Original Article

autism

Autism
1–10

DOI: 10.1177/1362361320968641
journals.sagepub.com/home/aut
SAGE

Developmental associations between joint engagement and autistic children's vocabulary: A cross-lagged panel analysis

Kristen Bottema-Beutel[1], So Yoon Kim[1], Shannon Crowley[1] and Paul J Yoder[2]

Abstract

Cross-lagged panel analysis was used to examine associations between two joint engagement variables; higher order supported joint engagement and higher order supported joint engagement that co-occurs with caregiver's follow-in talk (higher order supported joint engagement + follow-in), and expressive and receptive vocabulary in a group of young autistic children ($n = 91$) with language delay (mean chronological age = 39 months). Variables were measured twice, 8 months apart. Coefficients for cross-lagged variable pairs were derived from structural equation models. Early higher order supported joint engagement was significantly associated with later expressive and receptive vocabulary ($bs = 0.18$ and 0.26, respectively), and early higher order supported joint engagement + follow-in was significantly associated with later expressive and receptive vocabulary ($bs = 0.14$ and 0.15, respectively). Associations between early vocabulary and later joint engagement were not significant. Linear contrasts between cross-lagged associations did not show a significantly superior association for any early joint engagement variables and later vocabulary variables. However, our results suggest that higher order supported joint engagement and higher order supported joint engagement + follow-in may be useful initial intervention targets for developmental interventions aimed at promoting autistic children's language development who are initially language delayed.

Lay abstract

In this study, we used a *cross-lagged panel analysis* to examine correlations over time between two types of engagement between children and their parents and children's later expressive and receptive vocabularies. This kind of design can help researchers understand which early developmental achievements "drive" later developmental achievements. It is important for intervention researchers to know which developmental achievements happen first, so that they can set intervention goals appropriately. The two joint engagement variables we examined were (a) higher order supported joint engagement, which occurs when caregivers influence their child's play with toys and the child reciprocally responds to the caregiver, but does not manage the interaction by shifting gaze between the toys and the caregiver, and (b) higher order supported joint engagement that co-occurs with caregiver's follow-in talk (higher order supported joint engagement + follow-in). Follow-in talk occurs when the caregiver talks about objects and events that the child is focused on. Ninety-one autistic children ($n = 91$) with language delay (mean chronological age = 39 months) participated, along with their primary caregivers. Each of the four variables was measured twice, 8 months apart. Our statistical procedures showed that early higher order supported joint engagement and early higher order supported joint engagement + follow-in were significantly associated with later expressive and receptive vocabulary. In contrast, associations between early vocabulary variables and later joint engagement variables were not significant. Our results suggest that higher order supported joint engagement and higher order supported joint engagement + follow-in may be useful initial intervention targets, for developmental interventions aimed at promoting language development in autistic children who are initially language delayed.

[1]Boston College, USA
[2]Vanderbilt University, USA

Corresponding author:
Kristen Bottema-Beutel, Lynch School of Education and Human Development, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA.
Email: kristen.bottema-beutel@bc.edu

Keywords
autism, caregiver–child interaction, follow-in talk, joint engagement, language, responsivity, vocabulary

Introduction

The ability to initiate and respond to bids for joint attention is critical for early language development, especially for autistic[1] children (Bottema-Beutel, 2016; Mundy, 2016; Yoder et al., 2015). When children engage in joint attention, which occurs when they shift their attention between referents of interest (e.g. objects or events) and their communication partners, they participate in a social-communicative context that facilitates linking words with meaning (Mundy & Jarrold, 2010) and developing interactive competence. Because autistic children have social-communication difficulties, any amount of joint attention routines they participate in may be critical for linguistic development. In contrast, non-autistic children participate in joint attention routines in abundance. This means that even if there is some variation in their joint attention abilities, it does not negatively impact their language development (Bottema-Beutel, 2016).

While some researchers conceptualize joint attention as a discrete set of child skills, Bakeman and Adamson (1984) describe joint *engagement states* that are constituted by mutual and extended caregiver–child interactions in the context of toy play. Engagement states are dyadic, capturing both child and caregiver contributions to the maintenance of ongoing interaction. Joint engagement (a dyadic state) may be a more useful developmental construct than joint attention (a child skill) for intervention researchers, because it invites attention to how caregivers and children mutually co-construct interactions that make joint attention to objects and events relevant. Interventions can then be designed around facilitating joint engagement between caregiver and child (e.g. Kasari et al., 2006). Alternatively, focusing on joint attention as a child skill can result in interventions that are overly focused on prompting the child to demonstrate joint attention abilities outside of meaningful interactive contexts, which is likely not helpful for language development. Recent research corroborates the developmental utility of the joint engagement construct, as it is a superior predictor of expressive language even when children's joint attention skills are accounted for (Adamson et al., 2019).

There are several different types of joint engagement, which appear at different stages in children's development (Bakeman & Adamson, 1984). *Coordinated joint engagement* (CJE) is the most developmentally advanced, and occurs when the caregiver influences the child's play with toys, and the child manages the interaction with the caregiver through explicit signaling of mutual attention such as by shifting eye gaze between the toys and the caregiver throughout the exchange. *Supported joint engagement* (SJE) appears earlier in development, and occurs when the caregiver influences the child's play with toys, but differs from CJE because in this state, the child does not explicitly manage the interaction through eye gaze to the caregiver.

During these engagement states, caregivers can add a symbolic dimension, either through pretense (e.g. using a toy to symbolize another object) or by communicating with the child via words or iconic gestures (e.g. by mimicking the use of a hammer to symbolize the act of hammering). Importantly, research has shown that symbol-infused SJE is a predictor of later vocabulary development, and even more so than symbol-infused CJE (Bakeman & Adamson, 1984). This was a surprising finding, as CJE is the more developmentally complex state for the child. The superiority of SJE over CJE in predicting later language is consistent for both non-autistic and autistic children (Adamson et al., 2009). One possible explanation for this finding is that SJE, which does not require the child to manage the interaction by shifting gaze between their caregiver and the toys, allows the child to allocate more cognitive resources for mapping their caregiver's words onto their referents than does CJE (Bloom et al., 2001). Consistent with research in non-autistic children, this line of research has contributed to a now widely adopted viewpoint that the interactive context in which caregivers talk to their children, and the opportunities for continued interaction that it produces, is important for determining the extent to which caregiver talk will facilitate children's language development (Bottema-Beutel et al., 2014; Crandall et al., 2019; Ramírez et al., 2020).

Higher order supported joint engagement and follow-in talk

In our research, we have made two adjustments to the conceptualization of symbol-infused SJE in an effort to further specify the types of caregiver–child interactions that are most predictive of autistic children's later language. First, we specify a state in which caregivers influence children's toy play and children display reciprocity with the caregiver through such actions as turn taking or imitating the caregiver (but are not managing these exchanges with gaze shifts to the caregiver's face). We refer to this as *higher order supported joint engagement* (HSJE). Children's reciprocity signals a level of coordination and involvement in the interaction that incorporates the caregiver as an agentive interaction partner and not simply an outside force operating on the toys. Furthermore, reciprocal interactions, such as when the child imitates their caregiver, are predictive of autistic children's later social functioning (Bottema, Kim, & Crowley, 2019). This suggests that reciprocity in interactions may be a key mechanism for developing social

and communicative competence. In sum, HSJE differs from CJE because in HSJE the child does not coordinate their attention with the caregiver through gaze to the caregiver's face, and HSJE differs from SJE because it includes *only* instances within SJE where the child is reciprocally interacting with the caregiver.

Second, we refine "symbol infusion" to specify instances when caregivers talk about what the child is doing or attending to, referred to as follow-in (FI) talk. This form of talk has been shown in multiple studies to predict autistic children's later expressive and receptive language, while talk about the caregiver's own focus of attention does not (e.g. McDuffie & Yoder, 2010; Siller & Sigman, 2002). FI talk may be especially beneficial for children's language development because, as in SJE, children are not required to shift their attention to make word-meaning connections. FI talk may also support development because it ensures that caregivers will pursue interactions around objects and activities in which children already show an interest. In prior work, HSJE was found to be a superior predictor of expressive language as compared to a form of SJE in which the child was not displaying reciprocity with the caregiver. Likewise, HSJE that co-occurred with FI talk (HSJE + FI) was found to be a superior predictor of receptive language development as compared to SJE that included FI talk, but did not involve reciprocity with the caregiver (Bottema-Beutel et al., 2014).

Our interest in homing in on specific forms of caregiver–child joint engagement that are developmentally associated with autistic children's language is motivated by promising (yet mixed) results from intervention research aiming to leverage joint engagement as an intervention context (e.g. Green et al., 2010; Kasari et al., 2006; Sandbank et al., 2020; Schreibman et al., 2015). The assumption in this research is that joint engagement between children and their caregivers may be the mechanism by which children's language and communication growth occurs (Green et al., 2010; Rahman et al., 2016). However, much of this research has been conducted with little specificity in regards to what is considered "joint engagement," with several different types often lumped into a single intervention target (see, e.g. Gulsrud et al., 2016; Kasari et al., 2006, which combines SJE and CJE into a single "joint engagement" variable). We argue that joint engagement should be conceptualized as consisting of several distinct types (as opposed to being combined into a single type) because only some highly specific forms are actually associated with autistic children's later language development (Adamson et al., 2009; Bottema-Beutel et al., 2014).

Research on caregiver talk as a means to improve child outcomes is also mixed. One well-designed primary study showed that caregiver's use of responsive talk was a mediator of intervention effects on autistic children (Pickles et al., 2015). However, a recent meta-analysis showed that synthesized intervention effects of interventions promoting responsive caregiver talk were significant and positive in regards to parents' use of responsive talk, but null in regards to downstream child gains (Edmunds et al., 2019). Continuing to examine the role of highly specified forms of joint engagement and caregiver talk (e.g. HSJE and HSJE + FI) will offer empirical evidence for streamlining these constructs within intervention programming.

Reciprocal influences of joint engagement and children's language

At present, there are open questions regarding the temporal ordering of emerging language and joint engagement development in autistic children. While evidence suggests that joint engagement impacts later language development (see Bottema-Beutel, 2016 and Mundy, 2016 for a meta-analysis and comprehensive summary, respectively), children's language may also impact joint engagement. For example, it is possible that children's emerging expressive vocabulary influences the caregiver–child dyad's propensity and capacity to engage in HSJE. When children are beginning to build expressive vocabularies, their word use may be a strong recruiter of caregiver scaffolding around play. Indeed, there has been at least one study suggesting that this may be the case. Adamson and colleagues (2019) showed that, in a sample of young children who screened as particularly likely to develop autism (via the M-CHAT-R/F; Robins et al., 2009), children's expressive language (defined as either talking or not talking) accounted for 45% of the variance in children's concurrent symbol-infused joint engagement with caregivers. In a longitudinal analysis of the same participants, they found that children who were not talking at the first visit but had started talking at the second visit increased SJE to a greater degree than children who were not talking at either visit. In contrast, research has shown associations between early HSJE + FI and later receptive language, suggesting that the opposite pathway may be particularly important for this pair of constructs (Bottema-Beutel et al., 2014; Bottema-Beutel, Woynaroski, et al., 2019).

If we are able to characterize the order of influence for joint engagement constructs and vocabulary, we can better plan developmental interventions that rely on sequenced developmental trajectories to derive intervention goals (Schreibman et al., 2015). That is, identifying the strongest longitudinal pathways can help intervention researchers know where to start. For example, if early HSJE + FI drives later receptive vocabulary development, caregivers and other adult interventionists could be coached on how to engage their child in HSJE + FI as a means to increase children's receptive vocabulary. This could be particularly impactful, as autistic children often have receptive vocabularies that are smaller than expected given their expressive vocabularies (Woynaroski et al., 2016). Furthermore, if it is the case that early expressive vocabulary is a driver

of later HSJE, this would mean that expressive vocabulary would need to be directly addressed by the intervention in addition to joint engagement. Intervention programming may need to support the use of alternative and augmentative communication (AAC) devices, including speech generating devices, for children who are still very early on in spoken language development (Kasari et al., 2014), and this may increase the extent to which autistic children are able to participate in joint engagement.

This study

This study is designed to determine if there is a superior direction of associations between combinations of HSJE, HSJE + FI, and children's expressive and receptive vocabularies. We examined four research questions, two of which were hypothesis driven based on previous research and two of which were exploratory. First, we examined whether early child expressive vocabulary impacts later HSJE to a greater extent than HSJE impacts later expressive vocabulary. When children talk about what they are playing with, caregivers may perceive this as an opportunity to shape the play as a joint interaction. Furthermore, expressive vocabulary may be employed interactionally to reciprocate caregiver actions around toys in increasingly abstract and flexible ways, enabling the development of more frequent and longer bouts of HSJE (Adamson & Bakeman, 2006; Adamson et al., 2009, 2019). We selected HSJE and not HSJE + FI for this hypothesis because of prior work indicating potentially stronger links between HSJE and expressive language as compared to HSJE + FI and expressive language (Bottema-Beutel et al., 2014).

Second, we examined whether HSJE + FI impacted later receptive vocabulary to a greater extent than early receptive language impacted later HSJE + FI. As described above, previous research has suggested that HSJE + FI may be an "optimal" engagement format for learning what words mean (Bottema-Beutel et al., 2014). Furthermore, children may be able to participate in complex forms of engagement (and therefore benefit from HSJE + FI) without initially high receptive vocabularies, because they may be able to use interactional cues such as caregivers' intonation and gesture even if they do not know what words mean (Bottema et al., 2018). Their participation in these interactions then offers opportunities for building receptive vocabularies.

Finally, we examined two exploratory questions to round out our analyses, but that were not associated with strong directional hypotheses. The first exploratory question is whether there was a superior pathway when comparing associations between early HSJE + FI and later expressive vocabulary, and associations between early expressive vocabulary and later HSJE + FI. The second was whether there is a superior pathway when comparing associations between early HSJE and later receptive vocabulary, and associations between early receptive vocabulary and later HSJE.

Method

Study design

A cross-lagged panel analysis (Allison et al., 2017) with relevant variables measured at two time points, 8 months apart, was used to answer our four research questions. This design can offer evidence about which construct "drives" the other over time. If, for example, the association between early HSJE + FI and later receptive vocabulary is significantly larger than the reverse pathway (early receptive vocabulary and later HSJE + FI), early HSJE + FI can be said to "drive" the development of receptive vocabulary. Important to this approach is that cross-lagged associations are adjusted for important confounds; concurrent associations between the two constructs of interest, and the temporal stability of individual constructs over time.

Participants

This study is a secondary analysis of previously collected longitudinal data for a project aiming to determine predictors of autistic children's spoken vocabulary (redacted for review). Participating families were recruited from two research sites, both located in metropolitan areas in the southern United States. All study procedures were approved by relevant Institutional Review Boards, and caregivers provided written informed consent prior to participating. Participants included 91 children ($n = 16$ girls) from English-speaking families with a confirmed diagnosis of autism or pervasive developmental disorder not otherwise specified (PDD-NOS) according to criteria indicated in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; DSM-IV-TR; American Psychiatric Association, 2000) and their caregivers. Child participants were screened out if they presented with (a) severe motor impairments, (b) metabolic or progressive neurological disorders (including visual and hearing impairments), or (c) identified genetic syndromes. At study entry (4 months prior to Time 1 for the current project), participants were between 24 and 51 months chronological age, had an average mental age of 21.1 months, said 20 words or fewer according to a parent report, and said fewer than five unique words in a 15-min semi-structured language sample.

The racial makeup of families was as follows: 1 identified as American Indian or Alaskan Native, 5 identified as Asian, 16 identified as Black or African American, and 65 identified as white. Four families identified as Hispanic. See Table 1 for additional participant characteristics. Formal educational levels of primary caregivers were distributed as follows: 4 completed less than a high school

Table 1. Participant characteristics at each time point.

	Study entry[a]			Time 1			Time 2		
	M	*SD*	Range	M	*SD*	Range	M	*SD*	Range
ADOS	22.62	3.79	6–28						
MSEL IQ	51.63	3.79	49–122						
MSEL mental age	12.11	4.73	3.75–26.5						
MSEL expressive LAE	7.96	4.15	1–21						
MSEL receptive LAE	6.41	6.17	1–30						
Chronological age				38.7	7.2	24–51	46.7	7.2	32–59
MCDI expressive vocabulary				18.43	29.86	0–181	61.22	85.41	0–396
MCDI receptive vocabulary				115.15	110.98	0–440	166.10	120.68	0–396
Duration HSJE (in seconds)				65.08	58.60	0–291	69.65	52.24	0–189
Number of intervals HSJE + FI				7.52	8.18	0–40	7.71	7.06	0–34

ADOS: autism diagnostic observation schedule; HSJE: higher order supported joint engagement; FI: follow-in talk; IQ: intellectual quotient; LAE: language age equivalent; MCDI: MacArthur–Bates communication development inventories; MSEL: Mullen scales of early learning; M: mean; *SD*: standard deviation.
[a]Study entry is 4 months prior to Time 1.

education, 19 completed a high school diploma or equivalent, 21 completed 1–2 years of college or technical school education, 28 had 3–4 years of college or technical school education, and 15 had attended graduate or professional school. Race, ethnicity, and education data were not available for four families.

Assessments

Confirmation of diagnosis. The Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) Module 1 was administered by a research reliable member of the study team. Scoring algorithms developed by Gotham and colleagues (2007) were used to identify cut points for diagnosing participants with autistic disorder or autism spectrum disorder. One child did not meet the cut-off according to this scoring algorithm, but was retained in the study because the licensed clinician on the study team determined that he met the criteria for PDD-NOS. Altogether, 95% of the sample met the criteria for autistic disorder and 5% met the criteria for PDD-NOS.

Mental age. To characterize the study sample, children's mental ages were examined at study entry using the Mullen Scales of Early Language (MSEL; Mullen, 1995). The MSEL is a standardized assessment of cognitive abilities that combines four subscales such as expressive language, receptive language, fine motor development, and visual discrimination. It is normed for children from birth to 68 months of age.

Child vocabulary. The MacArthur–Bates communicative development inventories (MCDI; Fenson et al., 2003) is a parent-completed checklist that was collected at Time 1 and Time 2. This checklist was used to inventory children's expressive vocabulary (words the child can say and understand) and receptive vocabulary (words the child can understand but may not say).

Parent child free play (PCFP) procedure. Observational variables were coded from the PCFP (Yoder et al., 2015), which was conducted at Time 1 and Time 2. In this procedure, children and caregivers were invited to play as they normally would in a small room containing a standard set of toys (e.g. doll and bottle, blocks, beads and container, a barn and animals). The procedure lasted 15 min and was video recorded from behind a two-way mirror.

Coding and deriving HSJE and HSJE + FI

HSJE and caregiver FI talk were both coded from the PCFP videos using established coding protocols developed and used in previous research (redacted for review) and with the aid of ProCoder DV software (Tapp, 2003). HSJE was coded using timed-event coding, which involves marking the onset and offset of a mutually exclusive and exhaustive set of engagement state codes, so that a total duration for each state can be calculated for each video file (because HSJE is conceptualized as a "state" where duration is important). FI comments (i.e. describing what the child is doing with the toys or the toys they are attending to) and directives (offering a suggestion for how the child might use toys the child is attending to in new ways) were both coded as FI talk using interval coding with a 5-s window. We coded FI talk using interval coding because we were interested in estimating the prevalence of such events and not the exact duration of their occurrence. We note that the overwhelming majority of FI talk was in response to children's focus of attention. While it is possible that FI talk could be coded in response to children's communication if it overlapped with children's focus of attention, this occurred infrequently due to the impaired social communication and language abilities of the

Table 2. Operational definitions of observational codes.

Code	Definition	Examples
Higher order supported joint engagement (HSJE)	The caregiver and child are engaged with the same materials, and the caregiver's actions influence the child's play. The child does not visually reference the adult. The child reciprocates the adult's play actions or collaborates with the adult in a play scheme	Turn-taking sequences (e.g. taking turns placing puzzle pieces on a board), child imitating caregiver (e.g. caregiver makes a toy pig drink water, child makes a toy cow drink water), or the child following through on a verbal directive made by the caregiver
Supported joint engagement (SJE)	The caregiver and child are engaged with the same materials, and the caregiver's actions influence the child's play. The child does not visually reference the adult. Includes instances when the child does and does not reciprocate the adult's play actions/ collaborate with the adult in a play scheme	The HSJE examples above, plus the following: the child and caregiver both play with stacking blocks. The child must adapt their play to accommodate the caregiver's selection and positioning of blocks, but there is no established turn taking
Coordinated joint engagement (CJE)	The caregiver and child are engaged with the same materials, and the caregiver's actions influence the child's play. The child visually references the adult during the interaction	The child and caregiver take turns stacking blocks, and the child looks to the caregiver after stacking each block and smiles
Follow-in talk (FI)	The caregiver provides linguistic input that is related to the child's current focus of attention, either by describing what the child is looking at or playing with, or making a suggestion about how the child could play with the toy	The child plays with a toy boat, and the caregiver says, "Look at the boat rocking in the water!" Or, the child plays with a baby doll, and the caregiver says, "Can the baby drink the bottle?"

children in our sample. See Table 2 for definitions and examples of FI talk and HSJE.

Two independent coders overlapped on 20% of videos to determine intercoder agreement. Individual intra-class correlation coefficients from two-way random-effects models were 0.95 for HSJE and 0.79 for FI, indicating acceptable agreement (Yoder et al., 2019). To create the HSJE + FI variable, HSJE timed-event codes and FI interval codes were combined using ProCoder Merger (Tapp, 2013), which allowed for computing the total number of intervals within which HSJE and FI co-occurred for each participant. Conceptually, HSJE + FI is a subcategory of HSJE (and therefore the instances in which these two variables were coded overlap). However, they are somewhat distinct, given the different metrics and procedures for their derivation.

All participant-level data have been deposited in the Open Science Framework, including the Stata data file and syntax file.

Statistical procedures

Preliminary analysis. Variables that were not normally distributed were transformed as follows: HSJE, HSJE + FI, and expressive vocabulary at Times 1 and 2 were log transformed. Receptive vocabulary at Time 2 was natural log transformed.

Main analysis. Structural equation modeling, using the structural equation modeling (SEM) macro in Stata 15 (StataCorp, 2017), was used to generate standardized coefficients and *p*-values for cross-lagged pairs relevant to each research question, controlling for concurrent correlations between variables at each time point and longitudinal correlations within each time point. Linear contrasts of cross-lagged coefficients were then computed using the lincom command in Stata (StataCorp, 2013) to determine whether there was a superior pathway for each variable pair (e.g. if the cross-lagged association between early receptive vocabulary and later HSJE was superior to the cross-lagged association between early HSJE and later receptive vocabulary). Full information maximum likelihood estimation was used to account for missing data (which was <20% for each variable), and robust standard errors were computed to account for violations in distributional assumptions.

Results

Standardized coefficients for cross-lagged associations, concurrent correlations, and longitudinal correlations are shown in Figures 1 to 4, with each figure number corresponding to respective research questions. All cross-lagged associations between early engagement variables and later vocabulary variables were significant. Standardized coefficients for the paths from Time 1 HSJE to Time 2 expressive and receptive vocabulary were 0.18 ($p = 0.02$) and 0.26 ($p < 0.0001$), respectively. Standardized coefficients for the paths from Time 1 HSJE + FI and Time 2 expressive and receptive vocabulary were 0.14 ($p = 0.05$) and 0.15 ($p = 0.03$), respectively. In contrast, none of the cross-lagged associations between early vocabulary variables and later engagement variables were significant. Table 3 provides coefficients, standard errors, and *p*-values for linear contrasts between cross-lagged coefficients generated from each SEM. None of the four linear contrasts between cross-lagged coefficients were significant, indicating that

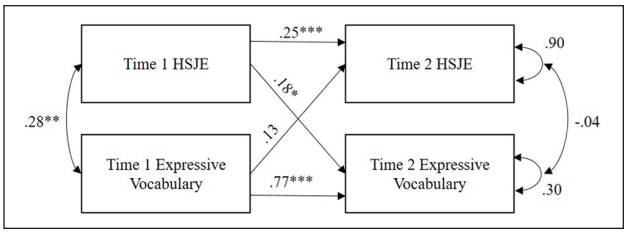

Figure 1. Cross-lagged analysis for higher order supported joint engagement (HSJE) and expressive vocabulary.
Standardized coefficients generated from SEM analysis using robust standard errors.
*$p < 0.05$; $p < 0.01$; ***$p < 0.001$.

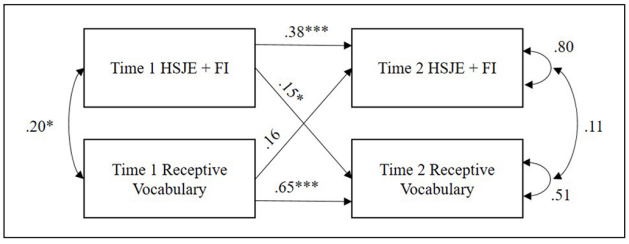

Figure 2. Cross-lagged analysis for higher order supported joint engagement with follow-in talk (HSJE + FI) and receptive vocabulary.
Standardized coefficients generated from SEM analysis using robust standard errors.
*$p < 0.05$; $p < 0.01$; ***$p < 0.001$.

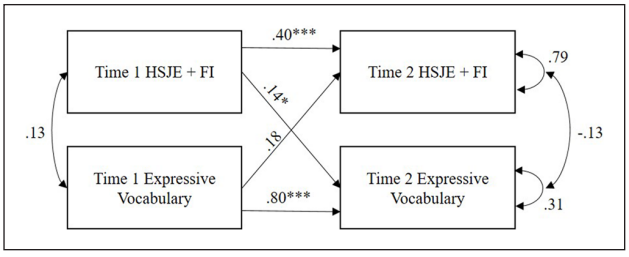

Figure 3. Cross-lagged analysis for higher order supported joint engagement with follow-in talk (HSJE + FI) and expressive vocabulary.
Standardized coefficients generated from SEM analysis using robust standard errors.
*$p < 0.05$; ***$p < 0.001$.

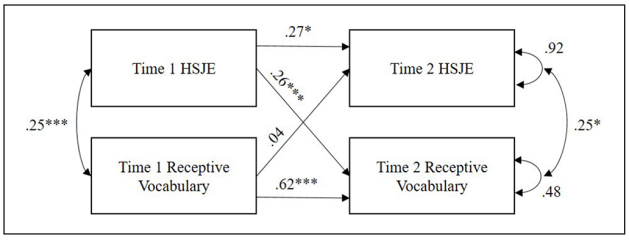

Figure 4. Cross-lagged analysis for higher order supported joint engagement (HSJE) and receptive vocabulary.
Standardized coefficients generated from SEM analysis using robust standard errors.
*$p < 0.05$; $p < 0.01$; ***$p < 0.001$.

Table 3. Linear contrasts between cross-lagged effects generated from SEM.

Linear contrast	Coefficient	*SE*	95% CI
(Early exp → late HSJE)– (early HSJE → late exp)	−0.05	0.13	[−0.31, 0.22]
(Early rec → late HSJE + FI)– (early HSJE + FI → late rec)	0.01	0.14	[−0.27, 0.29]
(Early rec → late HSJE)– (early HSJE → late rec)	−0.22	0.12	[−0.46, 0.03]
(Early exp → late HSJE + FI)– (early HSJE + FI → late exp)	0.03	0.12	[−0.21, 0.28]

CI: confidence interval; Exp: expressive vocabulary; FI: follow-in talk; HSJE: higher order supported joint engagement; Rec: receptive vocabulary;
SE: standard error; SEM: structural equation modeling.

within each SEM, there were no superior directional pathways between variables.

Discussion

Our findings expand on previous work regarding the developmental pathways between children's joint engagement with caregivers and their developing vocabularies. We focus on a form of joint engagement where the caregiver influences the child's toy play, and the child reciprocally engages with the caregiver but does not manage the interaction with gaze. We also focus on a specific form of caregiver talk that occurs when the caregiver talks about the child's focus of attention.

Neither of our two hypotheses was entirely supported. For our first research question, the opposite pathway relative to our prediction was significant; early HSJE was significantly associated with children's later expressive vocabulary, but associations in the reverse direction were not significant. HSJE may support expressive vocabulary by providing an interactional substrate within which spoken language is relevant. Our findings differ from Adamson and colleagues (2019), who showed a positive association between children's emerging expressive language on children's later joint engagement. We showed concurrent associations between HSJE and expressive vocabulary, but not between HSJE + FI and expressive vocabulary, and no longitudinal correlations between early measures of either construct and later expressive vocabulary. There are a variety of differences in study designs that could account for these divergent findings. First, we treated expressive vocabulary as a continuous variable as opposed to dichotomizing children into talking and not talking. Second, our sample only included children with confirmed autism diagnoses, while Adamson and colleagues' study included all children who were considered at heightened probability of later diagnosis. Third, we used different definitions of joint engagement (our variable was restricted to HSJE) and symbol infusion (our variable was restricted to follow-in talk). Finally, we

controlled for Time 1 correlations between HSJE and expressive vocabulary and longitudinal correlations within each variable.

Our second prediction was partially supported; we replicate previous findings that HSJE + FI is positively associated with autistic children's later receptive vocabularies and expand on these findings by showing that these correlations hold when accounting for concurrent associations between variables and longitudinal associations within each variable. We also showed that the reverse association (between early receptive vocabulary later HSJE + FI) was not significant. However, there was no significant difference between coefficients for each pathway, preventing us from making a strong conclusion that HSJE + FI drives the development of receptive vocabulary.

Findings for our exploratory research questions followed similar patterns as our first two research questions; the pathways between early joint engagement variables and later vocabulary variables were significant, while pathways between early vocabulary variables and later joint engagement variables were not. Again, we did not find significant differences between cross-lagged coefficients, so we cannot make strong conclusions that any particular pathway is superior to the reverse pathway.

Limitations

The results of this study should be interpreted in light of some limitations. First, like all correlational approaches, our analysis does not permit ruling out all alternative explanations for significant findings. However, consistent with our cross-lagged analytic approach, we do rule out several important confounds. Future work with larger sample sizes may permit researchers to control for additional potential confounds such as family socio-economic status. Second, this study is restricted to understanding the development of children's vocabulary. Joint engagement may also influence other aspects of language development, including syntactic, pragmatic, and turn-taking competence, which were not examined in this study. Third, research conducted by our team subsequent to the collection of this data set has indicated that two PCFP sessions per time point may need to be coded and averaged across to derive stable estimates of HSJE (redacted for review). Because this study coded HSJE from only one session, it is possible that measurement error was introduced into our estimates, which increases the probability of type II error (i.e. failing to find a "true" effect when such an effect exists). Subsequent research designed with this in mind may detect significant comparisons between cross-lagged effects.

Conclusion

Using a cross-lagged panel analysis, we show that, for young autistic children who are in the early phases of language learning, early HSJE and HSJE + FI are positively associated with later expressive and receptive vocabulary. Our findings suggest that supporting HSJE, and HSJE that co-occurs with caregivers' FI talk, may be an appropriate initial intervention target for children with delayed language. While our findings may not generalize to autistic children who do not experience language delays, we believe our study is important, as it offers implications for intervention research targeted toward children who are most likely to be enrolled in and demonstrate need for early intervention services.

Given that our analyses control for relevant concurrent and longitudinal associations between and among variables of interest, our cross-lagged panel design is stronger than traditional longitudinal designs. However, because we must temper our claims in regards to the direction of effects (because coefficients for a given pathway were not significantly different from the coefficients for the corresponding pathway in the reverse direction), these intervention targets could be supported alongside more direct methods for increasing receptive and expressive vocabulary (such as supporting AAC use). This may be especially useful in instances when children are making little progress or are progressing slowly within joint engagement interventions (e.g. Rogers & Dawson, 2020; elaborated in Stahmer et al., 2011).

Future research with larger sample sizes could test additional pathways in the context of cross-lagged panel analyses, which could, for example, determine the relative effect of HSJE and HSJE + FI in single model. In addition, randomized controlled trials, where HSJE and HSJE + FI are tested as intervention mediators, would allow for stronger conclusions in regards to the potential causal associations between these two joint engagement variables and autistic children's developing vocabularies. While Gulsrud and colleagues (2016) have attempted to assess intervention practices as mediators that contribute to increasing joint engagement, studies are still needed that test joint engagement as a mediator of child outcomes.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by National Institute for Deafness and other Communication Disorders under Grant Nos. R03DC017013-1 (PI Kristen Bottema-Beutel) and R01DC006893 (PI Paul J Yoder).

ORCID iDs

Kristen Bottema-Beutel https://orcid.org/0000-0002-0009-9464
So Yoon Kim https://orcid.org/0000-0003-1349-8031

Paul J Yoder https://orcid.org/0000-0002-0337-0141

Note

1. We use "identity-first language" (e.g. "autistic person") because it is generally preferred by autistic people (Kenny et al., 2016) and is considered less stigmatizing than person-first language (Brown, 2011). See Bottema-Beutel et al. (2020) for an expanded rationale for this language choice.

References

Adamson, L., & Bakeman, R. (2006). Development of displaced speech in early mother-child conversations. *Child Development*, *77*(1), 186–200.

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.

Adamson, L. B., Bakeman, R., Suma, K., & Robins, L. (2019). An expanded view of joint attention: Skill, engagement, and language in typical development and autism. *Child Development*, *90*(1), e1–e18.

Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Socius*, *3*, 2378023117710578.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders-IV-TR*.

Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*, 1278–1289.

Bloom, L., Tinker, E., & Scholnick, E. K. (2001). The intentionality model and language acquisition: Engagement, effort, and the essential tension in development. *Monographs of the Society for Research in Child Development*, *66*(4), 1–89.

Bottema-Beutel, K., Kim, S. Y., & Crowley, S. (2019). A systematic review and meta-regression analysis of social functioning correlates in autism and typical development. *Autism Research*, *12*(2), 152–175.

Bottema-Beutel, K., Lloyd, B., Watson, L., & Yoder, P. (2018). Bidirectional influences of caregiver utterances and supported joint engagement in children with and without autism spectrum disorder. *Autism Research*, *11*(5), 755–765.

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.

Bottema-Beutel, K., Kapp, S. K., Lester, J. N., Sasson, N. J., & Hand, B. N. (2020). Avoiding ableist language: Suggestions for autism researchers. *Autism in Adulthood*. Advance online publication. https://doi.org/10.1089/aut.2020.0014

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

Bottema-Beutel, K., Yoder, P., Hochman, J. M., & Watson, L. (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*, 2162–2174.

Brown, L. (2011). The significance of semantics: Person first language: Why it matters. https://www.autistichoya.com/2011/08/significance-of-semantics-person-first.html

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.

Edmunds, S. R., Kover, S. T., & Stone, W. L. (2019). The relation between parent verbal responsiveness and child communication in young children with or at risk for autism spectrum disorder: A systematic review and meta-analysis. *Autism Research*, *12*(5), 715–731.

Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J. P., . . . Reilly, J. (2003). *MacArthur communicative development inventories: User's guide and technical manual*. Baltimore, MD: Paul H. Brookes.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*, 613–627.

Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, . . . Pickles, A. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): A randomised controlled trial. *The Lancet*, *375*(9732), 2152–2160.

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.

Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *Journal of Child Psychology and Psychiatry*, *47*, 611–620.

Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., . . . & Almirall, D. (2014). Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(6), 635–646.

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.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.

McDuffie, A., & Yoder, P. (2010). Types of parent verbal responsiveness that predict language in young children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, *53*, 1026–1039.

Mullen, E. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service.

Mundy, P. (2016). *Autism and joint attention: Development, neuroscience, and clinical fundamentals*. New York, NY: The Guilford Press.

Mundy, P., & Jarrold, W. (2010). Infant joint attention, neural networks and social cognition. *Neural Networks*, *23*(8–9), 985–997.

Pickles, A., Harris, V., Green, J., Aldred, C., McConachie, H., Slonims, V., . . . & PACT Consortium. (2015). Treatment mechanism in the MRC preschool autism communication trial: Implications for study design and parent-focussed therapy for children. *Journal of Child Psychology and Psychiatry*, *56*(2), 162–170.

Rahman, A., Divan, G., Hamadani, S. U., Vajaratkar, V., Tayor, C., Leadbitter, L., . . . Green, J. (2016). Effectiveness of the parent-mediated intervention for children with autism spectrum disorder in south Asia in India and Pakistan (PASS): A randomised controlled trial. *Lancet Psychiatry*, *3*(2), 128–136.

Ramírez, N. F., Lytle, S. R., & Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proceedings of the National Academy of Sciences*, *117*(7), 3484–3491.

Robins, D. L., Fein, D., & Barton, M. (2009). *The modified-checklist for autism in toddlers-revised with follow-up (M-CHAT-R/F)*.

Rogers, S. J., & Dawson, G. (2020). *Early Start Denver Model for young children with autism: Promoting language, learning, and engagement*. New York, NY: Guilford Publications.

Sandbank, M., Bottema-Beutel, K., Crowley, S., Cassidy, M., Dunham, K., Feldman, J. I., . . . Woynaroski, T. G. (2020). Project AIM: Autism intervention meta-analysis for studies of young children. P. *Sychological Bulletin*, *146*(1), 1–29.

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., . . . Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(8), 2411–2428.

Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children's communication. *Journal of Autism and Developmental Disorders*, *32*(2), 77–89.

Stahmer, A. C., Schreibman, L., & Cunningham, A. B. (2011). Toward a technology of treatment individualization for young children with autism spectrum disorders. *Brain Research*, *1380*, 229–239.

StataCorp. (2013). *Stata 13 base reference manual*. College Station, TX: Stata Press.

StataCorp. (2017). *Stata statistical software: Release 15*.

Tapp, J. T. (2003). ProcoderDV [Computer software and manual]. http://procoder.vueinnovations.com/sites/default/files/public/d10/ProcoderDVManual_0.pdf

Tapp, J. T. (2013). *Procoder merger* [computer software]. Vanderbilt Kennedy Center.

Woynaroski, T., Yoder, P., & Watson, L. R. (2016). Atypical cross-modal profiles and longitudinal associations between vocabulary scores in initially minimally verbal children with ASD. *Autism Research*, *9*(2), 301–310.

Yoder, P. J., Lloyd, B. L., & Symonds, F. J. (2019). *Observational measurement of behavior* (2nd ed.). Baltimore, MD: Paul H. Brookes Publishing.

Yoder, P. J., Watson, L., & 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.