Joint attention (JA), emerging as early as 6 to 9 months, is a prelinguistic social-communicative skill to share experiences of some third objects or events by directing (initiating JA; JA) or following (responding to JA; RJA) the eye gaze or pointing to social partners (Mundy & Jarrold, 2010). As an important developmental precursor to overall social and cognitive abilities (Mundy & Jarrold, 2010), JA is found to be associated with later language development in both typical and atypical development (e.g., Bottema-Beutel, 2016; Brooks & Meltzoff, 2005; Mundy et al., 2007), as well as the symptom severity in autism spectrum disorder (ASD), a neurodevelopmental disorder characterized by

Method

Participants

Given the novel research questions and data analysis methods in our study, it was difficult to rely on previous effect sizes when designing the current study. Thus, the current sample size was determined by previous studies using the gaze-contingent approach combined with the JA/gaze-following paradigm (Bayliss et al., 2013; Caruana et al., 2018; Mundy et al., 2016; Oberwelland et al., 2016, 2017). The sample size for one participant group ranged between 16 and 32 in those studies, with $M = 23.11$ and $SD = 6.64$. Furthermore, when we opted for a moderate effect size ($\eta_p^2 = .06$), 0.85 power, an alpha of .05, and 0.5 as a correlation among repeated measures to perform power analysis using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007), a total sample of at least 32 individuals was required by a repeated-measures ANOVA with Group (ASD and TD) as the between-subjects factor and Condition (congruent, incongruent, and closed-eye gaze) as the within-subjects factor.

In the current study, after excluding four children with ASD who had an IQ lower than 65 as measured by the Wechsler Intelligence Scale, 26 Chinese children with ASD (24 boys) and 24 Chinese TD children (23 boys) participated in our study. They were approximately 7 years old (see Table 1 for details). We selected children at this age since it is a potentially sensitive developmental period for gaze perception among children whose more basic visual mechanisms are presumably in place (Mihalache et al., 2019). Furthermore, Thorup, Kleberg, and Falck-Ytter (2017) found that children with ASD at approximately 7 years old could follow others' gaze based on measuring children's passive responses to others' gaze directions (Thorup et al., 2017). However, it remains unclear whether children with ASD at this age are sensitive to other's gaze following. Two children with ASD were excluded from our analysis due to the poor quality of the data on their eye movements (see the Data Analysis section for details), resulting in 24 children with ASD (22 boys) in the final sample (see Table 1 for details). The two groups were matched by chronological age and IQ (see Table 1). Detailed descriptions of participant characteristics are provided in Table 1 and the online supplemental materials. The present protocol (protocol number: 2016–03-03e) was approved by the Committee for Protecting Human and Animal Subjects at the School of Psychological and Cognitive Sciences at Peking University, China. We obtained oral consent from all of the children and written consent from all of their parents before conducting the experiment.

Materials

Sixty images of fruits and vegetables were taken from the Internet. Three images of male faces were selected from six images of faces created by FaceGen, a commercial software program (https://facegen.com/). The three images were rated on a scale from 1 (*very unattractive*) to 5 (*very attractive*) by a group of college students ($N = 20$) and matched for attractiveness (mean attractiveness = 2.6, 2.4, and 2.2, $SD = 1.19$, 0.88, and 0.83, respectively). We used virtual faces rather than real faces due to the advantages of a high degree of standardization and systematic manipulability. Each face was digitally edited using FaceGen to produce three versions of 26 continua for each of the following: direct gaze to left averted gaze, direct gaze to right averted gaze, and direct gaze to closed-eye gaze. These continua were used to present dynamic gaze-shifting.

Eye movement data were recorded using a Tobii Pro X3-120 eye tracker (sampling rate: 120 Hz; Tobiitech Technology, Stockholm, Sweden). The Psychtoolbox (http://psychtoolbox.org) and Tobii Analytics Software Development Kit (Tobiitech Technology, Stockholm, Sweden) on the MATLAB platform were used to control stimulus presentation and data recording.

Procedure

The children sat approximately 60 cm away from a 21.5-in. LCD monitor (1920 × 1080 pixels resolution). Their eye movements were first calibrated using a 5-point calibration procedure. During the calibration, an animated cartoon character paired with an engaging sound appeared sequentially in the center and four

Table 1
Characteristics of the Participants

Variable	ASD (N = 24) *M*	ASD *SD*	ASD Range	TD (N = 24) *M*	TD *SD*	TD Range	*t*	*p*
Age (years)	7.22	1.58	5.08–11.57	7.49	.66	6.50–8.65	−.769	.446
Full Scale IQ[a]	98.54	18.58	69–136	95.96	10.5	77–117	.592	.557
ADOS[b] total severity	8.37	1.53	5–10					
SA severity[c]	8.46	1.47	5–10					
RRB severity[d]	7.75	1.11	5–10					
ADI-R[e]								
Social interaction	21.88	5.57	10–30					
Communication	17.75	4.80	9–26					
RRB	8.67	2.08	5–12					
D Scale[f]	3.25	1.15	1–5					

Note. ASD = autism spectrum disorder; TD = typically developing. [a] IQ was measured using the abbreviated Chinese Fourth Edition version of the Wechsler Intelligence Scale for Preschool and Primary Children (Wechsler, 2014b) for children under 6 years old, and the abbreviated Chinese Fourth Edition version of the Wechsler Intelligence Scale for Children (Wechsler, 2014a) for children over 6 years old. [b] ADOS = Autism Diagnostic Observation Schedule. [c] SA Severity = ADOS Social Affect Severity. [d] RRB Severity = ADOS Restricted, Repetitive Behavior Severity. SA and RRB Severity were calculated according to Hus, Gotham, and Lord (2014). [e] ADI-R = Autism Diagnostic Interview—Revised. [f] The D Scale is abnormality of development evident at/before 36 months.

corners of the screen. The children were instructed to fixate on the character. The calibration process was repeated when necessary until both eyes achieved good mapping on all five test positions (smaller than 1° visual angle).

Each trial was preceded by an attention-getter (a cartoon character subtending a visual angle of 4°×4°) at the center of the monitor to attract children's attention. The attention-getter disappeared once the children's gaze was detected to be within the attention-getter region. Next, one face with a direct gaze (10°×10° visual angle) appeared at the center of the screen along with two objects (8°×8° visual angle) randomly chosen from the fruit and vegetable images pool, which appeared at the left and right sides of the face (the center of object images appeared approximately 10° from the center of screen; Figure 1). Children were instructed to look at one of the two objects they preferred. When gaze was detected continually within the first-looked-at object (FLO) for 30 ms, the virtual face began to shift its gaze to look at the FLO (congruent condition), to shift its gaze to look at another object (non-first-looked at object, NFLO; incongruent condition), or to close its eyes (closed-eye gaze condition). These dynamic gaze-shifting movements lasted approximately 1.2 s, followed by 3 s of the final gaze phase, during which the face gazing at the object continued as a still frame. The children were given no further instructions and could view the scene freely.

For a given child, each virtual face was randomly assigned to one type of condition (congruent, incongruent, or closed-eye gaze condition) and appeared 10 times, resulting in 30 trials in total. The trials were randomly presented with the constraint that the same condition could not occur more than three times in a row. Eye movement data were recorded during the whole experiment.

Eye Movement Data Analysis

Data preprocessing. Missing gaze data with a gap shorter than 75 ms were filled in using linear interpolation, whereas those with a gap that exceeded 75 ms, which was regarded as an eyeblink (Olsen, 2012), were kept and coded as looking at nothing. Trials for which more than 30% of the gaze data was interpolated were excluded from the analysis. After the exclusion, the average proportion of the interpolated data was similar for the ASD (M = 0.02, SD = 0.03) and TD (M = 0.02, SD = 0.01) groups, $t(46)$ = 0.66, p = .515, Cohen's d = 0.19, 95% CI [−0.38, 0.76]. The average gaze positions of the left and right eyes were used as an analytical unit. Trials were also excluded if the gaze shifting of the virtual faces was not induced by children's fixations (e.g., saccades, noises, etc.). The proportion of trials that were excluded for this reason was higher for the ASD group (M = 0.15, SD = 0.11) than that for the TD group (M = 0.08, SD = 0.08), $t(46)$ = 2.53, p = .015, Cohen's d = 0.73, 95% CI [0.14, 1.31]. Fixation was calculated based on an I-VT fixation filter (Olsen, 2012; Wang, Hu, et al., 2018) with the following parameter settings as follows: (1) Velocity threshold was set at 30°/s; (2) fixations that were spatially and temporally (<0.5°, <75 ms) close were merged to prevent longer fixations from being separated into shorter fixations because of data loss or noise; and (3) duration threshold was set to 60 ms.

To ensure the quality of the data, two children in the ASD group with fewer than five valid trials for each condition after trial rejection were excluded from further analyses. The average number of valid trials for each condition was quite high for both the ASD (M = 8.49, SD = 1.09) and TD (M = 9.21, SD = 0.83) groups, with a significant group difference, $t(46)$ = −2.59, p = .013, Cohen's d = 0.75, 95% CI [0.16, 1.33]. The average proportional total looking time on the screen relative to the trial duration analyzed (i.e., 4 s) was similar in the ASD (M = 0.74, SD = 0.16) and TD (M = 0.81, SD = 0.15) groups, $t(46)$ = −1.63, p = .110, Cohen's d = −0.47, 95% CI [−1.04, 0.11]. Areas of interest (AOIs) for the two objects (FLO and NFLO) and eyes are illustrated in Figure 2.

Attention to objects. Our major concern was whether children with and without ASD were sensitive to others' gaze re-

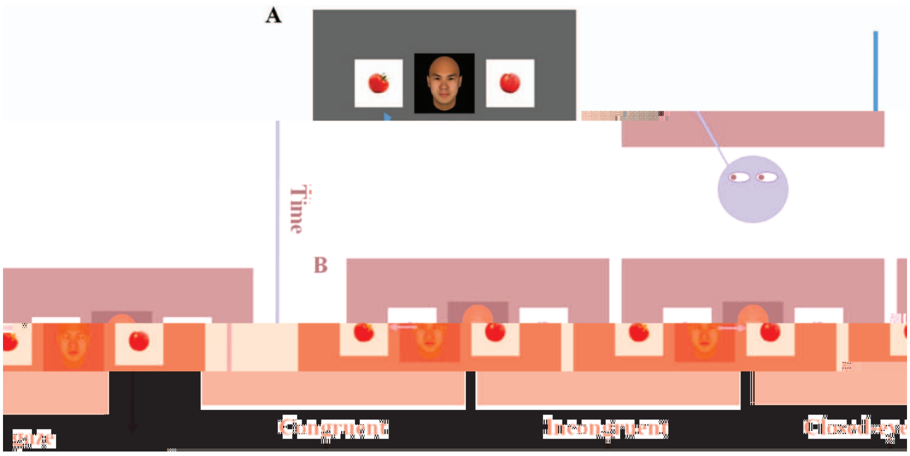

Figure 1. Experiment design. Children (cartoon face) initiate a joint attention by looking at one of the two objects (A), then the virtual face shifts its gaze (1.2 s) to follow the children's gaze in the congruent condition, to disregard the gaze and look at another object in the incongruent condition, or to close its eyes in the closed-eye gaze condition, followed by 3 s of the final gaze phase, during which the face gazing at the object continued as a still frame (B). Data from phase B were analyzed. See the online article for the color version of this figure.

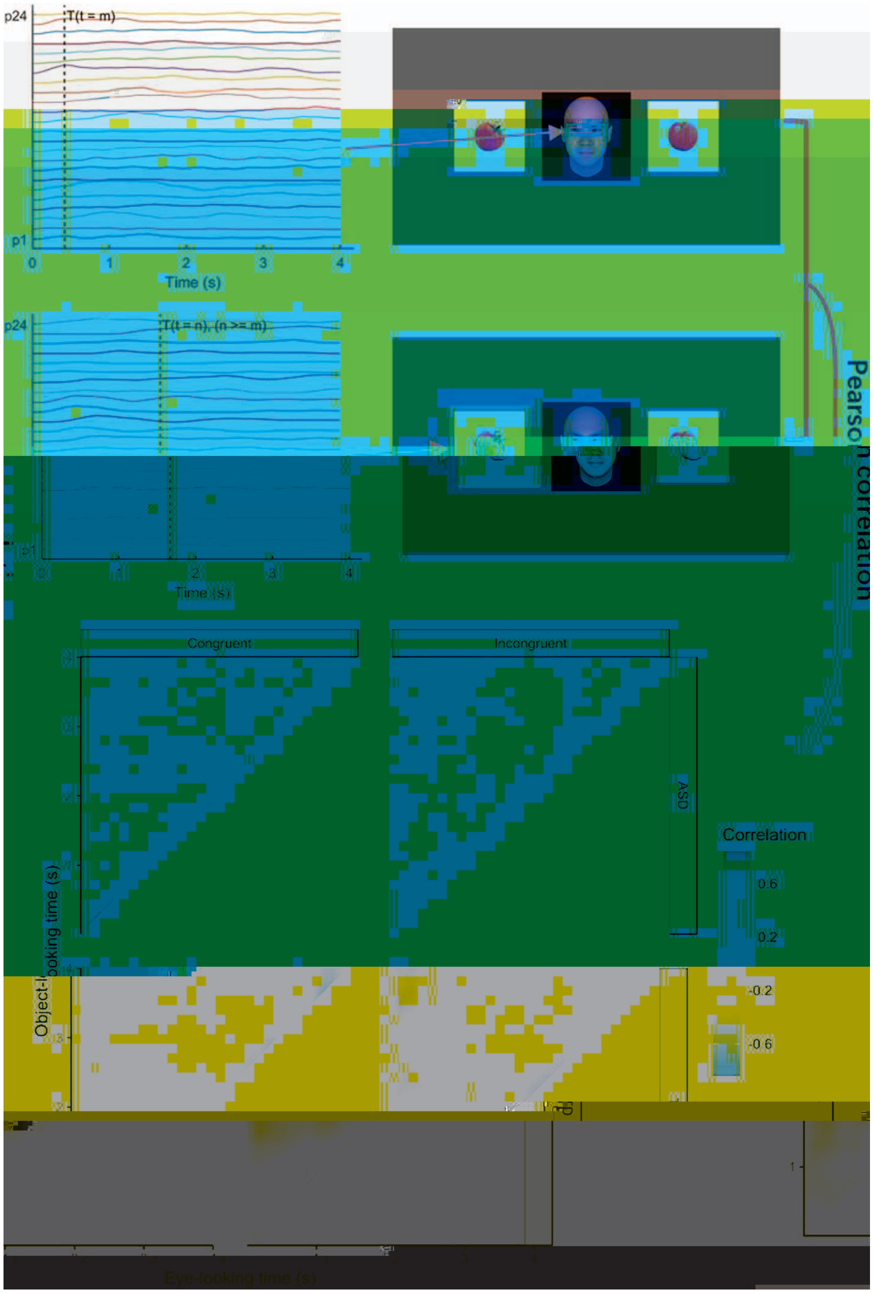

Figure 2. Results of the novel data-driven correlation method. For each data point of the time series signals, we correlated proportional eye-looking time across participants at time m with proportional looking time on the gazed-at object (FLO in the congruent condition or NFLO in the incongruent condition) at time n ($0 \leq m \leq n \leq 4$ s; top panel). This results in a 480 (data points) × 480 (data points) upper triangular matrix (bottom panel). Each value in the matrix represents a correlation coefficient between eye-looking time pattern at time m and looking time on gazed-at object at time n ($0 \leq m \leq n \leq 4$ s). That is, correlations are between the proportion of time spent looking at eyes and proportion of time spent looking at objects at any given time point throughout the trial, with the restriction that eye-looking time happens before object-looking time. Areas showing significant correlations are delimited by white borders (multiple comparisons were controlled by using the cluster-based permutation test). This analysis was done separately for each participant group and experimental condition. AOIs for the object and eyes (within the blue rectangles or regions pointed by arrows) are also illustrated in this figure. ASD = autism spectrum disorder; TD = typically developing. See the online article for the color version of this figure.

sponses, that is, whether their looking time at the object would be modulated by others' following or not following their own gaze. This attention effect, if any, could emerge at any time during a trial. For example, in the congruent condition, children might sustain their attention on the FLO when others follow their gaze, and the effect would therefore appear early in a trial. It is also possible that children might look at the face to extract gaze information after they have looked at the FLO, and then pay attention to the FLO again. In this case, the effect would appear late in a trial. Since we had no prior hypothesis about when the effect would happen, we employed a novel data-driven time-course analysis to investigate the effect (for a similar method, see Wang, Lu, et al., 2018). In brief, we created a time series signal of the proportional object-looking time by calculating the

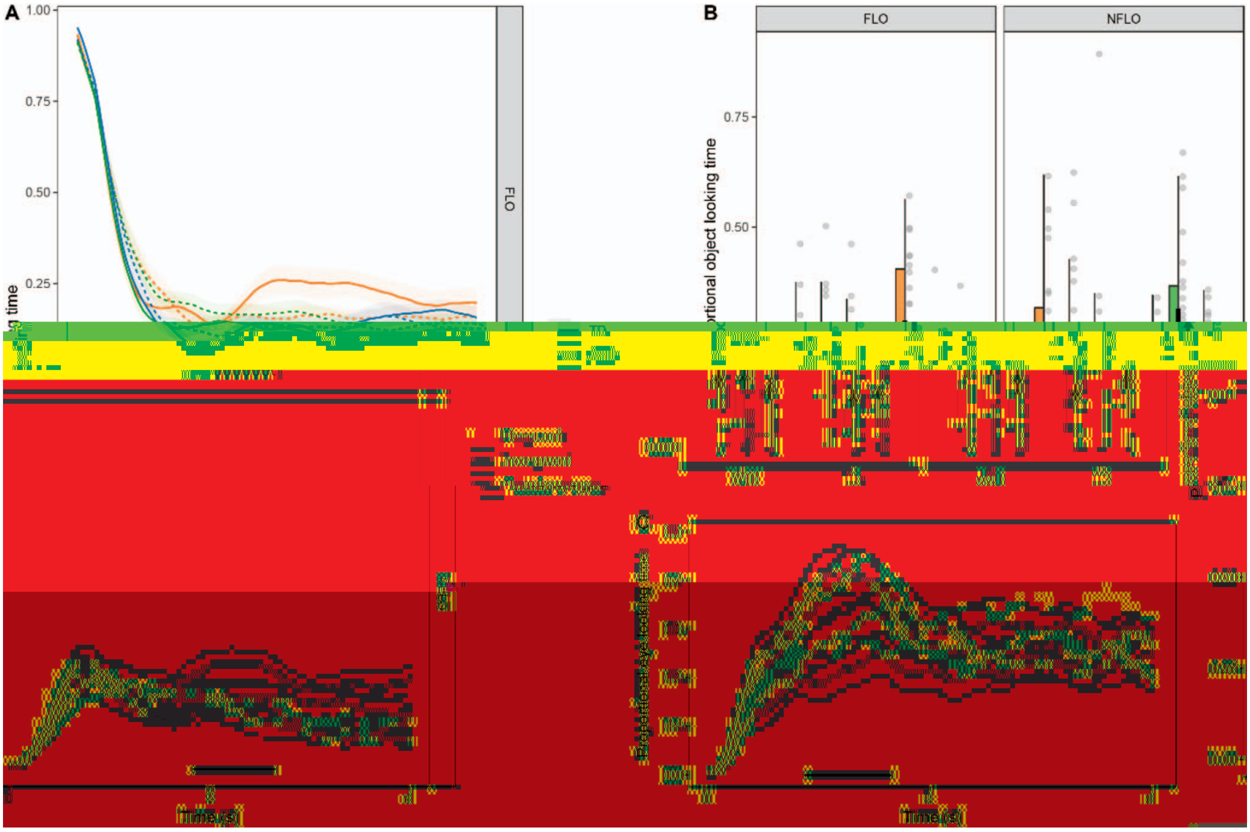

Figure 3. Results of the eye- and object-looking time. Time-course of proportional object-looking time (A) and eye-looking time (C). Time series signals of proportional AOI-looking time were created by calculating the proportional trial toward a particular AOI relative to the total number of valid trials for each data point. Multiple comparisons were corrected by using a cluster-based permutation test. Black horizontal line illustrates the cluster of time when the Condition × Group interaction effect (A) or Group main effect (C) is significant. Shaded area indicates standard errors. Time zero is the start of the face's gaze shifting. (B) Boxplot of object-looking time during significant time periods revealing interaction effect in Figure 3A, with each triangle representing mean value, each thick black vertical line representing error bar (standard error), and each point representing one child. ASD = autism spectrum disorder; TD = typically developing; FLO = first-looked-at object; NFLO = non-first-looked at object. See the online article for the color version of this figure.

Attention to Eyes

The time-course analysis only found the main effect of Group on the proportional eye-looking time during 0.88–1.66 s, $F_{\text{sum}} = 607.89$, $p = .046$. Data in the significant time period revealing the Group main effect were extracted and averaged accordingly. A 2 (Group: ASD vs. TD) × 3 (Condition: congruent, incongruent, closed-eye gaze) rmANOVA was conducted on the average data and confirmed the time-course analysis result: Only the main effect of Group was significant, $F(1, 46) = 6.93$, $p = .011$, $\eta_p^2 = 0.13$, 90% CI [0.02, 0.28]. This main effect remained significant when rerunning the analysis to include the average proportional total looking time on the screen as a covariate, $F(1, 45) = 4.26$, $p = .045$, $\eta_p^2 = 0.09$, 90% CI [0.00, 0.23]. Therefore, children with ASD looked at the eyes less than TD children during 0.88–1.66 s (after the virtual face shifted its gaze), which was earlier than the time period that revealed the Condition difference.

Correlation Between Eye-Looking Time and Object-Looking Time

Data-driven correlation analysis (see Figure 2) revealed that for the TD group, there was one cluster showing significant positive correlations between proportional eye-looking time and proportional looking time on the gazed-at object in the incongruent condition: Eye-looking time in the period from 0.07 to 1.29 s positively predicted object-looking time in the period from 1.51 to 4.00 s, $Z_{\text{sum}} = 54{,}092$, $p = .041$. Correlations were not significant after correction in the congruent condition.

For the ASD group, there was one cluster showing significant negative correlations between proportional eye-looking time and proportional looking time on the gazed-at object in the incongruent condition: Eye-looking time in the period from 0.66 to 2.48 s negatively predicted object-looking time in the period from 0.75 to 4.00 s, $Z_{sum} = -80{,}820$, $p = .046$. Correlations were not significant after correction in the congruent condition.

Discussion

Using a computer-based gaze-contingent design and novel time-course analyses, we investigated the eye movements in TD and ASD children in response to others' gaze following with respect to their own gazes. Specifically, we tested (1) how children attended to the objects in response to others' gaze following or failure to follow, (2) whether children with ASD displayed atypical attention to the partners' eyes during JA, and (3) whether attention to eyes influenced subsequent attention to objects.

First, we found that TD children's attention to the objects was modulated by others' gaze responses: They spent higher proportional FLO-looking time in the congruent condition than they did in the incongruent and closed-eye gaze conditions, and they spent higher proportional NFLO-looking time in the incongruent condieyes play during gaze-based interactions in both TD and ASD children. A related issue is whether the relationship between eye-looking time and object-looking time during JA is relevant to theory of mind. Since monitoring a person's gaze/attention is an example of monitoring a person's mental state (Baron-Cohen, 1991), the absence of positive correlations between eye-looking time and object-looking time in ASD children might be attributed to their deficits in theory of mind. However, we did not examine what kind of role theory of mind played in children's gaze following in our study, a topic that could be further investigated by follow-up studies. Fourth, as in a real-life situation, we did not instruct children to attend to the faces or eyes. Whether instructing children with ASD to attend to the interactive face's gaze will improve their JA is an interesting question and may shed light on developing intervention methods aiming to improve JA in individuals with ASD. Fifth, having one's own gaze followed affects how a social partner is perceived (Bayliss et al., 2013); for example, adults favor others who follow their gaze (Bayliss et al., 2013). Likewise, children could also learn and establish that association (e.g., face in the congruent condition = good face, face in the incongruent condition = bad face, and face in the closed-eye gaze condition = ignorant face). It would be interesting to test how learning outcome influences children's gaze following and how gaze following changes during learning course. However, these issues were not testable in our current study due to the limited trial numbers and absence of learning outcome measurements, making them a topic for future research. Lastly, previous fMRI studies using a similar paradigm set both the gaze-shift duration and the final gaze phase duration for 1 s (Oberwelland et al., 2016, 2017). We used similar gaze-shift durations (1.2 s) but longer final-gaze durations (3 s) to collect more eye-movement data. The length of the stimulus presentation time might influence the outcome, which could be examined in future investigations.

In conclusion, this study bridged a significant gap in the literature by studying gaze response to others' gaze following in children with and without ASD. TD children, but not ASD children, responded effectively and flexibly to others' gaze following of their own gazes. This study contributes to an understanding of the process of a more complex and reciprocal JA in TD children and abnormal social cognition in children with ASD in the context of ecologically valid social interactions.

References

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind* (pp. 233–251). Oxford, England: Basil Blackwell.

Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2013). "Gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General, 142,* 76–92. http://dx.doi.org/10.1037/a0029286

Bedford, R., Elsabbagh, M., Gliga, T., Pickles, A., Senju, A., Charman, T., . . . the BASIS team. (2012). Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *Journal of Autism and Developmental Disorders, 42,* 2208–2218. http://dx.doi.org/10.1007/s10803-012-1450-y

Bottema-Beutel, K. (2016). Associations between joint attention and language in autism spectrum disorder and typical development: A systematic review and meta-regression analysis. *Autism Research, 9,* 1021–1035. http://dx.doi.org/10.1002/aur.1624

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8,* 535–543. http://dx.doi.org/10.1111/j.1467-7687.2005.00445.x

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63,* i-174. http://dx.doi.org/10.2307/1166214

Caruana, N., Stieglitz, H. H., Brock, J., Woolgar, A., Kloth, N., Palermo, R., & Mcarthur, G. (2018). Joint attention difficulties in autistic adults: An interactive eye-tracking study. *Autism, 22,* 502–512.

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society of London Series: Biological Sciences, 358,* 315–324. http://dx.doi.org/10.1098/rstb.2002.1199

Chawarska, K., Klin, A., & Volkmar, F. (2003). Automatic attention cueing through eye movement in 2-year-old children with autism. *Child Development, 74,* 1108–1122. http://dx.doi.org/10.1111/1467-8624.00595

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences, 16,* 231–239. http://dx.doi.org/10.1016/j.tics.2012.02.007

Collis, G. M. (1977). *Visual co-orientation and maternal speech*. London, England: Academic Press.

Falck-Ytter, T., Thorup, E., & Bölte, S. (2015). Brief report: Lack of processing bias for the objects other people attend to in 3-year-olds with autism. *Journal of Autism and Developmental Disorders, 45,* 1897–1904. http://dx.doi.org/10.1007/s10803-014-2278-4

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism. *Journal of the American Academy of Child & Adolescent Psychiatry, 56,* 546–555. http://dx.doi.org/10.1016/j.jaac.2017.05.005

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133,* 694–724. http://dx.doi.org/10.1037/0033-2909.133.4.694

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48,* 1711–1725. http://dx.doi.org/10.1111/j.1469-8986.2011.01273.x

Hus, V., Gotham, K., & Lord, C. (2014). Standardizing ADOS domain scores: Separating severity of social affect and restricted and repetitive behaviors. *Journal of Autism and Developmental Disorders, 44,* 2400–2412. http://dx.doi.org/10.1007/s10803-012-1719-1

Kuhn, G., Benson, V., Fletcher-Watson, S., Kovshoff, H., McCormick, C. A., Kirkby, J., & Leekam, S. R. (2010). Eye movements affirm: Automatic overt gaze and arrow cueing for typical adults and adults with autism spectrum disorder. *Experimental Brain Research, 201,* 155–165. http://dx.doi.org/10.1007/s00221-009-2019-7

Kylliäinen, A., & Hietanen, J. K. (2004). Attention orienting by another's gaze direction in children with autism. *Journal of Child Psychology and Psychiatry, 45,* 435–444. http://dx.doi.org/10.1111/j.1469-7610.2004.00235.x

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190. http://dx.doi.org/10.1016/j.jneumeth.2007.03.024

Mihalache, D., Feng, H., Askari, F., Sokol-Hessner, P., Moody, E. J., Mahoor, M. H., & Sweeny, T. D. (2019). Perceiving gaze from head and eye rotations: An integrative challenge for children and adults. *Devel-*

opmental Science. Advance online publication. http://dx.doi.org/10.1111/desc.12886

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development, 78,* 938–954. http://dx.doi.org/10.1111/j.1467-8624.2007.01042.x

Mundy, P., & Jarrold, W. (2010). Infant joint attention, neural networks and social cognition. *Neural Networks, 23,* 985–997. http://dx.doi.org/10.1016/j.neunet.2010.08.009

Mundy, P., Kim, K., McIntyre, N., Lerro, L., & Jarrold, W. (2016). Brief report: Joint attention and information processing in children with higher functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders, 46,*