Infant Behavior and Development 71 (2023) 101840

Contents lists available at ScienceDirect

Infant Behavior and Development

journal homepage: www.elsevier.com/locate/inbede

Infant distractibility from social events mediates the relation between maternal responsiveness and infant language outcomes

Kaitlyn Testa [a], Myriah E. McNew [a], James Torrence Todd [a], Bret Eschman [b], Lorraine E. Bahrick [a,*,1]

[a] Florida International University, U.S.A
[b] University of Tennessee, Chattanooga, U.S.A

ARTICLE INFO

Keywords:
Maternal responsiveness
Intersensory matching
Receptive language
Bids for attention
Redirecting attention
Infant distractibility

ABSTRACT

Research demonstrates that contingent and appropriate maternal responsiveness to infant requests and bids for attention leads to better language outcomes. Research also indicates that infants who are less distracted by irrelevant competing stimulation and attend efficiently to audiovisual social events (e.g., faces and voices) show better language outcomes. However, few studies have assessed relations between maternal responsiveness, infant attention to faces and voices, and distractibility, and how together these factors lead to early language outcomes. A newly developed audiovisual protocol, the Multisensory Attention Assessment Protocol (MAAP; Bahrick et al., 2018), allows researchers to examine individual differences in attention to faces and voices and distractibility, and to assess relations with other variables. At 12 months, infants ($n = 79$) in an ongoing longitudinal study participated in the MAAP to assess intersensory matching of synchronous faces and voices and attention to an irrelevant competing visual distractor event. They also were observed in a brief play interaction to assess infant bids for attention and maternal responsiveness (accept, redirect, or ignore). At 18 months, receptive and expressive language were assessed using the Mullen Scales of Early Learning. Several noteworthy findings emerged: 1) mothers were generally responsive, accepting 74% and redirecting 14% of infant bids, 2) infants who had a greater number of their bids redirected by mothers, and who had better intersensory matching of synchronous faces and voices, showed less attention to the distractor, and 3) infants who showed less attention to the distractor had better receptive language. Findings demonstrate that maternal redirecting of infant attention by mothers who are generally responsive may promote better infant attentional control (lower distractibility) which in turn predicts better receptive language in toddlers.

1. Introduction

Maternal responsiveness is known to have a significant impact on a variety of domains including social-emotional, cognitive, and language development (e.g., Leerkes et al., 2009; Smith et al., 2006; Tamis-LeMonda et al., 2001; for a review, see Deans, 2020). Maternal sensitivity is a type of maternal responsiveness defined as mothers' prompt, contingent, and appropriate responses to infants'

* Correspondence to: Florida International University, 11200 South West 8th Street, Miami, FL 33199, United States.
E-mail address: bahrick@fiu.edu (L.E. Bahrick).
[1] This work was supported by National Institutes of Health (NIH) Grants R01-HD053776 and R01-HD094803 awarded to the last author

https://doi.org/10.1016/j.infbeh.2023.101840
Received 1 July 2021; Received in revised form 14 March 2023; Accepted 20 April 2023
Available online 19 May 2023

behaviors (e.g., Tamis-LeMonda et al., 2001; Baumwell et al., 1997). Maternal sensitivity has been shown to predict developmental outcomes across various domains and is a particularly robust predictor of infant language development. Specifically, maternal responsiveness that is contingent on infant behavior has been shown to predict infant production of mature vocalizations (Goldstein et al., 2003; Warlaumont et al., 2014) and infant language milestones (e. g., Tamis-LeMonda et al., 2001).

1.1. Maternal responsiveness to infant bids for attention and language outcomes

Maternal responsiveness has been assessed using both experimental manipulations and naturalistic or unstructured observations (e. g., free play). For instance, Goldstein and colleagues experimentally manipulated maternal responsiveness by instructing one group of mothers to smile and move closer to their 8-month-old infants contingent on (in response to) their infant's vocalization, whereas another group of mothers smiled and moved closer in a manner that was noncontingent with their infant's vocalization (Goldstein et al., 2003; Goldstein & Schwade, 2008). Infants produced more frequent and phonologically advanced vocalizations when their mothers smiled and moved closer contingent upon their vocalizations. In addition, Warlaumont and colleagues found that maternal responsiveness to infant vocalizations predicted infant production of speech-like vocalizations in typically and atypically developing infants between 8 and 48 months of age (Warlaumont et al., 2014). During unstructured play settings, Tamis-LeMonda and colleagues measured maternal responsiveness as any positive and meaningful change in the mother's behavior (e.g., naming or commenting about an object the infant was holding) in response to changes in the infant's behavior (e.g., exploring or manipulating a new object or toy; Tamis-LeMonda et al., 2001). Greater maternal responsiveness predicted earlier achievement of infant language milestones (e.g., first imitations, first words, 50 words in expressive language, use of combinatorial speech, the use of language to talk about the past).

Bids for attention are behaviors that are directed toward a social partner, presumably with the intent to share attention, including gesture, gaze and/or vocalization. Infants' ability to share attention with a social partner develops rapidly within their first year. By 6 months of age, infants are capable of making simple social bids (e.g., gaze, vocalizations) and begin to direct their behaviors to social partners. By 12 months, infants can flexibly shift gaze between social partners and objects, as well as point, vocalize, and gesture towards objects of interest (e.g., Tamis-LeMonda et al., 2001; Mundy et al., 2007). Given their communicative intent, infant bids may elicit greater interaction with caregivers, and in turn scaffold language development, providing an example of the active role infants play in their own language learning. We thus expected that infant bids and maternal responsiveness to infant bids may be an especially appropriate context for assessing language outcomes. However, few studies have characterized infant bids for attention and maternal responsiveness to those bids (see Tamis-LeMonda et al., 2001, for an exception).

Infant bids for attention have been defined in various ways: 1) verbal and nonverbal attempts to elicit mothers' attention or engagement, 2) smiling while gazing at the parent, or 3) using eye contact and/or gestures to spontaneously initiate coordinated joint attention (e.g., Wilson & Durbin, 2013; Ekas et al., 2013; Van Hecke et al., 2007). Infant bids have often been studied experimentally in the context of joint attention with a trained experimenter, requiring the experimenter to elicit behaviors from infants (e.g., Mundy et al., 2007; Van Hecke et al., 2007), or during the still-face paradigm, in which mothers are required to remain passive in response to infant bids for attention during a portion of the task (still-face phase; e.g., Wilson & Durbin, 2013; Bigelow & Power, 2016).

Within the context of joint attention studies with an experimenter, infant attempts to initiate triadic joint attention (defined as shifting gaze from an object to a partner and back to the object, or vice versa) has predicted vocabulary size in later development (Mundy et al., 2007). Earlier forms of infant bids for attention (dyadic joint attention; e.g., pointing to an object and vocalizing without shifting gaze between object and partner), have also been shown to predict infant vocabulary size. For example, initiating behavioral requests (IBR; using gaze and gestures to elicit aid from social partners) with an experimenter has been shown to predict expressive and receptive language in typically developing children and children with Down syndrome (Mundy et al., 1995); although several studies have found no direct relation between IBR and language outcomes (e.g., Mundy et al., 2007, Mundy et al., 1990).

In contrast with experimental studies, only a few studies have assessed infant bids for their mother's attention in unstructured dyadic interactions or how maternal responses to infant bids relate to infant language development. For example, Tamis-LeMonda et al. (2001) assessed maternal responsiveness (e.g., affirmations of child actions; descriptions or questions about an object, event, or activity; prompts or demonstrations during play) to various infant behaviors (vocalizations, bids, exploration, and play). Whereas maternal responsiveness to infant vocalizations, exploration, and play predicted children's achievement of language milestones, maternal responsiveness to infant bids was not a predictor of language outcomes. However, in this study and other studies (Baumwell et al., 1997; Bornstein et al., 2020), infant bids were characterized only as looks to the mother (and not necessarily directed, or bids for attention). A single behavior alone (e.g., gaze or smile to the mother) may not necessarily reflect an attempt to engage the mother. For example, looking to the mother may reflect social referencing or visual exploration, whereas a gaze in combination with vocalizing or pointing to an object, may be more likely to reflect a bid for attention. For this reason, we used a more stringent definition of infant bids requiring at least two behaviors in combination (e.g., vocalizations directed to the mother, points to objects, looks to the mother, and vocalizations directed towards the mother). Further, combinations of infant behaviors may be more likely to elicit maternal responses that promote subsequent language development. For example, pointing to an object while vocalizing may elicit object labeling from mothers, creating opportunities for object-label mapping.

Mothers' use of strategies such as following in (maintaining attention) and redirecting attention have also been related to language outcomes. There is general agreement that following in to the focus of infant attention during joint attention episodes is associated with better language outcomes (e.g., Tomasello & Farrar, 1986; Akhtar et al., 1991; Laakso et al., 1999). In contrast, there is some disagreement in the literature about the effects of redirecting attention on later language outcomes. Some studies indicate that redirecting attention is associated with poorer language outcomes (e.g., Tomasello & Farrar, 1986; Akhtar et al., 1991). However, others suggest that redirecting infant attention may have some beneficial effects on language (see, Saxon, 1997; Legerstee et al., 2002;

Mendive et al., 2013; Shimpi & Huttenlocher, 2007). For example, Saxon (1997) found that maternal redirection of infant bids for joint attention during a play interaction was positively associated with increases in bouts of joint attention between infants and mothers, and joint attention in turn was positively associated with better language outcomes. Shimpi and Huttenlocher (2007) found that successful maternal redirections (in which the infant followed the mother's direction) were positively related to infant vocabulary size. This may suggest that maternal redirecting of infant attention following an infant's bid may help refocus infant attention which in turn may have a positive effect on language.

Thus, mothers' use of maintaining and redirecting attention strategies may also have an effect on the development of infant attention skills. One study found that maternal redirecting of infant attention preceded infant focused attention on objects (Legerstee et al., 2002). Mothers use of attention redirecting strategies are also followed by dyadic joint attention (e.g., mother directs infant attention, infant looks to object, but infant does not look to mother), whereas attention maintaining strategies are often followed by coordinated (triadic) joint attention (infant shifts gaze between mother and object) between infants and their caregivers (Mendive et al., 2013). Moreover, it is unclear whether or how these maternal attention strategies cascade to later language outcomes, or what role infant attention plays in this cascade.

In sum, whereas experimental studies have demonstrated links between maternal responsiveness to infant bids and language outcomes (e.g., Mundy et al., 2007; Bigelow & Power, 2016), there has been little evidence from studies of unstructured mother-infant interactions. However, definitions of infant bids and maternal responsiveness to those bids have varied widely across studies. In the present study, we characterized infant bids as instances in which infants directed two or more behaviors (e.g., smiling, gazing, vocalizing, gesturing) to the mother, and we characterized maternal responsiveness as whether mothers accepted, redirected, or ignored those bids.

1.2. Sustained attention and distractibility in early development

Research indicates that infants who show greater sustained attention have better language outcomes (e.g., Bahrick et al., 2018; Yu et al., 2018). Sustained attention entails continuous and active attentional engagement with a stimulus or task and typically involves inhibiting attention to a distracting object or event (for reviews, see Colombo, 2001; Colombo & Cheatham, 2006). Thus, greater sustained attention reflects better attention control and focus, as well as lower levels of looking to a distractor event (i.e., lower distractibility). Distractibility has been assessed in a variety of ways, most often involving the "capacity to maintain attention in the context of competing events" (Salley et al., 2013; p. 464). Research supporting the link between sustained attention and reduced distractibility comes from a variety of domains including studies that have assessed looking time between two screens (where one screen served as the target event and the other as the distractor; e.g., Casey & Richards, 1988; Salley et al., 2013), those assessing infant visual attention during object manipulation in the presence of a video being played on a TV (e.g., Ruff & Capozzoli, 2003; Lansink & Richards, 1997), and those assessing heart rate defined measures of attention to target events and distractors (e.g., Richards & Casey, 1991). Here, we define infant "distractibility" in terms of looking to a competing visual distractor event during the presentation of two audiovisual events. Specifically, distractibility is operationalized as looking to a silent competing nonsocial visual distractor event during the presentation of two audiovisual social events (women speaking). Overall, this research indicates that infants are less distractible during phases of heightened sustained attention, reflecting greater endogenous control of attention. Further, research has shown that lower distractibility (i.e., better attention control) is associated with better receptive and expressive language outcomes in infants and children (Dixon & Salley, 2007; Salley et al., 2013).

1.3. Intersensory matching of faces and voices

Selective attention to social events, particularly audiovisual speech events, also predicts language outcomes (for a review, see Bahrick et al., 2020). Infant selective attention to audiovisual speech can be indexed by assessing intersensory matching of visual and auditory stimulation from faces and voices (e.g., Bahrick et al., 2018; Bahrick, 2010). In traditional intersensory matching tasks (e.g., intermodal matching paradigm; e.g., Spelke, 1976; Bahrick, 1983), infants are presented with visual events on two screens, one of which is synchronous with its natural soundtrack and the other is asynchronous. Typically, intersensory matching is inferred for a group of infants by comparing the proportion of total looking to the sound-synchronous event with the chance value of looking to either screen (50%). Thus, in the traditional intermodal matching paradigm, the sound-asynchronous event serves as a distractor that competes for visual attention.

However, traditional paradigms for assessing intersensory matching are inappropriate for assessing fine-grained individual differences because they include a small number of trials (1 or 2) and their psychometric properties (e.g., reliability) have not been established. Further, they typically include a single condition (e.g., social or nonsocial events) and provide just one measure (a preference for the sound-synchronous display). Recently, a new protocol, the Multisensory Attention Assessment Protocol (MAAP; Bahrick et al., 2018), has been developed to assess fine-grained individual differences in early attention skills. It assesses three foundational attention skills (sustaining attention, disengaging/shifting attention, intersensory matching) to both social and nonsocial events. It shows very good reliability in both infants and young children (Bahrick et al., 2018). Prior research with 2- to 5-year-old children using the MAAP demonstrates that sustained attention to social events (women speaking) predicts intersensory matching of synchronous faces and voices, which in turn predicts receptive and expressive language outcomes (Bahrick et al., 2018). Also, the MAAP assesses attention to social and nonsocial events in both the presence of a silent dynamic central visual distractor event (morphing geometric shapes) and in its absence, in order to characterize the impact of a distracting event on attention. Children's attention to the central visual distractor event during the presentation of the audiovisual social or nonsocial peripheral events indexes distractibility.

1.4. The present study

The purpose of this study is to characterize relations between maternal responsiveness to infant bids, infant attention to audiovisual social events, and infant distractibility, as well as pathways from these predictors to later language outcomes. At 12 months, we assessed maternal responsiveness to infant bids for attention during a mother-child play interaction. We also assessed infant attention to audiovisual social events (synchronous faces and voices) and distractibility from these social events (i.e., looking to a silent visual distractor event) using the MAAP at 12 months. Finally, at 18 months, we assessed receptive and expressive language using the Mullen Scales of Early Learning (MSEL; Mullen, 1995). Consistent with previous findings, we predicted that greater maternal responsiveness to infant bids, greater attention to social events (particularly intersensory matching of faces and voices; Bahrick et al., 2018; Edgar et al., 2022), as well as lower distractibility would all predict better receptive and expressive language outcomes (e.g., higher scores on the MSEL Receptive and Expressive Language scales). Further, we predicted that lower infant distractibility from social events would be associated with greater intersensory matching of faces and voices, as lower attention to an irrelevant distractor frees up resources for perceiving synchronous faces and voices. We also included assessments of infant attention to the nonsocial events in the MAAP. This served as a control condition for comparison with attention to the social events to determine whether relations with infant attention were specific to social contexts or general across the social and nonsocial domains. In addition, for the first time, we tested structural equation models (SEM) to characterize developmental pathways between maternal responsiveness and infant attention to social events, distractibility, and later language outcomes.

2. Methods

2.1. Participants

Seventy-nine infants (44 females, 35 males) who were enrolled in a larger, ongoing longitudinal study ($N = 104$) participated.[2] In the larger study, infants were recruited at 3 months of age and were tested at 3, 6, 12, 18, 24, and 36 months of age. Out of the $n = 79$, 47 infants and their caregivers participated and provided useable data at both 12 and 18 months, whereas 25 participated only at 12 months and 7 participated only at 18 months. The longitudinal study, entitled "Development of Intermodal Perception of Social and Nonsocial Events", received IRB approval from the Social and Behavioral Review Board of Florida International Univeristy (IRB-13–0448-CR06). See Table 1 for demographic information, including descriptive statistics for sex, ethnicity, race, and age. Infants were recruited from birth records obtained from the Florida Department of Health and caregivers were contacted via publicly listed phone numbers. Inclusion criteria included gestational age between 38 and 42 weeks at birth, Apgar scores of 9 or 10, and birthweight greater than 5 lbs. Further, only infants whose mothers (rather than fathers, grandparents, etc.) participated with them in the lab-based play interaction (see Section 2.2.3), were included in the study. There were no inclusion or exclusion criteria for mothers who participated in the study. Families received a $30 cash gift for their participation at each visit.

Infants participated in the MAAP (Bahrick et al., 2018) and an 8-minute mother-child interaction at 12 months age and were administered the Mullen Scales of Early Learning (Mullen, 1995) at 18 months of age. To deal with missing data in our correlation and structural modeling analyses, we use Full Information Maximum Likelihood estimation (FIML), which produces unbiased parameter estimates for data that is missing at random. FIML is preferable to traditional approaches for dealing with missing data (e.g., listwise and pairwise deletion). According to Enders (2010) and Graham & Shafer (1999) the validity of results is not compromised with up to 50% missing data. Based on the available statistical literature, what matters more than the amount of missing data across variables, is the amount of missing data within a variable (more than 50% available data). Thus, all infants who had data for one of more variables of interest were included in our study. All variables included in analyses had upwards of 51.9% available data. No variables were missing data greater than 50% (for details, see Table 3).

2.2. Assessments

2.2.1. The Multisensory Attention Assessment Protocol (MAAP)

At 12 months, infants participated in the MAAP (Bahrick et al., 2018). The MAAP is a 24-trial, 3-screen video-based protocol designed to index individual differences in three foundational multisensory attention skills (speed of shifting/disengaging, attention maintenance, and intersensory matching) to audiovisual social and nonsocial events. Infants were seated on the mother's lap about 101.6 cm away from a 116.84 cm widescreen monitor. Mothers were instructed to refrain from interacting with their infant during the procedure and were unaware of which side of the screen the sound-synchronous event was presented as they wore blackout glasses throughout the procedure.

On each 15 s trial, infants are presented with a 3 s silent dynamic central event (morphing geometric shapes), followed by two 12 s lateral events (see Fig. 1). Blocks of 12 social and 12 nonsocial trials are presented. For half of the trials in each block, the silent central

[2] With a sample size of $n = 79$, for bivariate correlations there is sufficient power to detect a medium effect size of $r = 0.31$ or greater (assuming a β of .80 and a two-tailed p value of .05). Further, there is sufficient power for structural equation models (SEM) to detect a non-zero path coefficient that accounts for 7.6% unique variance (assuming a β of .80, a two-tailed p value of .05, five predictors, and an R^2 of .30).

Table 1
Demographic information for the sample ($n = 79$).

	Age					
	Overall ($n = 79$)		12 months ($n = 72$)		18 months ($n = 54$)	
	n	Percent	*n*	Percent	*n*	Percent
Sex						
Female	44	55.7%	41	56.9%	27	50.0%
Male	35	44.3%	31	43.1%	27	50.0%
Ethnicity						
Hispanic	50	63.3%	46	63.9%	31	57.4%
Non-Hispanic	27	34.2%	24	33.3%	21	38.9%
Not Specified	2	2.5%	2	2.8%	2	3.7%
Race						
White / Caucasian	53	67.1%	50	69.4%	34	63.0%
Black/ African American	12	15.2%	11	15.3%	10	18.5%
Asian/ Pacific Islander	3	3.8%	2	2.8%	2	3.7%
Native American	1	1.3%	1	1.4%	1	1.9%
More than 1 race	7	8.9%	6	8.3%	4	7.4%
Not Specified	3	3.8%	2	2.8%	3	5.6%
			M	*SD*	*M*	*SD*
Infant Age			12.05	0.25	18.07	0.43
Maternal Age			33.61	5.57	33.62	5.87

Note: N is the number of observations within a category. Percentages are calculated out of the available sample size for an age or overall (*n = 79*). *M* is the mean age at each data collection point, measured in months. *SD* is the standard deviation of the mean age at each data collection point, measured in months.

event remains on during the presentation of the lateral events (high competition trials) and serves as a distractor, and on the other half of trials, it is turned off at the onset of the lateral events (low competition trials). On each trial, the lateral events appeared to the left and right of the central event and depicted either women reciting stories in infant directed speech (social trials), or objects striking a surface (nonsocial trials). The movements of one of the two lateral events were synchronous with its natural soundtrack while the other lateral event was asynchronous with the soundtrack. For an example video, visit: https://nyu.databrary.org/volume/326. For the present study, we included a new measure indexing *distractibility* to the silent central distractor event during presentation of the audiovisual events on high competition trials. Low competition trials were not used in analyses because the central distractor event is not present during the presentation of the lateral events, and thus, no measure of distractibility can be derived.

The social and nonsocial blocks (12 trials per block; 6 high competition trials and 6 low competition trials) were designed to be analyzed separately if needed. For the purposes of the present study, the social trials were used given they present audiovisual speech events, and our purpose was to predict language outcomes and these trial types predicted language outcomes in our prior studies (Bahrick et al., 2018; Edgar et al., 2022). Performance on the MAAP nonsocial events were also evaluated to assess the specificity of our effects with social stimuli. For additional information regarding stimuli, procedures, and counterbalancing in the MAAP, see Bahrick et al. (2018).

The MAAP was coded in real-time by trained observers who coded the duration of infant looking to the left, right, and center of the screen as the infant viewed the video displays. Infants were required to have at least 2 trials per block, with a minimum of at least 250 ms of looking time to at least one of the three events (central, synchronous, asynchronous) on each trial. For the present study, we calculated two measures—intersensory matching of faces and voices (social events) and distractibility in the presence of social

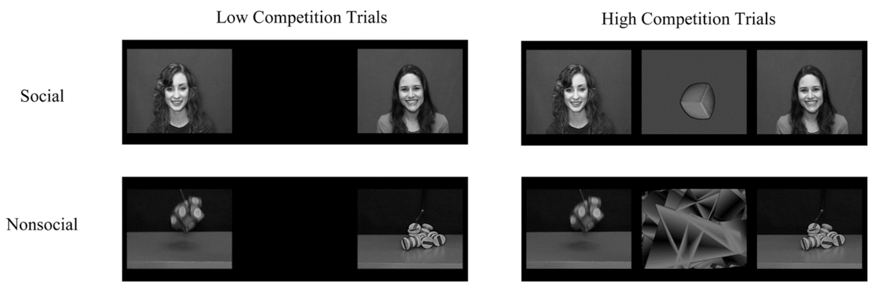

Fig. 1. Static images of the dynamic audiovisual events from the MAAP. On all trials, a three second central distractor event (computerized geometric shape) was followed by two side-by-side lateral events (social, nonsocial), one of which was synchronous with its appropriate soundtrack. On high competition trials (used in the present study; images on right), the central event remained on during the lateral events, providing a basis for assessing distractibility from the central distractor event. In contrast, low competition trials (images on left) provide no central distractor event and thus no basis for assessing distractibility.

events—on MAAP high competition trials (trials during which the central distractor event was available). Measures of intersensory matching and distractibility were calculated on each trial and then averaged to provide stable means for each measure (for descriptive statistics, see Table 2; preliminary analyses indicated that other MAAP measures, sustained attention and disengagement/shifting, were not related to our other predictors or outcomes). Intersensory matching was calculated in the standard way (see Bahrick, 1983; Bahrick et al., 2018; Spelke, 1976) as the proportion of the total looking time to the sound-synchronous social event divided by the total looking to the sound-synchronous and the sound-asynchronous events (PTLT Matching). PTLT Matching scores greater than 50% (chance value of looking to the sound-synchronous event out of two possible events) reflect preferences for the sound-synchronous social events (indexing face-voice matching). To ensure that infants attended to both left and right lateral events on each trial, trials on which infants spend the entire trial looking to one or the other event were excluded from analyses (for details, see Section 3.1.2). Our new measure of infant distractibility was operationalized as looking to a silent competing nonsocial visual distractor event during the presentation of two audiovisual social events (women speaking). Distractibility was calculated as the proportion of total looking time to the central distractor event divided by the total looking time to the central event and left and right audiovisual events (PTLT Distractor). Unlike PTLT matching (which is derived from attention to two events), PTLT Distractor is derived from attention to three events. Thus, PTLT Distractor scores greater than 33% (chance value of looking to the central distractor event out of three possible events) reflect preferences for the central distractor event in the presence of the audiovisual events. Interobserver reliability was assessed for 49% ($n = 39$) of participants on the MAAP at 12 months. Pearson correlation coefficients were calculated between primary and secondary observers: PTLT Matching = 0.93, PTLT Distractor = 0.93. In addition, to assess agreement between observers, we calculated the median absolute difference between primary and secondary observers' estimates. Estimates derived using this method are less subject to artifact than traditional estimates of reliability (Goodwin & Leech, 2006; Jaccard & Becker, 2009). Results indicated strong agreement (i.e., differences close to zero) between observers: PTLT Matching = 0.02 (range: .00 to 1.00), PTLT Distractor = 0.02 (range: .00 to 1.00).

2.2.2. Mother-Child Interaction

At 12 months, infants and their mothers participated in an 8-minute mother-child interaction (MCI). The MCI took place in a laboratory-based playroom with a 101.6 cm × 71.12 cm table in the center of the room. Mothers were seated across the table from their infants. Children were placed in a seat clamped to the edge of the table directly across from their mothers. Three toys were placed on the table between them (see Fig. 2). Mothers were instructed to "play with your child like you would at home." No specific instructions were given regarding how or whether to play with the toys.

The MCI was recorded by 3 video cameras (Canon VIXIA HF R50) simultaneously. One was placed behind the mother's shoulder to record the infant's face, one was placed behind the infant's shoulder to record the mother's face, and one was placed to the right of the table to provide a side-view of the interaction. Videos were synchronized using a mixer (Roland VR-3EX) and were coded for the number of *infant bids* and *maternal responses to infant bids* (for descriptive statistics, see Table 2). We modeled and adapted our coding scheme for infant bids based on the Early Social Communication Scales (Mundy et al., 2003) definitions for "initiating behavioral requests" (IBR) and "initiating joint attention" (IJA; see Mundy et al., 2003). To identify a bid, a trained observer watched the videotaped interaction and recorded each instance in which the infant directed two or more of the following behaviors—smiling, gazing, vocalizing, and gesturing—towards the mother. This definition was used to clearly operationalize a bid, and as noted earlier, to increase the likelihood that bids assessed were in fact directed toward the mother. Each infant bid was then coded for *maternal responses*. Maternal responses were categorized as accepted, redirected, or ignored based on mothers' behaviors in the 5 s following the end of each bid (similar to the time frame used by Tamis-LeMonda et al., 2001). Accepted responses occurred when the mother followed in appropriately to the infant's request for attention to a toy or interaction (e.g., the infant pointed and vocalized while looking at a toy and the mother handed the infant the toy). Redirected responses occurred when the mother failed to follow the infant's lead and directed the infant's attention to a new activity or object (e.g., the infant vocalized and pointed to a toy and the mother directed the infant's attention to another toy). Ignored responses occurred when the parent failed to acknowledge the bid for social interaction (e.g., the infant pointed and vocalized while looking at a toy and the mother did not acknowledge and continued her previous task). Interobserver reliability was assessed for 37% ($n = 22$) of the participants. Pearson correlation coefficients were calculated between primary and secondary coders: infant bids: .99, infant bids accepted: .98, redirected: .96, ignored: .89. In addition, median absolute differences between primary and secondary coders indicated strong agreement: infant bids = 0.04 (range: 0–14); infant bids accepted = 0.00 (range: 0–13); redirected= .00 (range: 0–3); ignored = .00 (range: 0–3).

2.2.3. The Mullen Scales of Early Learning

When infants were 18 months old, they were administered the Mullen Scales of Early Learning (MSEL; Mullen, 1995). The MSEL is a standardized assessment of nonverbal (Visual Receptive, Gross Motor, Fine Motor) and verbal (Receptive Language, Expressive Language) functioning appropriate for children from birth to 68 months of age (33 months of age for Gross Motor). In the present study, we used standardized *T* scores from the Receptive and Expressive Language scales, given our focus on predicting language outcomes. We also calculated the Early Learning Composite, a global index of cognitive functioning, to assess whether results were specific to language functioning or overall cognitive functioning (for descriptive statistics, see Table 2).

Table 2
Descriptive statistics—means (*M*), standard deviations (*SD*), and ranges—for measures from the Multisensory Attention Assessment Protocol (MAAP), the mother-child interaction (MCI) at 12 months of age, and the Mullen Scales of Early Learning (MSEL) at 18 months of age ($n = 79$).

	n	*M*	*SD*	Range
MAAP (12 months)	71			
Social Events				
Face-Voice Matching	62	0.50	0.11	0.32-0.92
Distractibility: Social	69	0.47	0.20	0.08-0.91
Nonsocial Events (control)				
Object-Sound Matching	66	0.53	0.12	0.22-0.73
Distractibility: Nonsocial	70	0.49	0.19	0.16-0.92
MCI (12 months)	59			
Infant Variables				
Infant Bids	59	5.90	4.52	0–22
Maternal Variables				
Bids Redirected	59	1.02	1.06	0–4
Bids Accepted	59	4.37	3.72	0–18
Bids Ignored	59	1.02	1.22	0–5
MSEL (18 months)	54			
ELC	54	87.83	19.14	49–130
Receptive Language	54	43.44	15.04	20–79
Expressive Language	54	43.89	12.51	20–80

Note: Sample sizes vary because some participants completed all assessments whereas others completed just one or two assessments. Total sample size was 79. Face-Voice Matching: percentage of total looking time to sound-synchronous speaking face in the presence of a distractor event. Distractibility Social: percentage of total looking time to the distractor event on social trials. Object-Sound Matching: percentage of total looking time to sound-synchronous moving object in the presence of a distractor event. Distractibility Nonsocial: percentage of total looking time to the distractor event on nonsocial trials. Infant bids: number of bids for attention infants produced. Bids Redirected: number of infant bids redirected by the mother. Bids Accepted: number of infant bids accepted by the mothers. Bids Ignored: number of infant bids ignored by the mother. Receptive Language: MSEL Receptive Language T scores. Expressive language: MSEL Expressive Language T scores. ELC: MSEL Early Learning Composite.

Fig. 2. Toys provided to infant-mother dyads during the Mother-Child Interaction (MCI) when infants were 12 months old: xylophone (left), wooden farm animal puzzle (center), and hard plastic interlocking blocks (right).

3. Results

3.1. Data analysis overview, available data, and descriptive statistics

After presenting descriptive statistics and available data, we first conducted bivariate correlations between all our measures to narrow down those for inclusion in our structural equation models (SEMs). Our primary analyses consisted of SEMs to assess pathways between infant bids, maternal responses to infant bids, infant attention measures, and infant language outcomes. SEMs were conducted using full information maximum likelihood (FIML) to address missing data and estimate pathways between model variables simultaneously (which is not possible with bivariate correlations).

3.1.1. Available data

All participants from the larger sample ($N = 104$) who had data for one or more of the variables of interest and whose mother (rather than father or grandparent) participated in the MCI ($n = 79$) were included in the analyses. Of the 79 total infants, 72 contributed data at 12 months (71 for MAAP and 59 for MCI) and 54 provided data for MSEL at 18 months (see Tables 2 and 3 for details). Thirty-four (43%) infants completed all three assessments (MAAP, Mother-Child Interaction, Mullen), 34 (43%) completed two assessments, and 11 (14%) completed one assessment. Overall, most infants completed at least two assessments (86%; see Table 3). Available data and descriptive statistics for each variable of interest are presented in Table 2. To address missing data, all analyses were conducted using Full Information Maximum Likelihood using the lavaan package (Rosseel, 2012) in R.

For MAAP, infants were required to have at least 2 trials per block, with a minimum of at least 250 ms of looking time to at least one of the three events (central, synchronous, asynchronous event) on each trial. On average, infants contributed useable data on 22.69 (*SD* = 3.17; range: 8–24) trials (94.5%) and spent 56.6% (*SD* = 11.5%; range: 27–86%) of their looking time attending to the displays. Also, for the PTLT Distractor measure, 69 (97.2%) infants contributed data for social events, and 70 (98.6%) contributed data for the

Table 3
Sample Sizes (n) and Percentages of Available ($n = 79$) and Missing Data ($n = 25$) for the Different Assessments Out of the Total Sample Enrolled ($N = 104$).

	12 months ($n = 72$)		18 months ($n = 54$)	
	n	Percent	n	Percent
MAAP				
Available data	71	68.3%	———	———
Missed visit or withdrew	16	15.4%	———	———
Data excluded (e.g., fussiness, sleepy, did not administer)	17	16.3%	———	———
MCI				
Available data	59	56.7%	———	———
Missed visit or withdrew	16	15.4%	———	———
Data excluded (e.g., mother not in MCI, AV issues, did not administer)	29	27.9%	———	———
MSEL				
Available data	———	———	54	51.9%
Missed visit or withdrew	———	———	32	30.8%
Data excluded (e.g., fussiness, did not administer)	———	———	18	17.3%

Note: Percentages are calculated out of the total sample enrolled, including participants who withdrew ($N = 104$). Out of the $n = 79$ infants, 72 participants contributed data at 12 months ($n = 71$ participants contributed useable MAAP data and $n = 59$ participants provided useable MCI data). However, 7 participants who did not contribute data at 12 months did contribute data at 18 months.

nonsocial events. For the PTLT Matching measure, in addition to the above criteria, infants were required to look to both lateral events on each trial, such that 62 (88.6%) infants contributed data for the social events, and 66 (93%) contributed data for the nonsocial event.

3.1.2. *Descriptive statistics and group-level analyses*

Descriptive statistics for the MAAP, MCI, and MSEL measures are presented in Table 2.

3.1.2.1. MAAP: face-voice matching and distractibility. To evaluate if infants showed significant matching of faces and voices (although not relevant to our primary research questions regarding developmental pathways based on analyses of individual differences), group level analyses of face-voice matching and distractibility for social events on high competition trials on the MAAP were conducted. For face-voice matching, at 12 months, infants spent an average of 50.48% of their looking time (to the sound synchronous and sound asynchronous events) during a trial fixating the sound-synchronous event (face-voice matching; $SD = 11.19\%$; See Fig. 3A). A single sample t-test revealed that, as a group, infants showed no evidence of face-voice matching (significant preference for the sound-synchronous event; 50% chance), $t(61) = 0.34$, $p = .74$. For distractibility, at 12 months, infants spent an average of 46.95% of their looking time (to the sound synchronous, sound asynchronous and distractor events) fixating the central distractor event in the presence of social events ($SD = 20.06\%$; See Fig. 3B). A single sample t-test revealed that, as a group, infants showed a significant preference for the central distractor event, $t(68) = 5.78$, $p < .001$, greater than chance (33%). Thus, as a group, infants showed no significant preference for the sound-synchronous event but did show a significant preference for the central distractor event. Importantly, there was substantial variability in PTLT Matching and PTLT Distractor scores that served as a meaningful basis for predicting individual differences.

Further, infants distributed their looking to the central distractor event, as well as the synchronous and asynchronous social events, across the entire length of each trial of the MAAP. We calculated total looking time (in seconds) and the number of looks to the three displays during the first and last half (6 s) of the 12 s MAAP trials. Results indicated that looking to all three displays was distributed across the entire trial, and the central distractor continued to disrupt attention to the synchronous and asynchronous events across the entire trial.[4]

[4] For social events, results indicated an average of 8.96 s (out of 12 s; $SD = 2.30$) of looking time across all three displays, with an average of 3.82 s ($SD = 1.54$) to the central distractor, 2.49 s ($SD = 1.28$) to the sound-synchronous event, and 2.31 s ($SD = 1.06$) to the asynchronous event. Results for the number of looks indicated a similar pattern, with an average of 8.81 looks ($SD = 5.79$) across all three displays, 3.63 looks ($SD = 2.78$) to the central distractor, 2.58 looks ($SD = 1.60$) to the sound-synchronous event, and 2.60 looks ($SD = 1.74$) to the asynchronous event. In order to determine how looking to the central distractor was distributed across trial time relative to looking to the sound-synchronous and asynchronous displays, we calculated looking time and the number of looks during the first and last half of each 12 s trial separately. Results indicated a significant decline in looking time to each of the three displays from the first to last half of the trial: central distractor (first half: $M = 2.89$ s, $SD = 1.41$; last half: $M = 1.49$ s, $SD = 0.84$), $t(77) = 10.53$, $p < .001$; sound-synchronous (first half: $M = 1.69$ s, $SD = 1.02$; last half: $M = 1.03$ s, $SD = 0.70$), $t(77) = 5.57$, $p < .001$; asynchronous (first half: $M = 1.61$ s, $SD = 0.86$; last half: $M = 0.88$, $SD = 0.52$), $t(77) = 7.46$, $p < .001$. Results for number of looks indicated a similar pattern of decline from the first to last half of the trial: central distractor (first half: $M = 2.10$, $SD = 1.79$; last half: $M = 1.76$, $SD = 1.16$), $t(77) = 2.50$, $p = .015$; sound-synchronous (first half: $M = 1.58$, $SD = 1.02$; last half: $M = 1.14$, $SD = 0.80$), $t(77) = 4.33$, $p < .001$; asynchronous (first half: $M = 1.64$, $SD = 1.19$; last half: $M = 1.10$, $SD = 0.76$), $t(77) = 4.94$, $p < .001$. In sum, looking to all three displays was distributed across the entire trial, and looking time and the number of looks appeared to decline across trial time in a similar manner for the central distractor, synchronous, and asynchronous events

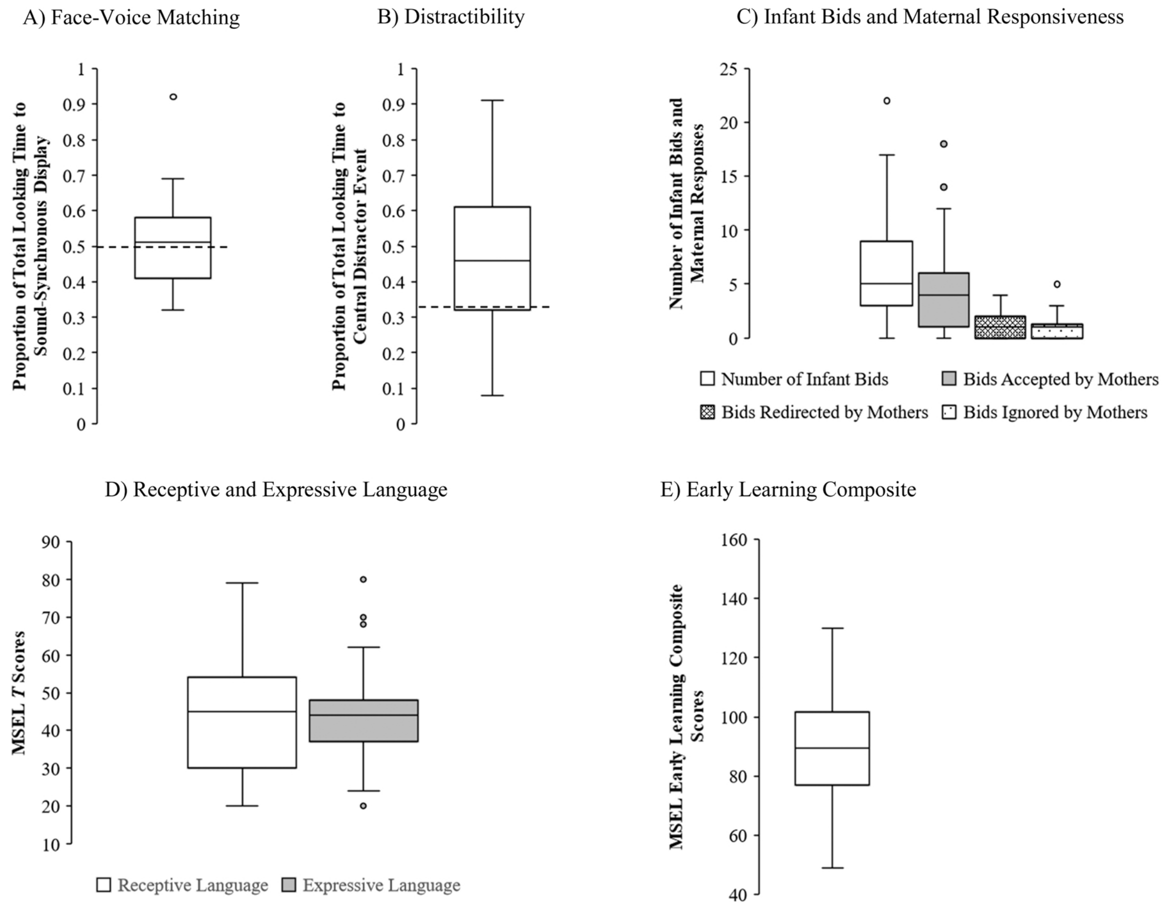

Fig. 3. Boxplots depicting distribution of scores for the variables in the study. Fig. 3A: Boxplot for Infant Face-Voice Matching with dashed lines depicting chance levels of looking (50%). Fig. 3B: Boxplot for Infant Distractibility from Social Events with dashed lines depicting chance levels of looking (33%). Fig. 3C: Boxplots for Number of Infant Bids, Bids Accepted by Mothers, Bids Redirected by Mothers, and Bids Ignored by Mothers. Fig. 3D: Boxplots for MSEL Receptive and Expressive Language Scores. Fig. 3E: Boxplot for the MSEL Early Learning Composite. The horizontal line within each box indicates the median value of the distribution. The upper and lower ends of the box indicate the 75th and 25th quartile, respectively. The whiskers (lines extending out of the box) indicate values that may lie 1.5 x the interquartile range (75th quartile minus 25th quartile)[31]

3.1.2.2. Infant bids for attention and maternal responses to infant bids. At 12 months, infants produced an average of 5.90 bids during the 8-minute MCI ($SD = 4.52$). Six infants (8%) did not make any bids for attention. However, these infants were included in our analyses as they provided data on our other measures (e.g., MAAP and/or MSEL data). Mothers accepted the overwhelming majority of bids produced by their infants. Out of an average of 5.90 bids, an average of 4.37 ($SD = 3.72$) were accepted (74%), 1.02 ($SD = 1.06$) were redirected (14%), and 1.02 ($SD = 1.22$) were ignored (12%; See Fig. 3 C). The majority of mothers (67.2%) redirected or ignored at least one of their infant's bids.

3.1.2.3. MSEL: Receptive, Expressive language, and ELC scores. The subscales of the MSEL have standardized T scores with a mean of 50 ($SD = 10$). In comparison, the average standardized T score for our sample for Receptive Language was 43.44 ($SD = 15.04$), and for Expressive Language was 43.90 ($SD = 12.51$) for Expressive Language (see Fig. 3D), both of which are significantly lower than norming sample, $ts > 3.21$, $ps < 0.002$. Four infants (7.7%) scored at floor on the Receptive Language scale (T score of 20), and two infants (4.1%) scored at floor on the Expressive Language scale. These infants were included in analyses. Also, MSEL standardized ELC score (a composite of Visual Reception, Fine Motor, Receptive Language, and Expressive Language scores) has a mean of 100 ($SD = 15$). In comparison, the average ELC scores in our sample was 87.83 ($SD = 19.14$; see Fig. 3E), which is significantly lower than the norming sample, $t(53) = 4.67$, $p < .001$.

3.2. Correlations among infant bids, maternal responses to infant bids, infant intersensory matching skills, distractibility, and language outcomes

We first conducted bivariate correlations as preliminary analyses to inform the construction of our SEMs that address our main research questions. We included variables with significant ($p < .05$) and marginally significant correlations ($p < .10$) with other variables in our SEMs. SEM allows us to take into account the effects of all variables simultaneously and assess pathways between variables, while at the same time compensating for missing data.

3.2.1. Correlations among infant bids, maternal responses to bids, and language outcomes

We first examined bivariate Pearson correlations among variables of interest (see Table 4), and we corrected for familywise error rates using a modified Bonferroni procedure (Holm, 1979).[5] Not surprisingly, during the MCI at 12 months, infants who produced a greater number of bids had a greater number of bids accepted, $r = 0.92$, $p < .001$, redirected, $r = 0.45$, $p = .002$ (see Fig. 4A), and ignored, $r = 0.48$, $p = .001$, by mothers.[6] Further, there was no evidence of significant correlations between infant bids at 12 months and MSEL Receptive Language, Expressive Language, and overall Early Learning Composite (ELC) scores at 18 months, or between maternal responses to infant bids at 12 months and MSEL scores at 18 months.

3.2.2. Correlations among intersensory matching, distractibility, infant bids, and maternal responses to infant bids

Infants who had more bids *redirected* by mothers during the MCI at 12 months showed lower attention to an irrelevant visual distractor on the MAAP high competition social trials at 12 months, $r = -0.39$, $p = .005$ (see Table 4 and Fig. 4B). In contrast, there was no evidence that infant attention to the distractor or face-voice matching on the MAAP at 12 months was significantly correlated with infant bids and other maternal responses to infant bids (accepted, ignored) during the MCI at 12 months.

3.2.3. Correlations among intersensory matching, distractibility, and language outcomes

On MAAP high competition social trials at 12 months, infants with greater face-voice matching showed lower attention to the irrelevant visual distractor (i.e., less distractibility) in the presence of social events, $r = -0.26$, $p = .03$ (see Table 4 and Fig. 4C). Infants who showed lower distractibility had higher ELC scores, $r = -0.30$, $p = .05$. However, when examining relations between infant distractibility from social events and the language subscales of the MSEL, distractibility was only correlated with Receptive (but not Expressive) Language scores, $r = -0.33$, $p = 0.03$ (this correlation failed to reach the cutoff for the adjusted familywise significance level: .05 / 2 = .025; see Fig. 4D).[7] In contrast, there was no evidence of significant correlations among face-voice matching at 12 months and MSEL Receptive or Expressive Language scores at 18 months.[8]

In sum, correlational analyses revealed that distractibility in the presence of social events at 12 months was the only significant predictor of language outcomes. However, the number of infant bids redirected (but not accepted or ignored) by mothers and infant face-voice matching predicted infant distractibility. Finally, the number of bids produced by infants predicted the number of infant bids redirected by mothers.

3.3. The relation between maternal responsiveness to infant bids and infant receptive language is mediated by infant distractibility in the presence of social events

Results of the prior correlational analyses served to narrow down and guide our decision about which variables to include in a model of developmental pathways between our measures (Cohen et al., 2003; Kline, 2005). Our model included pathways based on the pattern of significant ($p < 0.05$, or marginally significant ($p < .10$) correlations between predictor variables at 12 months (infant bids, maternal redirection of infant bids, infant face-voice matching, infant distractibility) and outcome variables at 18 months (receptive

[3] Though univariate outliers were evident in the sample, a greater concern for analysis of individual differences is the presence of bivariate outliers. We used the outliers/leverage approach recommended by Rousseeuw & van Zomeren (1990); see also Wilcox (2017). We found no instances of bivariate outliers in the data, and as a result, did not exclude any data from analyses.

[5] We created a family of two for correlations with our two language outcomes (Receptive, Expressive Language). Thus, the correlation with the smallest p value is compared against a critical value of $p < .025$ (0.05 / 2), and the remaining p value is tested against $p < .05$. We created a family of three for correlations with our maternal responsiveness variables (accept, redirect, ignore), so that the correlation with the smallest p value is compared against a critical value of $p < .017$ (0.05 / 3), the next smallest p value against $p < .025$, and the remaining p value against $p < .05$.

[6] Inspection of histograms revealed that only infant bids and maternal responses to infant bids (redirected, accepted, ignored) had skewed distributions. To assess whether the non-normality in these variables impacted the data, we conducted Spearman's rank correlations between these and the other variables. We found few differences between Spearman's and Pearson's correlations (small changes in magnitude and significance that did not change the overall relations between variables). Thus, we chose to keep Pearson correlations in Table 4 as they are more appropriate for ratio level data, whereas Spearman correlations are more appropriate for rank or ordinal level data.

[7] The correlation between infant distractibility from social events and MSEL Receptive Language remained significant event after holding MSEL Visual Reception constant, $r = -0.34$, $p = 0.03$. Thus, infant distractibility from social events specifically predicts language functioning, rather than general developmental functioning.

[8] Findings were similar regardless of whether MSEL standardized T scores or raw scores were used. Face-voice matching was not related to raw receptive ($r = .128$, $p = 0.426$) or raw expressive ($r = .152$, $p = 0.363$) language scores. The relation between raw receptive language scores and social distractibility was marginally significant ($r = -.273$, $p = .058$). Social distractibility was not related to raw expressive language scores ($r = -0.172$, $p = .276$). Object sound matching was significantly related to raw receptive ($r = -0.307$, $p = .030$) and raw expressive ($r = -0.307$, $p = .032$) language. Nonsocial distractibility was not significantly correlated with raw receptive ($r = -0.110$, $p = .433$) or raw expressive ($r = -0.045$, $p = .753$) language scores. Thus, relations between our attention measures and raw scores on the MSEL were very similar to the relations between our attention measures and standardized scores.

Table 4
Bivariate Pearson's correlation coefficients among variables from the Multisensory Attention Assessment Protocol (MAAP) and the mother-child interaction (MCI) at 12 months of age, and the Mullen Scales of Early Learning (MSEL) at 18 months of age.

	FVM	DSoc	OSM	DNon	Bids	Bids-R	Bids-A	Bids-I	RLang	ELang	ELC
MAAP (12 months)											
Social Events											
Face-Voice Matching (FVM)	–										
Distractibility: Social (DSoc)	-0.26[a]	–									
Nonsocial Events											
Object-Sound Matching (OSM)	0.06	-0.10	–								
Distractibility: Nonsocial (Dnon)	-0.03	0.36[b]	-.02	–							
MCI (12 months)											
Infant Variables											
Infant Bids (Bids)	-0.03	-0.20	-0.09	-0.13	–						
Maternal Variables											
Bids Redirected (Bids-R)	-0.06	-0.39[b]	-.08	-0.07	0.45[c]	–					
Bids Accepted (Bids-A)	-0.03	-0.14	-0.10	-0.11	0.92[c]	.15	–				
Bids Ignored (Bids-I)	-0.05	-0.03	-0.10	-0.02	0.48[c]	.40[b]	.17	–			
MSEL (18 months)											
Receptive Language (RLang)	0.19	-0.32[a]*f*	-0.30[a]*f*	0.18	-0.01	0.04	0.04	-0.20	–		
Expressive Language (ELang)	0.16	-0.18	-0.29[a]*f*	-0.07	-0.27	-0.05	-0.28	-0.31	0.60[c]	–	
Early Learning Composite (ELC)	0.30	-0.30[a]*f*	-0.35[a]	-.07	-0.23	0.06	-0.20	-0.34	0.77[b]	.74[b]	–

Note: Face-Voice Matching: percentage of total looking time to sound-synchronous speaking face in the presence of a distractor event. Distractibility Social: percentage of total looking time to the distractor event on social trials. Object-Sound Matching: percentage of total looking time to sound-synchronous moving object in the presence of a distractor event. Distractibility Nonsocial: percentage of total looking time to the distractor event on nonsocial trials. Infant bids: number of bids for attention infants produced. Bids redirected: number of infant bids redirected by the mother. Bids accepted: number of infant bids accepted by the mothers. Bids ignored: number of infant bids ignored by the mother. Receptive language: infants' standardized receptive language scores. Expressive language: infants' standardized expressive language scores. Early Learning Composite: Standardized index of infant cognitive functioning.

[a] $p < .05$,
[b] $p < .01$,
[c] $p < .001$; *f* did not meet familywise significance level of $p < .025$.

language). Infant distractibility in the presence of social events at 12 months was included as the main predictor of receptive language at 18 months (holding constant bids redirected by mothers and infant face-voice matching). Further, infant face-voice matching and bids redirected by mothers at 12 months were included as predictors of infant distractibility, and the number of infant bids was included as a predictor of bids redirected by mothers. This resulted in 2 developmental pathways between predictor variables and language, each mediated by infant distractibility: 1) infant bids for attention predicts bids redirected by mothers, which in turn predicts infant distractibility, which in turn predicts receptive language, and 2) infant face-voice matching predicts infant distractibility, which in turn predicts receptive language. Our proposed model is presented in Fig. 5.

We tested our proposed mediation model using structural equation modeling. We also tested several preliminary models but rejected them because they were not a good fit to the data.[9] We used the lavaan package (Rosseel, 2012) in R with full information maximum likelihood (FIML) estimation to address missing data. Indices of fit indicated that this mediation model was an excellent fit to the data, $X^2(5) = .883$, $p = .971$.[10] Bids Redirected by Mothers, Infant Face-Voice Matching, and Infant Distractibility from Social events explained 10% of the total variance in Infant Receptive Language. Also, Infant Face-Voice Matching and Bids Redirected by Mothers together explained 28% of the total variance in Infant Distractibility from Social events. Finally, Infant Bids explained 20% of the total variance in Bids Redirected by Mothers.

Results of the model confirmed our predictions that Infant Distractibility from Social events mediates the relation between our other predictors and Infant Receptive Language (see Fig. 3 for standardized and unstandardized path coefficients between variables). More Bids Redirected by Mothers predicted lower Infant Distractibility from Social Events, $p < 0.001$, which in turn predicted greater Infant Receptive Language, $p < 0.05$. However, Bids Redirected by Mothers was not a significant predictor of Infant Receptive Language. Also, greater Infant Face-Voice Matching predicted lower Infant Distractibility from Social events, $p < .01$, which in turn predicted greater Infant Receptive Language. However, Infant Face-Voice Matching was not a significant predictor of Infant Receptive

[9] For Preliminary Model 1, we tested a model in which all 12-month variables directly predicted Infant Receptive Language (no mediated relations). Only Infant Distractibility from Social emerged as a significant predictor of Infant Receptive Language, holding all other predictors constant, $p = 0.04$. For Preliminary Model 2, we tested a model in which Bids Redirected by Mothers mediated the relation between Infant Distractibility from Social and Infant Receptive Language. There was no evidence of a mediated effect, and only Infant Distractibility from Social significantly predicted Infant Receptive Language. We thus chose to move forward with our main model in which relations Infant Distractibility from Social events mediates relations between other 12-month predictors and Infant Receptive Language.

[10] Non-significant chi square tests reflect good model fit to the data. Other indices of fit also reflected excellent model fit: Comparative Fit Index (CFI) = 1.000; Tucker-Lewis Index (TLI) = 1.404; Standardized Root Mean Square Residual (SRMR) = .037, Root Mean Square Error of Approximation (RMSEA) < .001

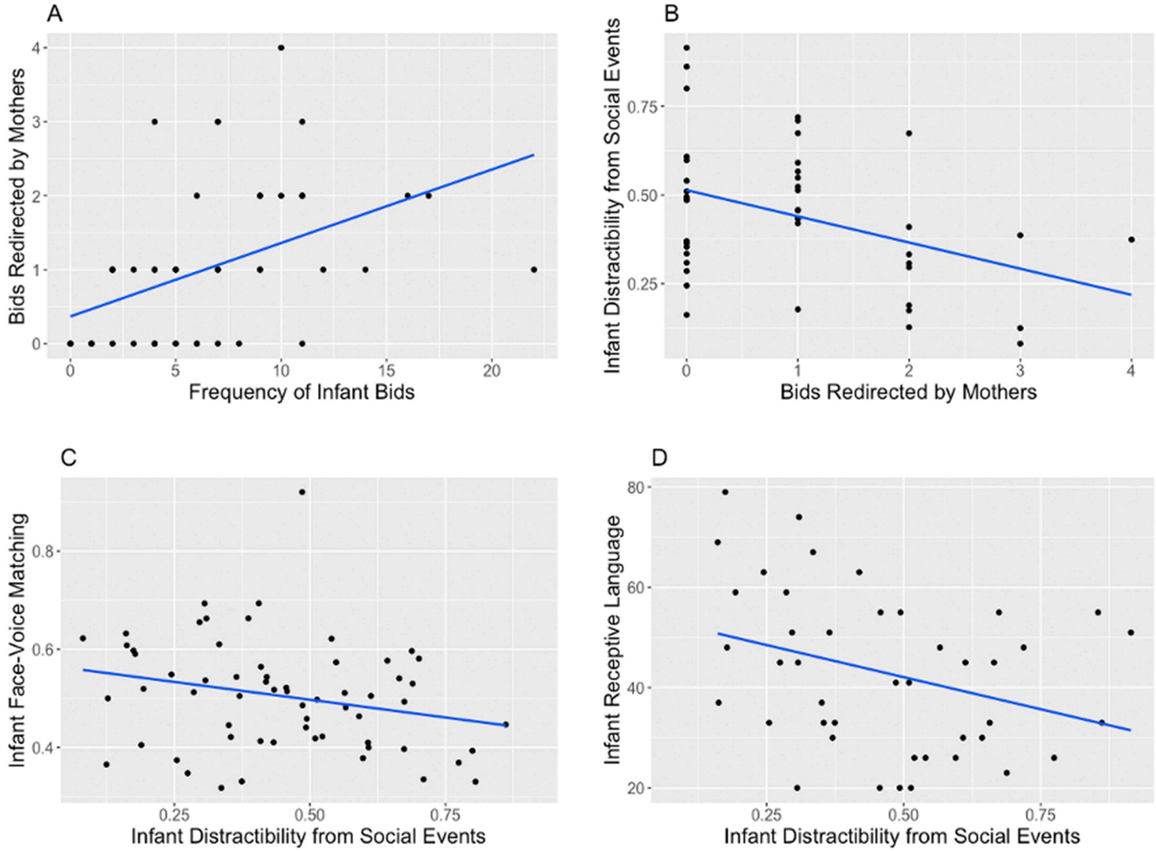

Fig. 4. Scatterplots depicting the relations between variables. Fig. 4A: Frequency of infant bids and the number of infant bids redirected by mothers. Fig. 4B: Number of infant bids redirected by mothers and infant distractibility from social events. Fig. 4C: Infant face-voice matching and infant distractibility from social events. Fig. 4D: Infant distractibility from social events and infant receptive language.

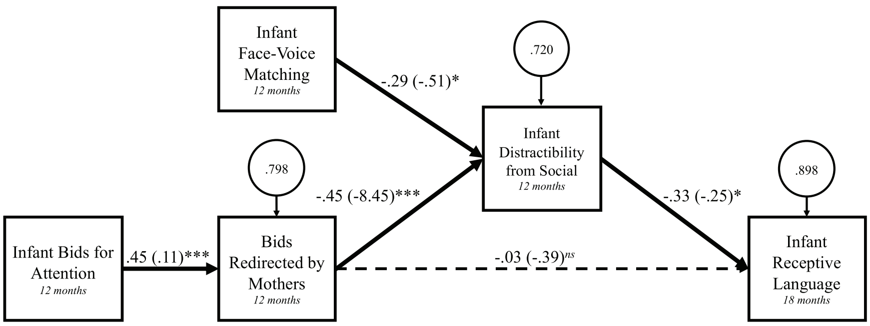

Fig. 5. Structural equation model: Infant Distractibility from Social events mediates the relation between predictors (Infant Bids for Attention, Bids Redirected by Mothers, Infant Face-Voice Matching) at 12 months and Infant Receptive Language outcomes at 18 months. Face-Voice Matching: percentage of total looking time to sound-synchronous speaking faces. Infant Bids for Attention: number of bids for attention infants produced. Bids Redirected by Mothers: number of infant bids redirected by the mother. Infant Receptive Language: infants' standardized Receptive Language scores from the MSEL. Rectangles represent observed variables, arrows represent path coefficients between variables, and circles represent error variances (proportion of unexplained variance). Solid lines represent significant paths, dashed lines represent non-significant paths. Standardized regression coefficients are presented outside of parentheses and unstandardized coefficients are presented in parentheses. $* p < .05$, $** p < .01$, $*** p < .001$, $^{ns} p \geq .10$.

Language.

Given our novel finding of a link between infant bids redirected by mothers and reduced infant distractibility, we explored our data set further for possible mechanisms underlying this relation. We reasoned that perhaps maternal redirecting of infant bids, at least within a moderate range for mothers who are generally responsive (accepting a majority of infant bids), reflects appropriate social scaffolding and/or promotes better infant attention control. If so, successfully redirecting infant attention following a bid may lead infants to disengage from their current focus of attention and look back to the mother more often for guidance. Although coding and analyzing infant behaviors (e.g., looks to mother, looks away from mother) following each maternal response to an infant bid (accept, redirect, ignore) is beyond the scope of this paper, we were able to assess the frequency of infant looks to the mother during the 8-minute MCI (the same interaction from which infant bids and maternal responses were measured). A frame-by-frame coding of infant gaze revealed that on average, infants looked to their mothers 29.80 times ($SD = 18.90$) during the 8-minute interaction. More bids redirected by mothers was marginally related to more infant looks to mothers ($r = .38, p = .06$). Also, infants who looked more to their mothers showed lower distractibility from social events on the MAAP ($r = -0.31 p = .05$), and greater receptive language scores on the MSEL at 18 months ($r = 0.36$, $p = .05$). Findings of these exploratory analyses suggest that more bids redirected by mothers during social interactions may be associated with greater infant social attention, and this is reflected both in greater looking to the mother during the MCI as well as greater looking to faces and voices (i.e., less distractibility) during the MAAP.

3.4. The specificity of distractibility to social events as a mediator of maternal responsiveness to infant bids and infant receptive language

To assess whether this mediated relationship is unique to distractibility in the presence of social events or a more general form of attention distractibility, we tested an alternative model that included infant distractibility in the presence of nonsocial events as a potential mediator between maternal responsiveness and receptive language. Our alternative model tested the same pathways as our primary model but used measures from the nonsocial trials of the MAAP (infant distractibility from nonsocial events, object-sound matching) to test this hypothesis. This model is presented in Fig. 6. Unlike our main mediation model, findings from this alternative model indicated that Bids Redirected by Mothers was not a significant predictor of Infant Distractibility from Nonsocial events, which in turn was not a significant predictor of Infant Receptive Language. Also, Infant Object-Sound Matching was not a significant predictor of Infant Distractibility from Nonsocial events. Together, the findings from our main and alternate model demonstrate that the mediating role of distractibility between maternal responsiveness and receptive language is only evident when distractibility is assessed in the context of social events.

4. Discussion

The current study investigated how maternal responses to infant bids for attention, infant attention to audiovisual speech events (face-voice matching), and infant distractibility from social events at 12 months relate to infant receptive language outcomes at 18 months. Similar to prior studies (e.g., Kannass et al., 2006; Salley et al., 2013), we defined distractibility as looking to a distractor in the presence of target events. Specifically, distractibility was operationalized as looking to a silent competing nonsocial visual distractor event during the presentation of two audiovisual social events (women speaking).

A number of novel findings emerged. First, we found that infant distractibility from social (but not nonsocial) events predicted infant receptive language, and that it mediated the relations between our other potential 12-month predictors of language outcomes (infant bids, maternal responses to those bids, and infant face-voice matching) and infant receptive language at 18 months. This mediated pathway was only evident for infant distractibility from social, but not nonsocial, events, indicating that this pattern of relations is specific for attention to social events and not a general attentional process. Second, we found that the number of infant bids redirected by mothers (but not bids accepted or ignored) was the only measure of maternal responsiveness that predicted infant distractibility. Findings indicate that redirecting infant bids for attention, by mothers who are generally responsive to their infants (i.e., accept a majority of infant bids by following in), is associated with greater attentional control, and reduced distractibility, in infants. Infants whose mothers redirect more bids may be more attentive (i.e., less distractible) to social information from their mothers or to

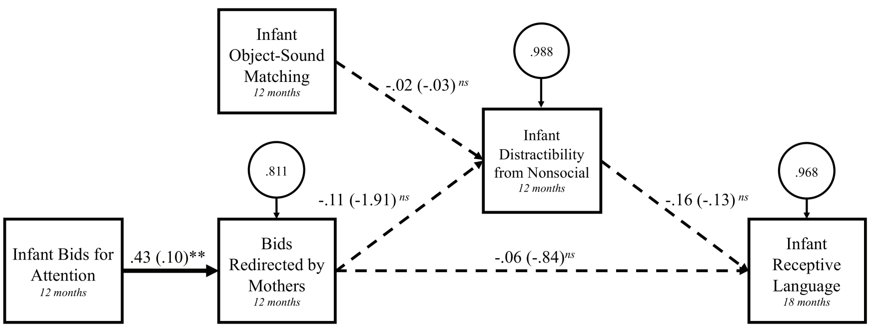

Fig. 6. Alternative Model: Infant Distractibility from Nonsocial events does not significantly mediate the relation between predictors (frequency of Infant Bids for attention, Bids Redirected by mothers, Infant Object-Sound Matching) at 12 months and Infant Receptive Language outcomes at 18 months. Object-Sound Matching: percentage of total looking time to sound-synchronous moving objects. Infant bids: number of bids for attention infants produced. Bids redirected: number of infant bids redirected by the mother. Receptive language: infants' standardized Receptive Language scores from the MSEL. Rectangles represent observed variables, arrows represent regressions, circles represent error variances. Solid lines represent significant paths, dashed lines represent non-significant paths. Standardized regression coefficients are presented with unstandardized coefficients in parentheses. * * $p < 0.01$, ^{ns} $p \geq .10$.

the faces and voices of people in the environment, which may allow infants to take greater advantage of language learning opportunities.

4.1. Infant distractibility from social events mediates the relation between maternal responsiveness and language outcomes

Results from our correlational analyses revealed that out of all of our measures at 12 months, only infant distractibility from social events predicted infant receptive language at 18 months (though distractibility did predict overall ELC scores, this relation was driven by the relation between distractibility and receptive language, but not expressive language, fine motor, and visual reception). In contrast, infant bids, maternal responses to bids, and infant face-voice matching were not significantly related to receptive language. Instead, these measures predicted infant distractibility. At 12 months, the number of infant bids redirected by mothers as well as infant face-voice matching significantly predicted infant distractibility from social events, and the number of bids infants made significantly predicted the number of bids mothers redirected. Using these correlations as a guide, we developed and tested a model in which infant distractibility from social events mediated the relation between the other predictors and infant receptive language. Two different pathways to receptive language at 18 months emerged. First, more infant bids predicted more infant bids redirected by mothers, which predicted lower infant distractibility from social events, which in turn predicted greater receptive language. Second, greater infant face-voice matching predicted lower infant distractibility from social events, which in turn predicted greater receptive language. Thus, the developmental pathways from maternal responsiveness and infant face-voice matching to infant language were mediated by infant distractibility from social events. At 12 months, infants whose mothers redirected more of their bids for attention and who had better intersensory matching of faces and voices showed lower attention to an irrelevant visual distractor in the presence of social events on the MAAP, and in turn, at 18 months, showed better receptive language outcomes.

Maternal responsiveness is a well-established predictor of child language outcomes. However, unlike prior research, our findings showed no evidence of a direct relationship between maternal responsiveness and infant language outcomes (e.g., Tamis-LeMonda et al., 1998; Tamis-LeMonda et al., 2001). Instead, our measure of maternal responsiveness, redirecting infant bids (but not accepting or ignoring infant bids), was indirectly related to receptive language through infant distractibility from social events.

There were several differences between our study and prior studies that likely contributed to the different findings. First, maternal responsiveness has been measured in a variety of ways, including global measures of maternal sensitivity, verbal responses to infants' actions, and responses to infant bids for attention. In the present study, we assessed maternal responses to infant bids for attention, a measure few studies have used (but see Tamis-LeMonda et al., 2001). Second, our definition of infant bids was stricter than that of other studies: we required that infants combine two or more behaviors (e.g., vocalizing plus looking to mother), whereas other studies defined bids as behaviors in just a single modality (e.g., looks to mother). Third, unlike prior studies that have found links between maternal redirection and language outcomes (e.g., Saxon, 1997; Shimpi & Huttenlocher, 2007), we defined maternal responsiveness as whether and how often mothers accepted, redirected, or ignored infant bids for attention. Finally, we focused on infants of 12 months, and few studies have assessed relations between maternal responsiveness in infants of 12-month-olds or younger and later language outcomes (for a review, see Tamis-LeMonda at el., 2014).

4.2. Infant distractibility mediates the relation between infant face-voice matching and language outcomes

Why might infant distractibility from social events predict language outcomes? Prior research has demonstrated that developments in attentional control are characterized by improvements in sustaining attention as well as decreases in distractibility with age (e.g., Colombo, 2001). We think of sustained attention and distractibility as complementary skills that both reflect attentional control. Infants must learn to maintain attention to relevant events and their properties, while at the same time inhibiting attention to irrelevant events and properties. We predicted that intersensory matching and distractibility would be related given that both reflect attentional control. Consistent with our predictions, greater intersensory matching of faces and voices was associated with less attention to an irrelevant central distractor on the MAAP, which was in turn associated with better infant receptive language outcomes. Prior research has demonstrated that attention to social events, particularly redundant audiovisual speech across faces and voices, directly predicts language outcomes, particularly expressive language (Bahrick et al., 2018; Edgar et al., 2022; for reviews, see Bahrick et al., 2020; Bahrick & Todd, 2012). In contrast, our findings are the first to demonstrate the intersensory matching of faces and voices is indirectly related to receptive language through distractibility from social events. Together with prior findings, our current findings suggest that different measures of infant attention may predict different aspects of infant language outcomes. For example, better attentional control and lower distractibility may be a stronger predictor of receptive language (understanding the meaning of words), whereas greater intersensory matching of social events may be a better predictor expressive language (understanding and producing words; Edgar et al., 2022; Bahrick et al., 2018). Given that the MAAP now provides a fine-grained assessment of attention to audiovisual events, characterizing these developmental pathways will be an important goal for future research.

4.3. Infant distractibility to social, but not nonsocial, events predicts language outcomes

Our main model assessed relations between infant attention to social events (distractibility and face-voice matching) and receptive language. To assess whether the developmental pathways from infant attention to receptive language were specific to social events or reflected more general attentional processes, we tested an alternative model in which we included distractibility from nonsocial events and infant intersensory matching of nonsocial events (object-sound matching). Unlike our main model, there was no evidence that infant distractibility from nonsocial events predicted infant receptive language. Further, there was no evidence that the number of infant bids redirected by mothers or infant intersensory matching of nonsocial events predicted infant distractibility from nonsocial events. Together, findings across the two models indicate that: maternal responsiveness (redirecting) to infant bids and intersensory matching in social contexts (face-voice matching) at 12-months are related to infant attention (distractibility) during social but not nonsocial events at 12-months, and that individual differences in infant distractibility during social, but not nonsocial, events are related to infant receptive language outcomes at 18-months of age. Thus, attentional control in the context of social events (e.g., focusing attention on faces and voices while ignoring irrelevant distractors) may lead to increased opportunities for language learning. This is not to say that greater attentional control and less distractibility in the presence of nonsocial events is unimportant in early development: it may predict other developmental outcomes, such as nonverbal cognitive functioning, or it may predict outcomes at earlier ages.

4.4. Relations among infant bids, maternal responsiveness, infant distractibility, and infant language

Prior research demonstrates that prompt, contingent, and appropriate maternal responsiveness is associated with better language outcomes (Tamis-LeMonda et al., 2001; Baumwell et al., 1997). In contrast with previous research, maternal responsiveness did not directly predict language outcomes in our study. Rather, it predicted infant distractibility from social events, which in turn predicted receptive language. Further, only one measure of maternal responsiveness—infants bids redirected by mothers—predicted infant distractibility from social events. Why might maternal redirection (rather than acceptance) of infant bids in our sample of generally responsive mothers predict infant distractibility and in turn receptive language?

In this present study, maternal redirections of infant bids for attention occurred when the mother failed to follow the infant's lead and instead directed the infant's attention to a different activity (e.g., the infant vocalized and pointed to a toy and the mother directed the infant's attention to another toy). Our findings indicate that the infants of these mothers were less distracted during social events on the MAAP, which means they had greater attentional resources for sustaining attention and further processing the social events (including synchronous faces and voices). There are several possible explanations for this finding. The first, and most compelling, is that maternal redirection of infant bids leads to greater attentional control in infants. When an infant's bid for attention is redirected by the mother, the infant may attempt to re-engage the mother's attention and establish joint attentional focus. Infants of mothers who are generally responsive (accepting a majority of infant bids and redirecting fewer) may be adept at orienting to mothers' occasional redirections and re-establishing joint attention. In our sample, the vast majority of infant bids were accepted by mothers (74%). This process of re-establishing shared attention following a maternal redirection may entail greater control of attention. Consistent with this proposal, Mason and colleagues (Mason et al., 2019) found that infants of highly responsive caregivers (as measured by sensitive responses to infant's vocalizations and looking to the caregiver) were quicker to attend to their caregivers' occasional redirections, suggesting greater attention control. Our exploratory analyses show some support for this hypothesis. In the 8-minute mother-child interaction, a greater number of infant bids redirected by mothers was associated with a greater frequency of infant looks to their mother, suggesting infants were prioritizing attention to mothers in order to re-establish joint attention. Further, infants who made more shifts to their mother's face showed lower distractibility from social events on the MAAP and had better receptive language scores at 18 months.

Another explanation for the link between maternal redirecting infant bids and reduced infant distractibility is that there may be an optimal midrange in which maternal redirecting predicts greater attentional control and reduced distractibility. Too little redirection may not foster enough engagement with a social partner, leading to fewer opportunities for training attention control, and too much redirection may be a deterrent to infant initiation of interactions with a social partner. A mid-range of attentional redirection by caregivers may thus be optimal for scaffolding attention control in infants, leading to better sustained attention to social events and lower attention to irrelevant distractors. Further, infants may benefit from the variability of both following in and redirecting to teach them how to interact with a social partner in a reciprocal manner. Interactions with a social partner are generally characterized by a pattern of turn taking. Attentional redirecting may help infants learn to attend to the attentional focus of others, while acceptance of infant bids for attention may help infants learn that others can share the focus of their attention.

4.5. Limitations and future directions

This study has several limitations. First, there was a relatively small sample ($n = 34$) of infants whose mothers redirected their bids. Future research will assess whether this distribution of maternal redirecting, accepting, or ignoring infant bids is evident at younger (e.g., 6 months) or older (e.g., 18 months) ages to better characterize the relation between maternal responsiveness to infant bids and infant attentional control. Second, maternal responsiveness and infant attention were only assessed at 12 months of age, and infant language outcomes were only assessed at 18 months of age. Thus, it is unclear, whether infant language at an earlier age predicts maternal responsiveness, infant attentional control, and later language outcomes. Consistent with prior research that has demonstrated a social feedback loop between infant vocalizations and maternal responsiveness (e.g., Warlaumont et al., 2014), future research will characterize the development of maternal responsiveness, infant attentional control, and early language outcomes across the first 2 years of life. Third, the majority of our participants were Hispanic, and thus the extent to which our findings would generalize to infants of other ethnicities is not known. Although there were no differences between Hispanic and non-Hispanic participants in terms of maternal sensitivity, infant distractibility, and infant receptive language, future studies could examine these differences in a larger sample. Fourth, on average, receptive and expressive language scores, as well as overall scores, from the MSEL were significantly lower than the norming sample and thus the extent to which these findings might generalize to a sample with higher language and overall cognitive scores is unknown. Finally, mothers' pattern of responding to infant bids may have been influenced by the lab-based setting of the mother-child interaction in our study. Mothers may have been more likely to redirect the focus of their infants' attention to encourage the infant to engage with a specific object or task, since they were aware of being observed in the lab setting. It is unclear the extent to which this relation between maternal redirecting infant attention and attention control/distractibility will be evident in other contexts (e.g., in the home).

5. Conclusion

This study adds to our understanding of the relations between maternal responsiveness, infant attention, and infant language outcomes. Maternal responsiveness to infant bids (number of bids redirected by the mother) was shown to positively influence infant distractibility from social events. In turn, lower distractibility in infants was positively related to infant language. This suggests that maternal responses to infant's bids may influence language through infant attention skills. Further, greater accuracy of intersensory matching in infants contributed to lower distractibility. These findings suggest that occasional and successful redirection of infant attention by mothers may promote better attention control and, in turn, better language outcomes. These findings will further our understanding of the relation between maternal responsiveness and infant attention, and their joint influence on infant language development.

Declaration of Competing Interest

None.

Data Availability

Data are not available but may be available in the future.

Acknowledgements

This work was supported by National Institutes of Health (NIH) Grants R01-HD053776 and R01-HD094803 awarded to the last author. We thank our research team for their assistance with data collection and coding, including Ruth Zuraw-Moya, Maria Rodriguez-Vargas, Amanda Delgado, Kaityn Contino, Amanda Costales, Erica Barr, Valeria Martinez Villalba, & Kasey Soska.

Author statement

This research was conducted in accordance with APA ethical standards in the treatment of the study sample. The authors have no conflicts of interest to declare.

References

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language, 18*, 41–49.

Bahrick, L. E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development, 6*, 429–451.

Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. In G. Bremner, & T. D. Wachs (Eds.), *The Wiley-Blackwell handbook of infant development* (second ed., pp. 120–166). Oxford, England: Blackwell Publishing.

Bahrick, L., Lickliter, R., & Todd, J. (2020). The Development of Multisensory Attention Skills: Individual Differences, Developmental Outcomes, and Applications. In J. Lockman, & C. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development: Brain, behavior, and cultural context (Cambridge Handbooks in Psychology* (pp. 303–338). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108351959.011.

Bahrick, L. E., & Todd, J. T. (2012). Multisensory processing in autism spectrum disorders: Intersensory processing disturbance as a basis for atypical development. In B. E. Stein (Ed.), *The new handbook of multisensory processes* (pp. 657–674). Cambridge, MA: MIT Press.

Bahrick, L. E., Todd, J. T., & Soska, K. C. (2018). The multisensory attention assessment protocol (MAAP): Characterizing individual differences in multisensory attention skills in infants and children and relations with language and cognition. *Developmental Psychology, 54*(12), 2207–2225.

Baumwell, L., Tamis-LeMonda, C. S., & Bornstein, M. H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant Behavior and Development, 20*(2), 247–258.

Bigelow, A. E., & Power, M. (2016). Effect of maternal responsiveness on young infants' social bidding-like behavior during the still face task. *Infant and Child Development, 25*, 256, 276.

Bornstein, M. H., Putnick, D. L., Hahn, C. S., Tamis-LeMonda, C. S., & Esposito, G. (2020). Stabilities of infant behaviors and maternal responses to them. *Infancy, 25*, 226–245.

Casey, B., & Richards, J. E. (1988). Sustained visual attention in young infants measured with an adapted version of the visual preference paradigm. *Child Development, 59*(6), 1514–1521.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analyses for the behavioral sciences* (*third ed.*). Lawrence Erlbaum Associates.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*, 337–367.

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and childhood. In R. Kali (Ed.), *Advances in child development and behavior* (Vol. 34, pp. 283–322). San Diego, CA: Academic Press.

Dixon, W., & Salley, B. J. (2007). "Shhh! We're tryin' to concentrate.": Attention and environmental distractors in novel word learning. *The Jounral of Genetic Psychology, 167*(4), 393–414.

Deans, C. L. (2020). Maternal sensitivity, its relationship with child outcomes, and interventions that address it: A systematic literature review. *Early Child Development and Care, 190*(2), 252–275.

Edgar, E. V., Todd, J. T., & Bahrick, L. E. (2022). Intersensory matching of faces and voices in infancy predicts language outcomes in young children. *Developmental Psychology, 58*(8), 1413–1428.

Ekas, N. V., Haltigan, J. D., & Messinger, D. S. (2013). The dynamic still-face effect: Do infants decrease bidding over time when parents are not responsive? *Developmental Psychology, 49*(6), 1027–1035.
Enders, C. (2010). *Applied missing data analysis* (2nd ed). Guilford Press.
Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *PNAS, 100*(13), 8030–8035.
Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*(5), 515–523.
Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *The Journal of Experimental Education, 74*, 249–266. https://doi.org/10.3200/JEXE.74.3.249-266
Graham, J. W., & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In *Statistical strategies for small sample research* (pp. 1–29). Sage.
Jaccard, J. J., & Becker, M. A. (2009). *Statistics for the behavioral sciences* (5th ed.). Wadsworth Publishing.
Kannass, K. N., Oakes, L. M., & Shaddy, D. J. (2006). A longitudinal investigation of the development of attention and distractibility. *Journal of Cognition and Development, 7*(3), 381–409.
Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage,.
Laakso, M. L., Poikkeus, A. M., Eklund, K., & Lyytinen, P. (1999). Social interactional behavior and symbolic play competence as predictors of language development and their association with maternal attention directing strategies. *Infant Behavior and Development, 22*(4), 541–556.
Lansink, J. M., & Richards, J. E. (1997). Heart rate and behavioral measures of attention in six-, nine-, and twelve-month-old infants during object exploration. *Child Development, 68*(4), 610–620.
Leerkes, E. M., Blankson, A. N., & O'Brien, M. (2009). Differential effects of maternal sensitivity to infant distress and nondistress on social-emotional functioning. *Child Development, 80*(3), 762–775.
Legerstee, M., Varghese, J., & van Beek, Y. (2002). Effects of maintaining and redirecting infant attention and the production of referential communication in infants with and without down syndrome. *Journal of Child Language, 29*, 23–48.
Mason, G. M., Kirkpatrick, F., Schwade, J. A., & Goldstein, M. H. (2019). The role of dyadic coordination in organizing visual attention in 5-month-old infants. *Infancy*, (162 186).
Mendive, S., Bornstein, M. H., & Sebastian, C. (2013). The role of maternal attention-directing strategies in 9-month-old infants attaining joint engagement. *Infant Behavior and Development, 36*, 115–123.
Mullen, E. M. (1995). *Mullen Scales of Early Learning (AGS ed.)*. Circle Pines, MN: American Guidance Service Inc.
Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development, 78*(3), 938–954.
Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., & Seibert, J. (2003). *Early social communication scales (ESCS)*. Coral Gables, FL: University of Miami.
Mundy, P., Kasari, C., Sigman, M., & Ruskin, E. (1995). Nonverbal communication and early language acquisition in children with down syndrome and in normally developing children. *Journal of Speech and Hearing Research, 38*, 157–167.
Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders, 20*(1), 115–128.
Richards, J. E., & Casey, B. (1991). Heart rate variability during attention phases in young infants. *Psychophysiology, 28*(1), 43–53.
Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://www.jstatsoft.org/v48/i02/.
Rousseeuw, P. J., & van Zomeren, B. C. (1990). Unmasking Multivariate Outliers and Leverage Points. *Journal of the American Statistical Association, 85*(411), 633–639. https://doi.org/10.1080/01621459.1990.10474920.
Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology, 39*(5), 877–890.
Salley, B., Panneton, R. K., & Colombo, J. (2013). Separable attentional predictors of language outcomes. *Infancy, 18*(4), 462–489.
Saxon, T. F. (1997). A longitudinal study of early mother-infant interaction and later language competence. *First Language, 17*, 271–281.
Shimpi, P. M., & Huttenlocher, J. (2007). Redirective labels and early vocabulary development. *Journal of Child Language, 34*, 845–859.
Smith, K. E., Landry, S. H., & Swank, P. R. (2006). The role of early maternal responsiveness in supporting school-aged cognitive development for children who vary by birth status. *Pediatrics, 117*(5), 1608–1617.
Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8*(4), 553–560.
Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748–767.
Tamis-LeMonda, C. S., Bornstein, M. H., Kahana-Kalman, R., Baumwell, L., & Cyphers, L. (1998). Predicting variation in the timing of linguistic milestones in the second year: An events-history approach. *Journal of Child Language, 25*, 675–700.
Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language facilitated by parental responsiveness? *Current Directions in Psychological Science, 23*(2), 121–126.
Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development, 57*(6), 1454–1463.
Van Hecke, A. V., Mundy, P. C., Acra, C. F., Block, J. J., Delgado, C. E. F., Parlade, M. V., Neal, A. R., Meyer, J. A., & Pomares, Y. B. (2007). Infant joint attention, temperament, and social competence in preschool children. *Child Development, 78*(1), 53–69.
Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and it's reduction in autism. *Psychological Science, 25*(7), 1314, 1324.
Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing* (4th edition). Academic Press.
Wilson, S., & Durbin, C. E. (2013). Mother-child and father-child dyadic interaction: Parental and child bids and responsiveness to each other during early childhood. *Merril-Palmer Quarterly, 59*(3), 249–279.
Yu, C., Suanda, S. H., & Smith, L. B. (2018). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science, 22*(1), 1–12.