we further conducted simple analyses to examine the condition difference of children's looking time on the eyes and the mouth. Results showed that children's looking time was significantly different among the three conditions for both the eyes, $F(2, 80) = 30.60$, $p < 0.001$, $\eta_p{}^2 = 0.43$, and the mouth region, $F(2, 80) = 19.29$, $p < 0.001$, $\eta_p{}^2 = 0.33$. Multiple comparisons showed that children spent longer time viewing the eyes in the cue-to-eyes condition than in the cue-to-mouth condition, $t(80) = 7.82$, $p < 0.001$, Cohen's $d = 0.46$, in the cue-to-eyes condition than in the free-viewing condition, $t(80) = 3.46$, $p = 0.003$, Cohen's $d = 0.18$, in the free-viewing condition than in the cue-to-mouth, $t(80) = 4.79$, $p < 0.001$, Cohen's $d = 0.28$ (all *p*s were corrected by Bonferroni corrections; Figure 5a). Multiple comparisons also showed that children spent longer time viewing the mouth in the cue-to-mouth condition than in the cue-to-eyes condition, $t(80) = 6.11$, $p < 0.001$, Cohen's $d = 0.45$, and in the cue-to-mouth condition than in the free-viewing condition, $t(80) = 5.39$, $p < 0.001$, Cohen's $d = 0.43$. In addition, children spent similar time viewing the mouth in the cue-to-eyes condition and free-viewing condition, $t(80) = 0.48$, $p > 0.99$, Cohen's $d = 0.03$ (all *p*s were corrected by Bonferroni correction; Figure 5b). We also explored whether cuing to the mouth or eyes could change children's looking time on the nose and the other area. Results only showed a significant effect of condition for each of both areas, please see the Supplemental materials for detail (Figure S3).

3.2.2 | Cuing to mouth enhanced the McGurk effect in autism

We tested the group differences of the three types of responses in the three conditions separately and found that the autistic group showed less McGurk effect than the nonautistic group in all three conditions (see Figure S4 in the Supplemental materials for detailed information).

We performed a two-way repeated measures permutation ANOVA to test the condition and group differences of the McGurk effect ("da" response) using the *R* package "permuco" default method (Frossard & Renaud, 2019). The results showed a significant main effect of Group, $F(1, 80) = 10.00$, permutation $p = 0.003$, $\eta_p{}^2 = 0.11$, a significant main effect of Condition, $F(2, 160) = 7.51$, permutation $p = 0.0004$, $\eta_p{}^2 = 0.09$, and a significant Group × Condition interaction, $F(2, 160) = 3.61$, permutation $p = 0.03$, $\eta_p{}^2 = 0.04$ (Figure 6). We further examined the condition differences of the McGurk effect in each group. We found that the autistic group showed a stronger McGurk effect in the cue-to-mouth condition than in the cue-to-eyes condition, $Z = 3.33$, $p = 0.003$, $r = 0.53$, and in the cue-to-mouth condition than in the free-viewing condition, $Z = 3.15$, $p = 0.003$, $r = 0.50$, and showed a similar McGurk effect in the cue-to-eyes condition and the free-viewing condition, $Z = 1.70$, $p = 0.09$, $r = 0.27$, by conducting Wilcoxon signed-rank tests (all *p*s were corrected by FDR correction; Figure 6). However, we found that the nonautistic group showed similar McGurk effects in three conditions, $F(1, 80) = 0.64$, permutation $p = 0.53$, $\eta_p{}^2 = 0.02$, by performing a one-way repeated measures permutation ANOVA. That is, the McGurk effect in the autistic group increased in the cue-to-mouth condition compared with the other two conditions, but the McGurk effect in the nonautistic group did not differ in the three conditions.

To examine the potential practice effects for those participants who were in both experiments, we conducted a Wilcoxon signed-rank tests to examine the McGurk effect difference in the 26 nonautistic and 25 AC who participated in both Experiments 1 and 2. Particularly, we compared the McGurk effect in the baseline conditions in two experiments (i.e., clear-eyes condition in Experiment 1 and free-viewing condition in Experiment 2) for each group respectively. Results showed that the nonautistic group was similar in the percentages of McGurk effect in the two experiments, $z = 0.72$, $p = 0.47$, $r = 0.18$, indicating no practice effect in the nonautistic group. The autistic group, however, showed a larger McGurk effect in Experiment 2 than in Experiment 1, $z = 3.34$, $p = 0.001$, $r = 0.67$, indicating the existence of practice effects in the autistic group.

explanation is that face-looking patterns change with age; NAC spent more time viewing the mouth compared to adults (Nakano et al., 2010). For NAC, whose mouth-looking time is relatively long, it might be difficult for them to increase their audiovisual speech integration in the McGurk task by increasing their mouth-looking time (Nakano et al., 2010). For AC, who showed less mouth-looking time compared with NAC, their audiovisual speech integration in the McGurk task might be enhanced by increasing mouth-looking time (Nakano et al., 2010; Feng et al., 2021). NAC's audiovisual speech integration could depend on other factors, such as neural development in brain areas such as the superior temporal gyrus and inferior frontal gyrus (Jones & Callan, 2003; Tryfon et al., 2018). Future research should explore the potential influence of these factors on audiovisual speech integration in both NAC and AC.

Our finding that increasing mouth-looking time in AC could enhance their performance in McGurk task should be considered with caution. As the McGurk effect is only an indicator of audiovisual speech integration, the changes of McGurk performance in AC does not necessarily mean that there were changes in their general audiovisual speech integration ability. Improving the ability of general audiovisual speech integration in autism does require a relatively long time of professional supports. In future studies, we could conduct trainings that last for several weeks or months and examine whether changing face attention in AC could increase their general audiovisual speech integration. We could also further examine whether such supports could have cascading effects on their language development and communication abilities (Feldman et al., 2018; Lewkowicz & Hansen-Tift, 2012; Righi et al., 2018; Tenenbaum, et al., 2017; Young et al., 2009).

The present study revealed group and condition differences in McGurk percept. One might argue that these differences could reflect a decision bias rather than truly reflecting perceptual processing. We propose that a genuine audiovisual perceptual integration had taken place based on the following considerations. First, perceptual processing entails earlier, distinct interaction of audiovisual events before these events could be potentially handled by a later decision mechanism (Mercier & Cappe, 2022; van Wassenhove et al., 2005). In the McGurk task, previous studies have shown the audiovisual interaction in integration starts as early as 100 ms after the onset of audiovisual stimuli, which is characterized by the N1/P2 amplitude reductions (van Wassenhove et al., 2005). Second, the McGurk effect demonstrated in the two groups was comparable to previous studies (e.g., Feng et al., 2021), indicating that the McGurk effect was stable across different studies (e.g., Feng et al., 2021). Third, for the research method, we adopted a trial-by-trial, fully-randomized arrangement of the stimuli presentations in each condition. This arrangement largely prevented the participants from purposely adopting specific strategies (including potential expectations for a certain pattern of responses) in the present study. In sum, we believe both the research protocol and time window of cross-modal interaction in the present task favored an account of genuine perceptual processing.

We also found that there existed some practice effects in AC but not in NAC. The lack of practice effects in the nonautistic group might be explained by their relatively high performance of McGurk effect, which is difficult to enhance. The practice effect in the autistic group implicates the possibility of improving their McGurk effect through relevant supports by taking advantage of the practice effects.

The present study found that audiovisual speech integration in the McGurk task could be increased in AC by changing their mouth-looking time. This finding has several theoretical and practical implications. Theoretically, it implies that attentional allocation (i.e., spent less time viewing core features, like the mouth) in AC is one of the mechanisms underlying the less audiovisual speech integration in the McGurk task in AC. In addition, audiovisual speech integration in the McGurk task in AC could be enhanced but still could not catch up with that in NAC, which implies that the less audiovisual speech integration in AC could also be explained by other factors, such as their difficulties in central coherence (Happé & Frith, 2006), temporal processing (Stevenson et al., 2014), and predictive coding (Baum et al., 2015). In future studies, we could explore the factors underlying reduced audiovisual speech integration in AC. Practically, it implies that audiovisual speech integration supports could be carried out in AC, especially those at an early age, by taking measures to increase their mouth-looking time. Further, our findings indicate that blurring and cuing could effectively manipulate children's attention. In future studies, we could employ these measures to manipulate children's attention, especially in the support for AC.

The present study has some limitations. One limitation is that our participants were all boys, and our findings were only confined to boys. In future studies, we could recruit a group of autistic girls to explore the potential gender differences and further explore whether blurring the speaker's eyes or cuing to the speaker's mouth could enhance audiovisual speech integration in the McGurk task in autistic girls. Another limitation is that our participants were 4- to 8-year-old children who had developed a certain level of language ability. Blurring the speaker's eyes or cuing to the speaker's mouth might be more effective in improving audiovisual speech integration in the McGurk task in infants or toddlers who are in the initial stages of language development. In future studies, infants or toddlers could be recruited for further exploration. In addition, our participants have been diagnosed in licensed hospitals by professional pediatricians according to the criteria of the DSM-V. Although we have acknowledged the importance of including Autism Diagnostic Observation Schedule (ADOS), we did not do so given that the official Chinese version of ADOS has not yet been translated and published officially, and the reliable administrators are highly limited in China. Using this gold standard to confirm the diagnosis of autism is recommended for future studies. Last, we included only 12 incongruent trials for each condition of the two experiments considering AC's reduced cooperation and sustained attention to complete the task. The limited number of trials might make the task not sensitive enough to detect the differences between condition and groups. Future studies could include more trials to increase the sensitivity of the McGurk task.

In summary, the present study increased the mouth-looking time by blurring the eyes and cueing to the mouth, and these manipulations could increase audiovisual speech integration in the McGurk task in AC, but not in NAC. These findings contribute to a deeper understanding of the underlying mechanisms of audiovisual speech integration in autism. This finding could also provide insights for the development of supports to increase audiovisual speech integration in AC.

ACKNOWLEDGMENTS

The authors are grateful to Zipeng Ma, Fuli Liu, Yanhong Wu, Yinan Lv, and the staff in Qingdao Elim School, for their generous assistance in completing the study. This work was supported by Key-Area Research and Development Program of Guangdong Province (2019B030335001), National Natural Science Foundation of China (31871116, 62061136001, 31861133012, 32271116), Philosophy and Social Science Foundation of Hunan Province (19YBQ109), Outstanding Youth Foundation of Education Department of Hunan Province (21B0004), and Natural Science Foundation of Hunan Province of China (2022JJ40580).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data were available upon reasonable request.

ETHICS APPROVAL STATEMENT

The experiment was performed in compliance with the institutional guidelines set by the Ethics Committee of School of Psychological and Cognitive Sciences, Peking University, China, in accordance to the 1975 Declaration of Helsinki concerning human and animal rights.

REFERENCES

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839–843. https://doi.org/10.1016/j.cub.2005.03.046

Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, *183*, 399–404. https://doi.org/10.1007/s00221-007-1110-1

Altieri, N., Pisoni, D. B., & Townsend, J. T. (2011). Some behavioral and neurobiological constraints on theories of audiovisual speech integration: A review and suggestions for new directions. *Seeing and Perceiving*, *24*(6), 513–539. https://doi.org/10.3389/fpsyg.2014.00257

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th edn.). American Psychiatric Press.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism spectrum quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders*, *38*, 1230–1240. https://doi.org/10.1007/s10803-007-0504-z

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *38*, 813–822. https://doi.org/10.1111/j.1469-7610.1997.tb01599.x

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

Bebko, J. M., Schroeder, J. H., & Weiss, J. A. (2014). The McGurk effect in children with autism and Asperger syndrome. *Autism Research*, *7*, 50–59. https://doi.org/10.1002/aur.1343

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/10.1163/156856897x00357

Chawarska, K., & Shic, F. (2009). Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. *Journal of Autism Developmental Disorder*, *39*, 1663–1672. https://doi.org/10.1007/s10803-009-0803-7

de Wit, T. C. J., Falck-Ytter, T., & von Hofsten, C. (2008). Young children with autism spectrum disorder look differently at positive versus negative emotional faces. *Research in Autism Spectrum Disorders*, *2*, 651–659. https://doi.org/10.1016/j.rasd.2008.01.004

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

Feng, S., Lu, H., Wang, Q., Li, T., Fang, J., Chen, L., & Yi, L. (2021). Face-viewing patterns predict audiovisual speech integration in autistic children. *Autism Research*, *14*, 2592–2602. https://doi.org/10.1002/aur.2598

Frossard, J., & Renaud, O. (2019). permuco: Permutation tests for regression, (repeated measures) ANOVA/ANCOVA and comparison of signals. R package version 1.1.0. https://CRAN.R-project.org/package=permuco

Grynszpan, O., Simonin, J., Martin, J. C., & Nadel, J. (2012). Investigating social gaze as an action-perception online performance. *Frontiers in Human Neuroscience*, *6*(6), 94. https://doi.org/10.3389/fnhum.2012.00094

Gurler, D., Doyle, N., Walker, E., Magnotti, J., & Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye movements. *Attention, Perception, & Psychophysics*, *77*, 1333–1341. https://doi.org/10.3758/s13414-014-0821-1

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Iarocci, G., Rombough, A., Yager, J., Weeks, D. J., & Chua, R. (2010). Visual influences on speech perception in children with autism. *Autism*, *14*, 305–320. https://doi.org/10.1177/1362361309353615

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

Jones, J., & Callan, D. E. (2003). Brain activity during audiovisual speech perception: An fMRI study of the McGurk effect. *NeuroReport*, *14*(8), 1129–1133.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, ECVP Abstract Supplement.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1431–1436.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. https://doi.org/10.1023/A:1005592401947

Mercier, M. R., & Cappe, C. (2022). The interplay between multisensory integration and perceptual decision making. *NeuroImage*, *222*, 116970.

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 Developmental Disorder*, *38*, 1349–1358. https://doi.org/10.1007/s10803-007-0521-y

Moriuchi, J. M., Klin, A., & Jones, W. (2017). Mechanisms of diminished attention to eyes in autism. *American Journal of Psychiatry*, *174*(1), 26–35. https://doi.org/10.1176/appi.ajp.2016.15091222

Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., Ohta, H., Kato, N., & Kitazawa, S. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proceedings of the Royal Society B*, *277*, 2935–2943. https://doi.org/10.1098/rspb.2010.0587

Newman, R. S., Kirby, L. A., Von Holzen, K., & Redcay, E. (2021). Read my lips: Perception of speech in noise by preschool children with autism and the impact of watching the speaker's face. *Journal of Neurodevelopmental Disorders*, *13*, online ahead of print.

Olsen, A. (2012). The Tobii I-VT fixation filter: Algorithm Description. Tobii I-VT Fixation Filter-white paper [white paper]. Retrieved from: http://www.tobii.com/no/eye-tracking-research/global/library/white-papers/the-tobii-i-vt-fixation-filter/

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. https://doi.org/10.1163/156856897x00366

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*, 249–261. https://doi.org/10.1016/j.tics.2012.08.009

Righi, G., Tenenbaum, E. J., Mccormick, C., Blossom, M., & Sheinkopf, S. J. (2018). Sensitivity to audio-visual synchrony and its relation to language abilities in children with and without ASD. *Autism Research*, *11*, 645–653. https://doi.org/10.1002/aur.1918

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL http://www.R-project.org/

Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry*, *75*(3), 231–237. https://doi.org/10.1016/j.biopsych.2013.07.009

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. *Journal of Neuroscience*, *34*, 691–697. https://doi.org/10.1523/JNEUROSCI.3615-13.2014

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*, 400–410. https://doi.org/10.1016/j.tics.2010.06.008

Tenenbaum, E. J., Amso, D., Righi, G., & Sheinkopf, S. (2017). Attempting to "increase intake from the input": Attention and word learning in children with autism. *Journal of Autism and Developmental Disorders*, *47*, 1791–1805. https://doi.org/10.1007/s10803-017-3098-0

Tiippana, K., Andersen, T. S., & Sams, M. (2004). Visual attention modulates audiovisual speech perception. *European Journal of Cognitive Psychology*, *16*(3), 457–472. https://doi.org/10.1080/09541440340000268

Tryfon, A., Foster, N., Sharda, M., & Hyde, K. (2018). Speech perception in autism spectrum disorder: An activation likelihood estimation meta-analysis. *Behavioural Brain Research*, *338*, 118–127.

Tsang, T., Naples, A. J., Barney, E. C., Xie, M., Bernier, R., Dawson, G., Dziura, J., Faja, S., Jeste, S. S., McPartland, J. C., Nelson, C., Murias, M., Seow, H., Sugar, C., Webb, S. J., Shic, F., & Johnson, S. (2022). Attention allocation during exploration of visual arrays in ASD: Results from the ABC-CT feasibility study. *Journal of Autism and Developmental Disorders*, online ahead of print. https://doi.org/10.1007/s10803-022-05569-0

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, *102*(4), 1181–1186. https://doi.org/10.1073/pnas.0408949102

Wang, Q., Hoi, S. P., Wang, Y., Song, C., & Yi, L. (2020). Out of mind, out of sight? investigating abnormal face scanning in autism spectrum disorder using gaze-contingent paradigm. *Developmental Science*, *23*, e12856. https://doi.org/10.1111/desc.12856

Wang, Q., Wall, C. A., Barney, E. C., Bradshaw, J. L., Macari, S. L., Chawarska, K., & Shic, F. (2013). Promoting social attention in 3-year-olds with ASD through gaze-contingent eye tracking. *Autism Research*, *13*, 61–73. https://doi.org/10.1002/aur.2199

Wechsler, D. (2014). *Wechsler Preschool and Primary Scale of Intelligence—Fourth CN Edition (WPPSI-IV CN)*. (Li, Y., Trans, Z. J., & Li, Y., Eds.). King-May Company China.

Woynaroski, T. G., Kwakye, L. D., Foss-Feig, J. H., Stevenson, R. A., Stone, W. L., & Wallace, M. T. (2013). Multisensory speech perception in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(12), 2891–2902. https://doi.org/10.1007/s10803-013-1836-5

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, *12*(5), 798–814. https://doi.org/10.1111/j.1467-7687.2009.00833.x

Zhang, J., Meng, Y., He, J., Xiang, Y., Wu, C., Wang, S., & Yuan, Z. (2019). McGurk effect by individuals with autism spectrum disorder and typically developing controls: A systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, *49*, 34–43. https://doi.org/10.1007/s10803-018-3680-0

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: Feng, S., Wang, Q., Hu, Y., Lu, H., Li, T., Song, C., Fang, J., Chen, L., & Yi, L. (2022). Increasing Audiovisual Speech Integration in Autism Through Enhanced Attention to Mouth. *Developmental Science*, e13348. https://doi.org/10.1111/desc.13348