MIND, BRAIN, AND EDUCATION

# Mechanisms by Which Early Eye Gaze to the Mouth During Multisensory Speech Influences Expressive Communication Development in Infant Siblings of Children with and Without Autism

**Pooja Santapuram[1], Jacob I. Feldman[2], Sarah M. Bowman[2,3], Sweeya Raj[4], Evan Suzman[5], Shannon Crowley[6], So Yoon Kim[7], Bahar Keceli-Kaysili[2], Kristen Bottema-Beutel[6], David J. Lewkowicz[8], Mark T. Wallace[2,9,10,11,12,13,14], and Tiffany G. Woynaroski[2,9,10,11]**

ABSTRACT— Looking to the mouth of a talker early in life predicts expressive communication. We hypothesized that looking at a talker's mouth may signal that infants are ready for increased supported joint engagement and that it subsequently facilitates prelinguistic vocal development and translates to broader gains in expressive communication. We tested this hypothesis in 50 infants aged 6–18 months with the heightened and general population-level likelihood of autism diagnosis (Sibs-autism and Sibs-NA; respectively). We measured infants' gaze to a speaker's face using an eye-tracking task, supported joint engagement during parent–child free play sessions, vocal complexity during a communication sample, and broader expressive communication. Looking at the mouth was indirectly associated with expressive communication via increased higher-order supported joint engagement and vocal complexity. This indirect effect did not vary according to sibling status. This study provides preliminary insights into the mechanisms by which looking at the mouth may influence expressive communication development.

[1]Vanderbilt School of Medicine, Vanderbilt University,
[2]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center
[3]Augusta University/University of Georgia Medical Partnership at the Medical College of Georgia, Augusta University
[4]Neuroscience Undergraduate Program, Vanderbilt University
[5]Master's Program in Biomedical Sciences, Vanderbilt University
[6]Lynch School of Education and Human Development, Boston College
[7]Department of Teacher Education, Duksung Women's University
[8]Haskins Laboratories, Yale Child Study Center
[9]Vanderbilt Kennedy Center, Vanderbilt University Medical Center
[10]Vanderbilt Brain Institute, Vanderbilt University
[11]Frist Center for Autism and Innovation, Vanderbilt University
[12]Department of Psychology, Vanderbilt University
[13]Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Cente
[14]Department of Pharmacology, Vanderbilt University

Address correspondence to Tiffany Woynaroski, MCE 8310 South Tower, 1215 21st Avenue South, Nashville, TN 37232; e-mail: tiffany.g.woynaroski@vumc.org

Pooja Santapuram and Jacob I. Feldman contributed equally to this work.

Autism is a neurodevelopmental condition characterized by social communication differences in addition to restricted and repetitive patterns of behavior, interests, and activities (American Psychiatric Association, 2013). Many children with autism[1] also present with broader language and communication impairments (Tager-Flusberg, Paul, & Lord, 2005). In fact, difficulty acquiring expressive communication during the first few years of life is the single-most replicated predictor of long-term outcomes in this population (e.g., Billstedt, Carina Gillberg, & Gillberg, 2007; Eisenberg, 1956; Venter, Lord, & Schopler, 1992). These findings have spawned a large and ever-growing body of research seeking to identify factors that predict expressive communication in young children on the spectrum (see Yoder, Watson, & Lambert, 2015 for a review).

### Looking at the Face as a Replicated Predictor of Expressive Communication

One replicated predictor of expressive communication is eye gaze. Early eye gaze patterns follow a known developmental trajectory. Specifically, at birth, infants tend to look predominantly at the eyes of their communication partners (e.g., Haith, Bergman, & Moore, 1977; Maurer & Salapatek, 1976). Midway through the first year of life, however, infants begin to shift their gaze towards the mouth of their communication partners, in particular when such partners are producing audiovisual speech or “talking” (e.g., Lewkowicz & Hansen-Tift, 2012; Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013). Looking at a talker’s mouth is especially prominent during periods marked by qualitative shifts in expressive communication development (i.e., around 8–18 months; Hillairet de Boisferon, Tift, Minar, & Lewkowicz, 2018; Lewkowicz & Hansen-Tift, 2012; during this time, infants begin to imitate adult speech, communicate with purpose, and acquire their first words). The temporal coincidence of increased looking at the mouth with early communication milestones has led researchers to hypothesize that looking at audiovisual speech, in particular early in life, facilitates expressive communication development, perhaps by boosting speech processing efficiency and/or perceptual accuracy (e.g., Calvert, Brammer, & Iversen, 1998; Middelweerd & Plomp, 1987; Sumby & Pollack, 1954; van Wassenhove, Grant, & Poeppel, 2005).

There is increasing empirical support for the hypothesis that looking at the mouth during audiovisual speech bootstraps expressive communication development in individuals with autism. Several studies, for example, have demonstrated an association between looking at the mouth of a talker and aspects of language development and/or broader communicative competence in children and adults with autism both concurrently (e.g., Chawarska, Macari, & Shic, 2012; Habayeb et al., 2020; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Norbury et al., 2009) and prospectively (e.g., Campbell, Shic, Macari, & Chawarska, 2014). The mechanisms by which looking to the mouth may scaffold expressive communication development in autism, however, remain unclear.

Relations between looking to the eyes of a speaker during audiovisual speech and expressive communication have been observed less frequently. However, in some instances looking at both the mouth and the eyes of a talker have been shown to account for variance in communication skills (e.g., individual differences in word learning in Tenenbaum, Amso, Abar, & Sheinkopf, 2014). In addition, an emerging line of evidence suggests that looking at the eyes during audiovisual speech at 12 months is correlated with concurrent social skills (Pons, Bosch, & Lewkowicz, 2019) and that looking to the eyes at 6 months is correlated with later social skills (Wagner, Luyster, Moustapha, Tager-Flusberg, & Nelson, 2018) in non-autistic infants. These findings suggest that social skills may benefit from attention to the social information inherent in a talker’s eyes more than linguistic information inherent in a talker’s mouth. Thus, given the core impairment in social communication skills associated with autism, it is important to also examine links between looking at the eyes during audiovisual speech and communication skills.

### Putative Mechanisms by Which Looking at Audiovisual Speech May Influence the Development of Expressive Communication in Autism

We hypothesize that early looking to audiovisual speech may indirectly influence the development of expressive communication via parent–child engagement and children’s prelinguistic vocal complexity. Specifically, the shift in looking to a talker’s mouth in the first year of life is followed by a developmental period marked by increased joint engagement between parent–child dyads. This increase in joint engagement is evident first in what is termed supported joint engagement and, later in the course of development, via coordinated attention between a child and caregiver. We focus on supported joint engagement as the putative mechanism by which early looking to the mouth may support the development of expressive communication here because this type of joint engagement is most proximal, from a developmental perspective, to the predictor of interest.

#### *Supported joint engagement*

Supported joint engagement has been frequently evaluated as a potential predictor of later communication in children with autism. Early in life, caregivers support children’s engagement in joint play routines in a manner that scaffolds development across a number of domains (Bakeman & Adamson, 1984; Bottema-Beutel, Yoder, Hochman, & Watson, 2014). There is at least some evidence to suggest that caregiver supported joint engagement is linked with communication in children with or at increased likelihood for autism (Adamson, Bakeman, Deckner, & Romski, 2009; Gulsrud, Hellemann, Shire, & Kasari, 2016; Kasari, Paparella, Freeman, & Jahromi, 2008).

Recent work has differentiated between lower-order supported joint engagement (LSJE; wherein a caregiver appears to influence a child's play, but the child does not actively acknowledge the caregiver with gaze) and higher-order supported joint engagement (HSJE; wherein a caregiver influences a child's play, and the child acknowledges the caregiver by engaging in reciprocal play; Bottema-Beutel et al., 2014). Several studies have found HSJE may be superior to LSJE for predicting later expressive communication in children with autism (Bottema-Beutel et al., 2014; Bottema-Beutel et al., 2019; Crandall, Bottema-Beutel, McDaniel, Watson, & Yoder, 2019).

We suspect that a child's generalized tendency to look to the mouth during face-to-face interactions with an adult producing audiovisual speech may support the acquisition of communication skills by signaling to caregivers that children are increasingly "tuned in" to the parent's communication and developmentally ready for supported joint engagement, in particular HSJE, wherein the child actively acknowledges the communication partner. This may lead the caregiver to make increased efforts to engage the child in scaffolded interactions marked by reciprocal engagement around objects or events, a context in which child communication is routinely made relevant and which; thus, is especially likely to facilitate prelinguistic vocal development and subsequently translate to broader gains in expressive communication.

*Prelinguistic vocal complexity*

We hypothesize that the influence of looking to the mouth on expressive communication development through increased supported joint engagement will be detectable first in increased prelinguistic vocal complexity. Prelinguistic vocal complexity is commonly operationalized as the number of consonants and/or the proportion of canonical syllables (i.e., consonant-vowel combinations produced with adult-like speech timing) that a child produces in prelinguistic communication acts (i.e., consonant inventory and canonical syllabic communication; e.g., McDaniel, Woynaroski, Keceli-Kaysili, Watson, & Yoder, 2019; Saul & Norbury, 2020; Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski, Yoder, Fey, & Warren, 2014; Woynaroski et al., 2016, 2017). Recent research has shown that these indices of vocal complexity are useful for predicting language growth and outcomes in children with autism, even after controlling for a number of other factors previously linked with language development (e.g., parent linguistic input, responding to joint attention, intentional communication; Wetherby et al., 2007; Woynaroski et al., 2017; Yoder et al., 2015).

Thus, the links for both supported joint engagement, in particular HSJE, and prelinguistic vocal complexity with expressive communication development are well-established in autism. However, to our knowledge, no prior work has evaluated the hypothetical link between early looking to the mouth and expressive communication development through HSJE and prelinguistic vocal complexity. We aimed to do so by statistically testing whether HSJE and prelinguistic vocal complexity partially or completely accounted for (i.e., mediated; Hayes, 2017) associations between looking to the mouth during audiovisual speech and concurrent expressive communication, particularly early in life.

**Rationale for a Focus on Infants at Heightened Likelihood for Autism**

A challenge to testing this hypothesized indirect path through which early looking to audiovisual speech may influence expressive communication development is that autism cannot always be reliably diagnosed in the earliest stages of development (Ozonoff et al., 2015; Turner & Stone, 2007), in particular during the timeframe in which looking to audiovisual speech has been posited to predict vocal development (i.e., beginning at approximately 8 months of age; for discussion, see Lewkowicz & Hansen-Tift, 2012). A potential solution to this problem is to prospectively study infants who are known to be at increased likelihood for a future diagnosis of autism based on their status as younger siblings of children who are already diagnosed with autism (Sibs-autism; Costanzo et al., 2015). Approximately one-third of infant siblings of children with autism will go on to receive a diagnosis of autism and/or language impairment; many more will show sub-clinical characteristics associated with the disorder (Charman et al., 2017; Messinger et al., 2013; Ozonoff et al., 2011). These Sibs-autism are often contrasted with a group of infants at relatively lower, general population-level likelihood for autism (Sibs-NA; i.e., infant siblings of non-autistic, otherwise typically developing children). Prior studies taking this approach to date have found positive associations between looking to a talker's mouth and later expressive communication and negative associations between looking to a talker's eyes and later expressive communication in infants at increased likelihood of being diagnosed with autism (Elsabbagh et al., 2014; Wagner et al., 2018; Young, Merin, Rogers, & Ozonoff, 2009). Although concurrent links have between looking to the mouth and concurrent expressive communication have been reported in young children with autism (e.g., Chawarska et al., 2012; Tenenbaum et al., 2014), no study to our knowledge has evaluated these links concurrently in Sibs-autism. In the present study, we employ this approach in an attempt to test our hypotheses regarding *how* looking to audiovisual speech influences concurrent expressive communication development in infants with an elevated likelihood of autism diagnosis.

### Research Questions

The research questions for the present study were as follows:

(a) Do Sibs-autism differ from Sibs-NA in their early eye gaze patterns to audiovisual speech (i.e., looking to the mouth and eyes of a talker)?

(b) Do individual differences in looking to audiovisual speech indirectly influence concurrent expressive communication through parent–child engagement and prelinguistic vocal complexity? Does this indirect effect vary according to sibling status?

## METHODS

### Design Overview

Recruitment and study procedures were carried out in accordance with the approval of the Vanderbilt University Medical Center Institutional Review Board. Parents provided written informed consent prior to their child's participation in the study. All families were compensated for their participation.

### Participants

Participants were 50 infants (28 Sibs-autism, 22 Sibs-NA; see Table 1 for sample characteristics) drawn from an ongoing longitudinal study (e.g., Bottema-Beutel et al., 2019; Feldman et al., 2021). Participants made one to two visits to Vanderbilt University Medical Center over the course of a 2-week period during which all study measures were collected.

Inclusion criteria were: (a) chronological age 6–18 months, (b) predominantly English-speaking household, and (c) an older sibling (30 months–18 years) who was either diagnosed with autism (Sibs-autism) or who had no autism diagnosis (Sibs-NA). Groups did not differ at the group level on chronological age or biological sex ($p$-values >0.5). The diagnostic status of older siblings with autism was confirmed by record review (in the cases wherein the older sibling had been diagnosed at Vanderbilt) or by an independent assessment that included detailed history, parent interview, and research-reliable administration of the Autism Diagnostic Observation Schedule, second edition (Lord et al., 2012). The diagnostic status of non-autistic older siblings was confirmed by a score below the screening threshold for autism concern on the Social Communication Questionnaire (Rutter, Bailey, & Lord, 2003) and no history of developmental concerns per parent report. Exclusion criteria for all participants included: (a) adverse neurological history, (b) known genetic condition, and (c) pre-term birth (gestation <37 weeks). Communication assessments for the protocol were administered in English and, thus, families who did not speak English at home were excluded. The Mullen Scales of Early Learning (MSEL; Mullen, 1995) was administered to further characterize the sample.

**Table 1**
Participant demographics by sibling group

| | *Sibs-autism* (n = 28) M (SD) *Min–Max* | *Sibs-NA* (n = 22) M (SD) *Min–Max* |
|---|---|---|
| Age in months | 11.86 (3.78) 6–18 | 12.73 (3.51) 6–18 |
| Percent male | 53.6% | 50.0% |
| MSEL ELC* | 92.64 (12.19) 71–115 | 102.45 (9.85) 84–121 |
| Race | *n* | *n* |
| White | 28 | 20 |
| Black/African-American | 0 | 1 |
| Multiple | 0 | 1 |
| Ethnicity | *n* | *n* |
| Hispanic/Latino | 1 | 1 |
| Not Hispanic/Latino | 27 | 21 |
| Primary caregiver's highest level of education | *n* | *n* |
| High school diploma or GED | 2 | 0 |
| College/technical (1–2 years) | 10 | 3 |
| College/technical (3–4 years) | 8 | 7 |
| Graduate/professional school (1–2 years) | 5 | 5 |
| Graduate/professional school (3+ years) | 3 | 7 |

*Note*: MSEL ELC = Mullen Scales of Early Learning (Mullen, 1995) Early Learning Composite.
*Groups differed at $p < .01$.

### Procedures

#### *Looking to audiovisual speech*

Eye gaze to a talker's mouth, eyes, and broader face regions was measured while infants were exposed to a 50 s video of a monologue produced by a female actor speaking in an infant-directed manner in the infants' native language (English). The infant-directed speech was characterized by a slowed rate, as well as exaggerated speech prosody and pitch/frequency contours. This video has been utilized in prior research (Dunham et al., 2020; Lewkowicz & Hansen-Tift, 2012). Participants were seated 60–80 cm in front of a 22 in. computer monitor connected to an SMI REDn SensoMotoric Instruments (SMI, Teltow, Germany) remote eye tracker running on a laptop computer at a sampling rate of 60 Hz. The eye tracker's camera was connected to the bottom of the external computer monitor. We used SMI's iViewRed software to control the eye tracker camera and to process the eye gaze data and SMI's Experiment

Center software to control stimulus presentation and data acquisition. We presented the video on the external monitor and the auditory stimuli through an external speaker placed behind the monitor and tracked the participants' eye gaze. Prior to showing the video, we ran a 5-point calibration routine by showing a small star in the middle of the screen as well as in each corner of the screen.

Analysis of eye gaze data was completed using SMI's BeGaze program. To determine how much time each infant spent gazing to the eyes and mouth regions of the stimulus, ovular areas of interest (AOIs) were imposed onto the talker's mouth, eyes, and face in the video clip (see Figure 1). The AOI around the eyes was defined by a region that was demarcated on the top by a curved line slightly above the eyebrows and on the bottom by a curved line at the bridge of the nose. Left and right edges of the AOI were demarcated by curved lines at the edge of the hairline on either side of the speaker's face. The AOI around the mouth was defined by a region that was delineated on the top by a curved line at the bottom of the nose and at the bottom by a curved line at the middle of the chin. Both the left and right edges of the mouth AOI were delineated by curved lines at the middle of the cheek on either side of the face. Then, an AOI was created encompassing the talker's entire face. These AOIs were static (i.e., the AOIs did not change on a frame-by-frame basis).

Using these AOIs, we computed the percent of total looking time to the talker's eyes and mouth out of the total looking time to the talker's face. We chose to use these ratio metrics because (a) they are very stable in children (Dunham et al., 2020), (b) they are consistent with ratio metrics used in many prior studies (e.g., Hillairet de Boisferon et al., 2018; Lewkowicz & Hansen-Tift, 2012; Pons et al., 2019; Wagner et al., 2018; Wagner, Luyster, Yim, Tager-Flusberg, & Nelson, 2013), and (c) they allow us to compare time spent looking to the eyes and mouth regions. However, note that some previous studies have utilized other variables, such as an eye–mouth index (i.e., time spent looking at the eyes divided by total time looking at eyes and mouth combined; for example, Elsabbagh et al., 2014; Young et al., 2009), to quantify gaze to eye versus mouth regions of a talker. Therefore, we also present analyses based on this index in Supplemental Material.

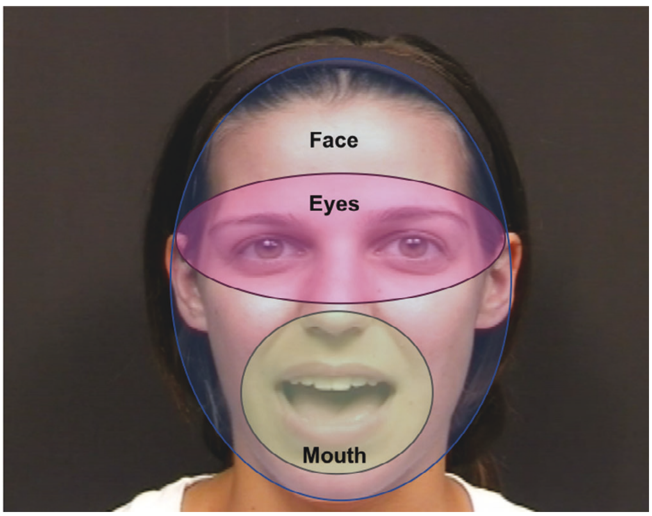


**Fig. 1.** Image of audiovisual speech stimulus video with eyes, mouth, and face areas of interest (AOIs) depicted.

*Parent–child engagement*

Two types of supported joint engagement—LSJE and HSJE—were coded from two 15 min semi-structured parent–child free play (PCFP) sessions. Procedures were identical to those used previously by Bottema-Beutel and colleagues (Bottema-Beutel et al., 2014; Bottema-Beutel, Kim, et al., 2019). During the PCFP sampling sessions, participants and their parents played with a standard set of toys (e.g., a doll, a toy barn and animals, beaded necklaces); parents were instructed to play as they normally do at home using any of the available toys, while taking care not to play with their backs to the camera.

Following the recommendations from Bottema-Beutel, Kim, et al. (2019), the PCFP was collected twice for each participant to increase the stability of derived metrics of supported joint engagement. Each sample was coded by a trained off-site coder naïve to sibling status and study hypotheses. Scores for HSJE and LSJE were averaged across the two coded samples for each participant. Reliability coding was completed on 20 (20%) of the PCFP samples. Intraclass correlation coefficients (ICCs) for HSJE and LSJE were .96 and .76, indicating excellent and good agreement, respectively.

*Prelinguistic vocal complexity*

Infants' prelinguistic vocal complexity was measured using the Communication and Symbolic Behavior Scales Developmental Profile Behavior Sample (CSBS; Wetherby & Prizant, 2002). This standardized, play-based assessment was approximately 30–35 min in duration and was administered by a member of the research team that was trained to fidelity. The CSBS was used as a behavior sampling context and was coded for intentional communication acts using a 5-s partial interval coding system, as in Woynaroski et al. (2017). Intentional communication acts were defined as: (a) nonword vocal or unconventional gestural acts (e.g., moving an object towards an adult and giving an object to an adult) in combination with coordinated attention to object and person; (b) conventional gestures (e.g., distal pointing and head nodding) with attention to adult; and (c) symbolic forms (e.g., nonimitative words and sign language). The intervals coded for intentional communication acts were subsequently coded for the presence/absence of vocalizations including canonical syllables and for the number of different consonants produced by the child. Canonical syllables were defined as a vocalization that consisted of three components: (a) consonantal sound (e.g., /m/, /b/, /t/); (b) fully resonant vowel sound; and (c) transition from consonant to vowel that is rapid and uninterrupted, as judged by a human observer (e.g., "ba" or "ug"). The number of different consonants used within communication acts was coded according to Wetherby's True Consonant Inventory List, which includes selected supraglottal consonants that emerge early and are relatively frequent and easy to code reliably (Wetherby & Prizant, 2002). The two metrics of vocal complexity that were derived from this communication sample were canonical syllabic communication (i.e., the proportion of intentional communication acts that included a canonical syllable) and consonant inventory (i.e., the total number of different consonants used in communication acts from Wetherby's True Consonant Inventory List).

One of every five CSBS samples (10 total; 20%) were randomly selected and coded for interrater reliability. ICCs quantifying interrater reliability for canonical syllabic communication and consonant inventory were .96 and .76, indicating excellent and good reliability, respectively.

*Expressive communication*

Expressive communication was assessed using three measures: the MSEL, a standardized test that assesses development in several domains, including expressive language, for children birth-68 months; the Vineland Adaptive Behavior Scales, second edition (VABS; Sparrow, Cicchetti, & Bella, 2005), a parent-report measure that assesses adaptive function in several domains, including expressive communication for children and adults from birth; and the MacArthur-Bates Communicative Development Inventories: Words and Gestures (MCDI; Fenson, Marchman, Thal, Dale, & Reznick, 2007), a parent report that assesses early-emerging vocabulary and broader language ability in infants. Expressive language age equivalency scores from the MSEL, expressive communication age equivalency scores from the VABS, and the expressive vocabulary raw scores from the MCDI were derived as indices of expressive communication.

### Analytic Plan

First, all variables used in analyses were examined for skewness $> |1|$ and kurtosis $> |3|$ and transformed as necessary to ensure assumptions were met for parametric statistics. MCDI expressive vocabulary and HSJE were square root transformed to correct for positive skew. In keeping with current recommendations (see Enders, 2010), we then imputed missing data (ranging from 0% to 10% missingness across variables used in analyses) using the *missForest* package (Stekhoven & Bühlmann, 2012) in R (R Core Team, 2020).

When indices from measures purported to tap the same construct were sufficiently correlated, aggregate variables were generated to increase their stability and potential predictive validity (Rushton, Brainerd, & Pressley, 1983). Thus, aggregates of prelinguistic vocal complexity (from consonant inventory and canonical syllabic communication; $r = 0.70$) and expressive communication (from all three measures of expressive communication; $r$ values $\geq 0.75$) were generated.

To answer our first research question, a mixed model ANOVA with group (Sibs-autism vs. Sibs-NA) as the between-subjects factor and AOI (mouth vs. eyes) as the within-subjects factor was carried out to evaluate whether eye gaze patterns to audiovisual speech varied by group. To answer our second and third research questions, mediation analyses were carried out using PROCESS (Hayes, 2017). These models statistically tested whether relations between looking to the mouth and eyes and expressive communication was mediated by supported joint engagement (i.e., HSJE or LSJE, tested in separate models) and prelinguistic vocal complexity (for simple mediation models testing indirect effects of each putative mediator in isolation, see Supplementary Material). Cook's $D$ was utilized to monitor for the presence of undue influence (defined as Cook's $D \geq 1$).

The analysis method that we planned to use in evaluating the validity of our measures of child vocal development assumes multivariate normality, and multivariate normality is more likely when univariate distributions do not grossly depart from the normal distribution (Tabachnick & Fidell, 2001). Thus, all variables were evaluated for normality. Variables showing univariate skewness > j1.0j or kurtosis > j3.0j were transformed prior to imputation and analysis.

## RESULTS

### Differences in Eye Gaze Patterns

Results of the 2 (group) × 2 (AOI) ANOVA revealed a main effect of AOI, such that infants looked significantly more to the mouth region compared to the eye region across groups, $F(1, 48) = 4.16, p = .047, \eta_p^2 = .08$ (see Figure 2). There was not a significant effect of group, $F(1, 48) = 0.31, p = .58, \eta_p^2 = .01$, nor was there a significant interaction between group and AOI, $F(1, 48) = 0.54$, $p = .47$, $\eta_p^2 = .01$. There was, however, a high degree of heterogeneity in looking to audiovisual speech across groups. Variance was not significantly different between groups ($p$-value for Levene's test = .34 and .62 for looking to the mouth and eyes, respectively).

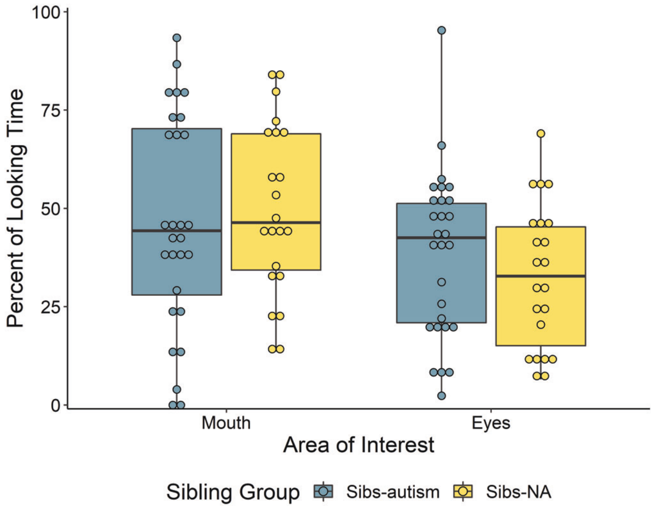


**Fig. 2.** Box plot for time spent looking to the mouth by sibling group. Across groups, infants looked more to the mouth than the eyes, $F(1, 48) = 4.16$, $p = .047$, $\eta_p^2 = .08$. The main effect of group and the interaction between group and area of interest were not significant. Sibs-autism = infant siblings of children diagnosed with autism. Sibs-NA = infant siblings of non-autistic, otherwise typically developing children.

### Direct Associations Between Looking to the Mouth, Expressive Communication, and Putative Mediators

For correlations between all variables of interest across groups, see Table 2 (see Table S1 for correlations between all variables within groups). Although the correlation between looking to the mouth and expressive communication only trended towards significance ($r = .27$, $p = .059$), the putative mediators are significantly associated with both looking to the mouth (zero-order correlations = .30 and .44 for HSJE and prelinguistic communication, respectively; *p*s = .033 and .001) and expressive communication (zero-order correlations = .65 and .76, respectively; *p*s < .001). The lack of a direct association between looking to the mouth and expressive communication does not preclude us from assessing an indirect association (see Hayes, 2009).

### Indirect Associations Between Looking to the Mouth and Expressive Communication

The mediation model testing the indirect effect of looking at the mouth on expressive communication through HSJE and prelinguistic vocal complexity indicated that the relation between looking to the mouth during audiovisual speech and expressive communication was mediated by HSJE and prelinguistic vocal complexity (95% CI = [0.001, 0.009]), consistent with our hypothesis. All paths comprising this indirect effect were statistically significant, with moderate to large effect sizes (see Figure 3).

The relation between looking to the mouth and expressive communication was not mediated by LSJE and prelinguistic vocal complexity (95% CI = [−0.004, 0.001]). In addition, the relation between looking to the eyes and expressive communication was not mediated by HSJE and vocal complexity (95% CI = [−0.009, 0.000]) or LSJE and prelinguistic vocal complexity (95% CI = [−0.002, 0.005]). Thus, the mediated relation appears to be specific to the association between looking to the mouth and expressive communication as mediated by HSJE and prelinguistic vocal complexity.

None of the aforementioned models were moderated by sibling group. In addition, the results of all mediation models (a) were consistent when controlling for the Mullen Early Learning Composite (IQ) and (b) did not vary according to chronological age.

## DISCUSSION

This study sought to examine whether eye gaze patterns, specifically looking to the mouth and eyes during audiovisual

**Table 2**
Zero-order correlations between variables used in analyses across groups

| *Variable* | *1* | *2* | *3* | *4* | *5* | *6* | *7* |
|---|---|---|---|---|---|---|---|
| 1. Looking to the mouth | — | | | | | | |
| 2. Looking to the eyes | −.81*** | — | | | | | |
| 3. HSJE | .30* | −.19 | — | | | | |
| 4. LSJE | .24 | −.33* | .26 | — | | | |
| 5. Prelinguistic vocal complexity | .44** | −.22 | .60*** | −.01 | — | | |
| 6. Expressive communication | .27 | −.20 | .65*** | .13 | .76*** | — | |
| 7. MSEL ELC | .05 | −.02 | −.16 | .25 | −.15 | −.07 | — |

*Note*: HSJE, higher-order supported joint engagement; LSJE, lower-order supported joint engagement, Prelinguistic Vocal Complexity = aggregate of canonical syllabic communication and consonant inventory from the Communication and Symbolic Behavior Scales Developmental Profile Behavior Sample (CSBS; Wetherby & Prizant, 2002), Expressive communication = aggregate of Mullen Scales of Early Learning (MSEL; Mullen, 1995) expressive language age equivalency score, Vineland Adaptive Behavior Scales, second edition (VABS; Sparrow et al., 2005) expressive communication age equivalency score, and the MacArthur-Bates Communicative Development Inventory: Words and Gestures (MCDI; Fenson et al., 2007) number of words the child says, ELC = Early Learning Composite.
*$p < .05$, **$p < .01$, ***$p < .001$.

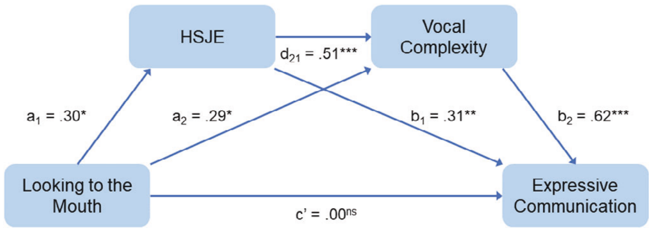


**Fig. 3.** Indirect effect of looking to the mouth on expressive communication through higher-order supported joint engagement (HSJE) and prelinguistic vocal complexity. $a_1$ = the relation between looking to the mouth and HSJE, not controlling for any other factors. $a_2$ = the relation between looking to the mouth and prelinguistic vocal complexity, not controlling for any other factors. $b_1$ = the relation between HSJE and expressive communication, controlling for looking to the mouth. $b_2$ = the relation prelinguistic vocal complexity and expressive communication, controlling for looking to the mouth. $d_{21}$ = the relation between HSJE and prelinguistic vocal complexity, controlling for looking to the mouth. c′ = the direct effect of looking to the mouth on expressive communication (i.e., the c′ path), controlling for HSJE and prelinguistic vocal complexity. Note that c′ is non-significant, meaning that the association between looking to the mouth and expressive communication is completely mediated by HSJE and prelinguistic vocal complexity. This indirect effect is not moderated by sibling group. All values are standardized coefficients. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$, *ns* = nonsignificant result.

speech (a) differed between younger siblings of children with and without autism at 6–18 months of age and (b) indirectly influenced expressive communication via parent–child engagement and child prelinguistic vocal complexity. In addition, we tested whether the aforementioned indirect effects varied according to a sibling group. Results suggest that looking behavior in infancy, on average, does not differentiate infants at high versus general population-level likelihood for autism. Looking to the mouth of a talker, however, may be useful for explaining individual differences in expressive communication, through its influence on one aspect of parent–child engagement—HSJE—and child prelinguistic vocal complexity.

### Failure to Find Between-Group Differences in Early Eye Gaze Patterns

Our null result for between-group differences in looking patterns during audiovisual speech was somewhat surprising. It is notable, however, that several studies which have focused on infants to date have failed to specifically find between-group differences in gaze to mouth and eye regions during audiovisual speech for infant siblings of children with autism in comparison to infant siblings of children without autism, particularly when using proportion metrics similar to those utilized in this study (i.e., proportion of time spent looking to the mouth to time spent looking to the face; e.g., Elsabbagh et al., 2014; Shic, Macari, & Chawarska, 2014; Young et al., 2009). The past studies that have found differences in looking at the mouth versus the eyes during audiovisual speech have mostly focused on children with autism and nonautistic comparison groups (e.g., Chawarska et al., 2014; Hosozawa, Tanaka, Shimizu, Nakano, & Kitazawa, 2012; Johnels, Gillberg, Falck-Ytter, & Miniscalco, 2014; Nakano et al., 2010), although a small number of studies have found differences in early eye gaze patterns between infant sibling groups (e.g., differences in looking to social scenes, people, and faces, as well as differences in scanning inner versus outer features of faces during speech; Chawarska et al., 2012; Chawarska, Macari, & Shic, 2013; Shic et al., 2014). Thus, although the finding of nonsignificant differences between groups did not support our a priori hypothesis, it is generally consistent with findings from previous studies of at-risk infants.

### Looking to the Mouth, Not Eyes, Is Positively Associated with Putative Mediators

Using a brief, passive remote eye-tracking task, we found moderately sized positive associations between looking to the mouth during audiovisual speech and our two putative mediators. Specifically, looking to the mouth was positively associated with HSJE. This specific form of supported joint engagement, wherein parents support the child's engagement and the child acknowledges the caregiver in the course of play, has been previously linked with language and communication development in young children with autism (Bottema-Beutel et al., 2014; Bottema-Beutel, Woynaroski et al., 2019; Crandall et al., 2019). The present study provides increased support for parsing out this particular form of supported joint engagement from LSJE, for which associations in our study did not hold.

In addition, looking to the mouth was moderately associated with the index of prelinguistic vocal complexity (comprised of canonical syllabic communication and consonant inventory) across the infants in our study sample. These results extend past work that has linked looking at the mouth to language and communication skills in infants at increased likelihood for autism, children with autism, and autistic adults (e.g., Chawarska et al., 2012; Klin et al., 2002; Norbury et al., 2009; Tenenbaum et al., 2014; Young et al., 2009) and provide empirical support for the view that increased looking to audiovisual speech cues is linked with qualitative changes in prelinguistic development that are observed early in life (Lewkowicz & Hansen-Tift, 2012).

Looking to the eyes was not positively associated with HSJE, prelinguistic vocal complexity, or expressive communication. Previous work has found evidence for concurrent associations in infants between looking to the eyes and communication, specifically positive correlations with concurrent and later social communication and negative correlations with later expressive communication in nonautistic infants (Pons et al., 2019) and in Sibs-autism and Sibs-NA (Wagner et al., 2018), although these studies collected their eye-tracking data at a narrower window of time (i.e., within a 1 month age range; this study utilized a 12 month age range). Thus, additional work is needed to determine how looking to the eyes relates to various aspects of communication across the first and second year of life.

### Looking to the Mouth Influences Expressive Communication Via Parent and Child Variables

Mediation analyses indicated that looking to the mouth influences expressive communication indirectly through its associations with supported joint engagement, specifically HSJE, and child prelinguistic vocal complexity. Although a direct link between looking to the mouth on concurrent expressive communication has been observed in several prior studies of children with autism (e.g., Chawarska et al., 2012; Habayeb et al., 2020; Klin et al., 2002; Norbury et al., 2009) and longitudinally in infants at increased likelihood for autism (e.g., Chawarska et al., 2012, 2013; Shic et al., 2014), it is notable that this indirect effect was "complete" in our study, meaning that the direct link between these variables was nonsignificant when controlling for HSJE and prelinguistic vocal complexity. In addition, this indirect effect did not vary according to sibling group. These findings provide empirical support for the notion that parent–child engagement and prelinguistic vocal behavior are mechanisms by which early eye gaze may influence expressive communication development in infants, regardless of sibling status. The present study provides some support for our hypothesis that attention to a speaker's face, specifically the speaker's mouth, may signal developmental readiness for more HSJE, which subsequently scaffolds increased prelinguistic vocal complexity in the child and translates to gains in his/her broader expressive communication. To our knowledge, these findings are the first to statistically test and find evidence in support of this theoretical possibility.

Notably, the indirect relation does not hold in models that include (a) LSJE as a mediator or (b) looking to the eyes as the predictor. In addition, the mediation models are still significant when a different looking metric is utilized (i.e., eye–mouth index; see Supplementary Materials) and when cognitive ability (i.e., Mullen Early Learning Composite) is included as a covariate. Thus, our hypothesis is strengthened by the specificity and robustness of these results.

### Limitations and Future Directions

Our study is limited in several ways. First, the present investigation is limited by our use of a concurrent correlational design. Although we were able to provide preliminary support for mediated associations between early eye gaze patterns and expressive communication, further research is needed to determine whether an early-emerging tendency to look to the mouth during audiovisual speech temporally precedes, or is causally related to, qualitative changes in expressive communication, through a cascade of changes in HSJE and child vocal complexity. Causality could be rigorously tested in the context of a clinical trial that attempts to intervene upon or directly target early eye gaze and systematically measures the effects of treatment on looking behavior and HSJE, as well as more distal prelinguistic and linguistic outcomes in infant siblings of children with autism.

Second, we cannot presently determine how the present findings apply to infant siblings of children with autism who do or do not go on to be diagnosed with autism. Nonetheless, our findings provide novel insights into the possible mechanisms underlying the replicated relation between looking to a talker's mouth and communication. Our innovative approach, focused on the elucidation of predictors of skills within discrete developmental domains (e.g., continuously quantified communication skill), has the potential to facilitate early identification and preventative intervention for infants known to be at increased likelihood of autism and other communication delays and disorders.

Third, some recent evidence suggests that gaze patterns obtained during live interactions may not correlate with looking during prerecorded videos, at least in adolescents with autism (Grossman, Zane, Mertens, & Mitchell, 2019). Given that our stimuli were specifically designed to limit head movements to accommodate static AOIs, it is possible that our looking task may not represent looking behaviors during live interactions. Some technologies do exist that may provide increased ecological validity to eye gaze paradigms in infants at increased likelihood of autism (Edmunds et al., 2017). These methods should be leveraged to explore the convergent validity between virtual and in vivo measures of eye gaze in infants at high and low likelihood for autism in future work.

Finally, additional research is needed to ascertain the neural substrates underlying the replicated relations between eye gaze patterns and linguistic development in children at increased likelihood for, or diagnosed with, autism. Looking to a talker's mouth provides infants with access to concurrent visual and acoustic articulatory cues whose dynamic properties are not only highly synchronized over time but also correspond in terms of variations in their intensity and prosody (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009). Access to such redundant audiovisual speech cues has been shown to increase the efficiency of speech processing as indexed by event-related potentials in nonautistic adults and children, as well as in children with autism (Knowland, Mercure, Karmiloff-Smith, Dick, & Thomas, 2014; van Wassenhove et al., 2005; Woynaroski et al., 2019). Ongoing work in our laboratory is exploring whether such brain-based measures of audiovisual speech processing may explain the associations that we have observed between looking behavior and expressive communication development in infancy.

## CONCLUSION

This study provides preliminary support for the hypothesis that infants' gaze patterns influence expressive communication via increased supported joint engagement which, in turn, scaffolds prelinguistic development (i.e., canonical syllables and for the number of different consonants produced by the child). This indirect effect is specific to the link between looking to the mouth during audiovisual speech and expressive communication as mediated by HSJE and prelinguistic communication. Additional research is needed to further characterize and increase confidence in the mediated relations observed, particularly longitudinally and with consideration to later diagnosis of autism, and to advance us towards the application of these findings in clinical practice.

*Acknowledgments*—This work was supported by NIH/NIDCD R21DC016144 (PI: Woynaroski), NIH/NCATS KL2TR000446 (PI: Woynaroski), NIH U54 HD083211 (PI: Neul), NSF NRT grant DGE 19-22697 (PI: Wallace), Vanderbilt Institute for Clinical and Translational Research award VR21082 (PI: Santapuram), a Vanderbilt Undergraduate Summer Research Program award to Pooja Santapuram (Primary Mentor: Woynaroski), and the Nancy Lurie Marks Family Foundation. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies. We would like to thank the Adventure Science Center of Nashville for their partnership in carrying out this work and all of the families for their participation in this project.

## SUPPORTING INFORMATION

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

**Appendix S1:** Supporting Information

## NOTE

1 There is some debate in the literature regarding whether researchers should utilize person-first language (e.g., individuals with autism) versus identity-first language (e.g., autistic individuals; see Robison, 2019). In this article, we refer to children with autism using person-first language and honor autistic adults' recent requests (Gernsbacher, 2017; Kenny et al., 2016) for the use of identity-first language while otherwise avoiding ableist language (Bottema-Beutel, Kapp, Lester, Sasson, & Hand, 2021), and acknowledge that this approach may not be received equally by all of the stakeholders of our research.

## REFERENCES

Adamson, L. B., Bakeman, R., Deckner, D. F., & Romski, M. (2009). Joint engagement and the emergence of language in children with autism and Down syndrome. *Journal of Autism and Developmental Disorders*, *39*(1), 84–96. https://doi.org/10.1007/s10803-008-0601-7

American Psychiatric Association. (2013) *Diagnostic and statistical manual of mental disorders-5*. Arlington, VA: American Psychiatric Association.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*(4), 1278–1289. https://doi.org/10.2307/1129997

Billstedt, E., Carina Gillberg, I., & Gillberg, C. (2007). Autism in adults: Symptom patterns and early childhood predictors. Use of the DISCO in a community sample followed from childhood. *Journal of Child Psychology and Psychiatry*, *48*(11), 1102–1110. https://doi.org/10.1111/j.1469-7610.2007.01774.x

Bottema-Beutel, K., Kapp, S. K., Lester, J. N., Sasson, N. J., & Hand, B. N. (2021). Avoiding ableist language: Suggestions for autism researchers. *Autism in Adulthood*, *3*(1), 18–29. https://doi.org/10.1089/aut.2020.0014

Bottema-Beutel, K., Kim, S. Y., Crowley, S., Augustine, A., Kecili-Kaysili, B., Feldman, J. I., & Woynaroski, T. G. (2019). The stability of joint engagement states in infant siblings of children with and without ASD: Implications for measurement practices. *Autism Research*, *12*, 495–504. https://doi.org/10.1002/aur.2068

Bottema-Beutel, K., Woynaroski, T., Louick, R., Keefe, E. S., Watson, L. R., & Yoder, P. J. (2019). Longitudinal associations across vocabulary modalities in children with autism and typical development. *Autism*, *23*(2), 424–435. https://doi.org/10.1177/1362361317745856

Bottema-Beutel, K., Yoder, P. J., Hochman, J. M., & Watson, L. R. (2014). The role of supported joint engagement and parent utterances in language and social communication development in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(9), 2162–2174. https://doi.org/10.1007/s10803-014-2092-z

Calvert, G. A., Brammer, M. J., & Iversen, S. D. (1998). Crossmodal identification. *Trends in Cognitive Sciences*, *2*(7), 247–253. https://doi.org/10.1016/S1364-6613(98)01189-9

Campbell, D. J., Shic, F., Macari, S., & Chawarska, K. (2014). Gaze response to dyadic bids at 2 years related to outcomes at 3 years in autism spectrum disorders: A subtyping analysis. *Journal of Autism and Developmental Disorders*, *44*(2), 431–442. https://doi.org/10.1007/s10803-013-1885-9

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

Charman, T., Young, G. S., Brian, J., Carter, A., Carver, L. J., Chawarska, K., … Zwaigenbaum, L. (2017). Non-ASD outcomes at 36 months in siblings at familial risk for autism spectrum disorder (ASD): A baby siblings research consortium (BSRC) study. *Autism Research*, *10*(1), 169–178. https://doi.org/10.1002/aur.1669

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry*, *53*(8), 903–913. https://doi.org/10.1111/j.1469-7610.2012.02538.x

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, *74*(3), 195–203. https://doi.org/10.1016/j.biopsych.2012.11.022

Chawarska, K., Shic, F., Macari, S., Campbell, D. J., Brian, J., Landa, R., … Bryson, S. (2014). 18-month predictors of later outcomes in younger siblings of children with autism spectrum disorder: A baby siblings research consortium study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(12), 1317–1327. https://doi.org/10.1016/j.jaac.2014.09.015

Costanzo, V., Chericoni, N., Amendola, F. A., Casula, L., Muratori, F., Scattoni, M. L., & Apicella, F. (2015). Early detection of autism spectrum disorders: From retrospective home video studies to prospective 'high risk'sibling studies. *Neuroscience & Biobehavioral Reviews*, *55*, 627–635. https://doi.org/10.1016/j.neubiorev.2015.06.006

Crandall, M. C., Bottema-Beutel, K., McDaniel, J., Watson, L. R., & Yoder, P. J. (2019). Children with autism spectrum disorder may learn from caregiver verb input better in certain engagement states. *Journal of Autism and Developmental Disorders*, *49*(8), 3102–3112. https://doi.org/10.1007/s10803-019-04041-w

Dunham, K., Feldman, J. I., Liu, Y., Cassidy, M., Conrad, J. G., Santapuram, P., … Woynaroski, T. G. (2020). Stability of variables derived from measures of multisensory function in children with autism spectrum disorder. *American Journal of Intellectual and Developmental Disabilities*, *125*(4), 287–303. https://doi.org/10.1352/1944-7558-125.4.287

Edmunds, S. R., Rozga, A., Li, Y., Karp, E. A., Ibanez, L. V., Rehg, J. M., & Stone, W. L. (2017). Brief report: Using a point-of-view camera to measure eye gaze in young children with autism spectrum disorder during naturalistic social interactions: A pilot study. *Journal of Autism and Developmental Disorders*, *47*, 898–904. https://doi.org/10.1007/s10803-016-3002-3

Eisenberg, L. (1956). The autistic child in adolescence. *American Journal of Psychiatry*, *112*(8), 607–612. https://doi.org/10.1176/ajp.112.8.607

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., Johnson, M. H., & The Basis Team (2014). What you see is what you get: Contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience*, *9*(4), 538–543. https://doi.org/10.1093/scan/nst012

Enders, C. K. (2010) *Applied missing data analysis*. New York, NY: Guilford Press.

Feldman, J. I., Raj, S., Santapuram, P., Golden, A. J., Daly, C., Dunham, K., … Woynaroski, T. G. (2021). Sensory responsiveness is linked with communication in infants siblings of children with and without autism. *Journal of Speech, Language, and Hearing Research*, *64*(6), 1964–1976. https://doi.org/10.1044/2021_JSLHR-20-00196

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual*. Baltimore, MD: Brookes.

Gernsbacher, M. A. (2017). Editorial perspective: The use of person-first language in scholarly writing may accentuate stigma. *Journal of Child Psychology and Psychiatry*, *58*(7), 859–861. https://doi.org/10.1111/jcpp.12706

Grossman, R. B., Zane, E., Mertens, J., & Mitchell, T. (2019). Facetime vs. screentime: Gaze patterns to live and video social stimuli in adolescents with ASD. *Scientific Reports*, *9*, 12643. https://doi.org/10.1038/s41598-019-49039-7

Gulsrud, A. C., Hellemann, G., Shire, S., & Kasari, C. (2016). Isolating active ingredients in a parent-mediated social communication intervention for toddlers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *57*(5), 606–613. https://doi.org/10.1111/jcpp.12481

Habayeb, S., Tsang, T., Saulnier, C., Klaiman, C., Jones, W., Klin, A., & Edwards, L. A. (2020). Visual traces of language acquisition in toddlers with autism spectrum disorder during the second year of life. *Journal of Autism and Developmental Disorders*, *51*, 2519–2530. https://doi.org/10.1007/s10803-020-04730-x

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, *198*(4319), 853–855. https://doi.org/10.1126/science.918670

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2017) *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. (2nd ed.). New York, NY: Guilford Press.

Hillairet de Boisferon, A., Tift, A. H., Minar, N. J., & Lewkowicz, D. J. (2018). The redeployment of attention to the mouth of a talking face during the second year of life. *Journal of Experimental Child Psychology*, *172*, 189–200. https://doi.org/10.1016/j.jecp.2018.03.009

Hosozawa, M., Tanaka, K., Shimizu, T., Nakano, T., & Kitazawa, S. (2012). How children with specific language impairment view social situations: An eye tracking study. *Pediatrics*, *129*(6), e1453–e1460. https://doi.org/10.1542/peds.2011-2278

Johnels, J., Gillberg, C., Falck-Ytter, T., & Miniscalco, C. (2014). Face-viewing patterns in young children with autism spectrum disorders: Speaking up for the role of language comprehension. *Journal of Speech, Language, and Hearing Research*, *57*(6), 2246–2252. https://doi.org/10.1044/2014_JSLHR-L-13-0268

Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: Randomized comparison of joint attention and play interventions. *Journal of Consulting and Clinical Psychology*, *76*(1), 125–137. https://doi.org/10.1037/0022-006X.76.1.125

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*, 809–816. https://doi.org/10.1001/archpsyc.59.9.809

Knowland, V. C. P., Mercure, E., Karmiloff-Smith, A., Dick, F., & Thomas, M. S. (2014). Audio-visual speech perception: A developmental ERP investigation. *Developmental Science*, *17*(1), 110–124. https://doi.org/10.1111/desc.12098

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*, *109*(5), 1431–1436. https://doi.org/10.1073/pnas.1114783109

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. L. (2012) *Autism diagnostic observation schedule, second edition (ADOS-2) manual (Part I): Modules 1–4*. Los Angeles, CA: Western Psychological Services.

Maurer, D., & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development*, *47*(2), 523–527. https://doi.org/10.2307/1128813

McDaniel, J., Woynaroski, T. G., Keceli-Kaysili, B., Watson, L. R., & Yoder, P. (2019). Vocal communication with canonical syllables predicts later expressive language skills in preschool-aged children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, *62*(10), 3826–3833. https://doi.org/10.1044/2019_JSLHR-L-19-0162

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., … Sigman, M. (2013). Beyond autism: A baby siblings research consortium study of high-risk children at three years of age. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(3), 300–308. https://doi.org/10.1016/j.jaac.2012.12.011

Middelweerd, M. J., & Plomp, R. (1987). The effect of speechreading on the speech-reception threshold of sentences in noise. *Journal of the Acoustical Society of America*, *82*(6), 2145–2147. https://doi.org/10.1121/1.395659

Mullen, E. M. (1995) *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service.

Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., … 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: Biological Sciences*, *277*(1696), 2935–2943. https://doi.org/10.1098/rspb.2010.0587

Norbury, C., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry*, *50*(7), 834–842. https://doi.org/10.1111/j.1469-7610.2009.02073.x

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., … Stone, W. L. (2011, Sep). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics*, *128*(3), e488–e495. https://doi.org/10.1542/peds.2010-2825

Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., … Stone, W. L. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: A baby siblings research consortium study. *Journal of Child Psychology and Psychiatry*, *56*(9), 988–998.

Pons, F., Bosch, L., & Lewkowicz, D. J. (2019). Twelve-month-old infants' attention to the eyes of a talking face is associated with communication and social skills. *Infant Behavior and Development*, *54*, 80–84. https://doi.org/10.1016/j.infbeh.2018.12.003

R Core Team (2020). R: A language and environment for statistical computing (Version 4.0.2). Vienna, Austria. https://www.R-project.org/

Robison, J. E. (2019). Talking about autism—Thoughts for researchers. *Autism Research*, *12*(7), 1004–1006. https://doi.org/10.1002/aur.2119

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin*, *94*, 18–38. https://doi.org/10.1037/0033-2909.94.1.18

Rutter, M., Bailey, A., & Lord, C. (2003) *Social Communication Questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.

Saul, J., & Norbury, C. (2020). Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment? *Autism*, *24*, 1217–1231. https://doi.org/10.1177/1362361319898560

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

Sparrow, S. S., Cicchetti, D. V., & Bella, D. A. (2005) *Vineland adaptive behavior scales*. (2nd ed.). Circle Pines, MN: AGS.

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*, 112–118. https://doi.org/10.1093/bioinformatics/btr597

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*, 212–215. https://doi.org/10.1121/1.1907309

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism, Handbook of autism and pervasive developmental disorders. In F. R. Volkmar, R. Paul, A. Klin & D. Cohen (Eds.), *Diagnosis, development, neurobiology, and behavior*. (Vol. *1*,3rd ed., pp. 335–364). Hoboken, NJ, US: John Wiley & Sons Inc.

Tenenbaum, E. J., Amso, D., Abar, B., & Sheinkopf, S. J. (2014). Attention and word learning in autistic, language delayed and typically developing children. *Frontiers in Psychology*, *5*, 490. https://doi.org/10.3389/fpsyg.2014.00490

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy*, *18*(4), 534–553. https://doi.org/10.1111/j.1532-7078.2012.00135.x

Turner, L. M., & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry*, *48*(8), 793–802.

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

Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry*, *33*(3), 489–507. https://doi.org/10.1111/j.1469-7610.1992.tb00887.x

Wagner, J. B., Luyster, R. J., Moustapha, H., Tager-Flusberg, H., & Nelson, C. A. (2018). Differential attention to faces in infant siblings of children with autism spectrum disorder and associations with later social and language ability. *International Journal of Behavioral Development*, *42*(1), 83–92. https://doi.org/10.1177/0165025416673475

Wagner, J. B., Luyster, R. J., Yim, J. Y., Tager-Flusberg, H., & Nelson, C. A. (2013). The role of early visual attention in social development. *International Journal of Behavioral Development*, *37*(2), 118–124. https://doi.org/10.1177/0165025412468064

Wetherby, A., & Prizant, B. M. (2002). *Communication and symbolic behavior scales developmental profile: First normed edition*. Baltimore, MD: Brookes.

Wetherby, A., Watt, N., Morgan, L., & Shumway, S. (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *Journal of Autism and Developmental Disorders*, *37*(5), 960–975. https://doi.org/10.1007/s10803-006-0237-4

Woynaroski, T. G., Feldman, J. I., Edmunds, S. R., Simon, D. M., Tu, A., Kuang, W., … Wallace, M. T. (2019) Audiovisual speech processing and attention are linked with language in children with and without autism *[paper presentation]*. Baltimore, MD, United States: Society for Research in Child Development.

Woynaroski, T. G., Oller, D. K., Keceli-Kaysili, B., Xu, D., Richards, J. A., Gilkerson, J., … Yoder, P. J. (2017). The stability and validity of automated vocal analysis in minimally verbal preschoolers with autism spectrum disorder. *Autism Research*, *10*, 508–519.

Woynaroski, T. G., Watson, L. R., Gardner, E., Newsom, C. R., Keceli-Kaysili, B., & Yoder, P. J. (2016). Early predictors of growth in diversity of key consonants used in communication in initially preverbal children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 1013–1024. https://doi.org/10.1007/s10803-015-2647-7

Woynaroski, T. G., Yoder, P., Fey, M., & Warren, S. (2014). A transactional model of spoken vocabulary variation in toddlers with intellectual disabilities. *Journal of Speech, Language, & Hearing Research*, *57*, 1754–1763.

Yoder, P., Watson, L. R., & Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *45*(5), 1254–1270. https://doi.org/10.1007/s10803-014-2286-4

Young, G., 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