Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-022-05752-3

ORIGINAL PAPER

Individual Differences in Multisensory Attention Skills in Children with Autism Spectrum Disorder Predict Language and Symptom Severity: Evidence from the Multisensory Attention Assessment Protocol (MAAP)

James Torrence Todd[1] · **Lorraine E. Bahrick**[1]

Accepted: 7 September 2022

Abstract

Children with autism spectrum disorders (ASD) show atypical attention, particularly for social events. The new Multisensory Attention Assessment Protocol (MAAP) assesses fine-grained individual differences in attention disengagement, maintenance, and audiovisual matching for social and nonsocial events. We investigated the role of competing stimulation on attention, and relations with language and symptomatology in children with ASD and typical controls. Findings revealed: (1) the MAAP differentiated children with ASD from controls, (2) greater attention to social events predicted better language for both groups and lower symptom severity in children with ASD, (3) different pathways from attention to language were evident in children with ASD versus controls. The MAAP provides an ideal attention assessment for revealing diagnostic group differences and relations with outcomes.

Keywords Multisensory attention skills · Individual differences · Social attention · Intersensory processing · Disengagement · Symptom severity

Introduction

Our world provides far more stimulation than can be attended at any given moment. Learning to selectively attend to relevant, meaningful information while filtering out irrelevant information is critical for the veridical development of perception (e.g., E. J. Gibson, 1969). Selective attention acts as a gateway for all that we perceive, learn, and remember. However, the development of selective attention and its relations with developmental outcomes, such as language and social functioning, remains poorly understood. Attention is a multi-faceted construct (e.g., Colombo, 2001; Ruff & Rothbart, 1996). However, various selective attention skills are typically studied in isolation, including sustaining attention to objects and events in the presence of competing stimulation, disengaging attention away from competing stimulation, and intersensory matching by selectively attending to temporally synchronous sights and sounds from events. Thus, little is known about which skills are foundational for which developmental outcomes, how different attention skills are interrelated, and how they cascade into more complex, later developing abilities such as language and social development. This is in part due to the lack of reliable individual difference measures of attention to multisensory social and nonsocial events for preverbal and nonverbal children, both typically developing and those with ASD. Despite a diverse literature on attention, most research has focused on developmental differences in these skills at the group level, limiting our knowledge of how they lead to later developmental outcomes, particularly language and social functioning, in typical and atypical development. Thus, reliable individual difference measures of attention to multisensory events are needed to characterize developmental trajectories and cascades, and to relate an individual's score to their performance on language or social outcomes. In the absence of individual difference measures, knowledge

Portions of these data were presented at the May 2009, 2010, 2011, 2012, 2013, and 2014 International Meeting for Autism Research.

✉ James Torrence Todd
jtodd@fiu.edu

[1] Department of Psychology, Florida International University, 11200 South West 8 Street, Miami, FL 33199, USA

Published online: 01 October 2022

about how disturbances in these foundational skills cascade to atypical social and language development in ASD is limited (e.g., Bahrick & Todd, 2012; Mundy & Burnette, 2005).

The present study uses a new measure designed to assess individual differences in what we have termed "multisensory attention skills" (MASks; Bahrick et al., 2020)—attention maintenance (duration), attention shifting/disengaging (speed), and intersensory matching[1] (accuracy)—to address these important questions. The Multisensory Attention Assessment Protocol (MAAP; Bahrick et al., 2018b) assesses these three MASks to audiovisual social and nonsocial events in a single protocol. Further, the MAAP assesses the cost of competing stimulation on each of the MASks. All three MASks reflect selective attention to audiovisual events. Attention maintenance indexes the duration of attention to audiovisual events in the presence of an irrelevant visual distractor. Shifting/disengaging indexes the speed of disengaging and shifting attention from competing stimulation to audiovisual events. Intersensory matching indexes the accuracy of matching synchronous audible and visual stimulation while filtering out stimulation from a concurrent asynchronous event, as well as a visual distractor. All three MASks reflect both exogenous ("bottom up", reflexive attention) and endogenous ("top down" control of attention). Examining this fundamental combination of skills in a single protocol can reveal more about how attention is allocated in the natural environment and in children with ASD. Further, by including multiple measures of attention in a single protocol, the MAAP can characterize patterns of attentional strengths and weaknesses in children with ASD, and which attention skills are most predictive of language outcomes and symptom severity. Here, we assess how the MAAP distinguishes 2–5-year-old children with ASD from a group of nonverbal mental age (MA) matched typically developing (TD) controls, and which MASks best predict language outcomes in children with ASD and TD children, as well as symptom severity in children with ASD.

We first detail our theoretical framework regarding the development of selective attention in TD children—the Intersensory Redundancy Hypothesis (IRH). We next summarize the developmental importance of MASks in TD children and summarize the prior literature on each MASk (attention maintenance, shifting/disengaging, intersensory matching) in children with ASD. We then describe our theoretical framework characterizing the nature of attention in children with ASD—the Intersensory Processing Disturbance Hypothesis (IDH). Finally, we introduce the MAAP and explain why it provides a more comprehensive individual difference measure of attention skills to audiovisual events than existing paradigms.

[1] We use "multisensory" as a general term to refer to stimulation impacting more than one sensory system (e.g., auditory, visual, proprioceptive, etc.). It serves as a name for our protocol (MAAP) and for the collection of the three attention skills it measures ("multisensory attention skills" (MASks): attention maintenance, shifting/disengaging, intersensory matching). In contrast, the term "intersensory" is more specific and is used here to refer to just one of these skills—intersensory matching. Intersensory matching is the detection of information that is common across auditory and visual stimulation such as synchrony, rhythm, tempo, or intensity patterns. We also refer to intersensory processing as the activity of perceiving, integrating, and further processing this information.

The Intersensory Redundancy Hypothesis: A Theory of Selective Attention

Our theory of selective attention, the *intersensory redundancy hypothesis* (for a review, see Bahrick et al., 2020; Bahrick & Lickliter, 2012) is consistent with Gibson's invariant detection view of perceptual development (E. J. Gibson, 1969) and developmental systems perspectives (e.g., Gottlieb et al., 1998; Thelen & Smith, 1994). The IRH describes a set of principles characterizing how detection of redundant stimulation across the senses guides and constrains typical perceptual development. Detection of intersensory redundancy (stimulation that is common across two or more sensory systems) creates salience hierarchies that guide selective attention to certain properties of objects and events at the expense of other properties. Specifically, during multisensory stimulation (e.g., the face of voice of a person speaking), attention and perceptual processing is promoted to properties that are redundantly specified across the senses (amodal properties; e.g., synchrony, rhythm, tempo common to the face and voice) at the expense of modality specific properties (e.g., facial configuration; pitch or timbre of voice). This principle is known as *intersensory facilitation*. Detection of intersensory redundancy allows infants to unitize audiovisual stimulation and perceive a single event (e.g., the face and voice of a person speaking). This then provides a foundation for further perceptual processing of the audiovisual event, including affect (e.g., Flom & Bahrick, 2007), prosody (e.g., Bahrick et al., 2019), and communicative intent (e.g., Vaillant-Molina & Bahrick, 2012). In contrast, during unisensory stimulation (e.g., hearing a voice alone or seeing a silent face), attention is promoted to modality specific properties at the expense of amodal properties (e.g., Bahrick et al., 2006, 2013). This principle is known as *unisensory facilitation*. Intersensory and unisensory facilitation are adaptive processes that organize perceptual development by ensuring that a more general framework is perceived prior to more specific detail (e.g., Bahrick, 2001, 2002). Further, they are most pronounced in early development and when tasks are difficult (Bahrick et al., 2010). Later in development and when tasks are easy in relation to the skills of the perceiver, children can detect both amodal and modality-specific properties in multisensory and unisensory stimulation. These principles have been demonstrated in the domains of affect, prosody, temporal synchrony, tempo, and rhythm perception (for reviews, see Bahrick et al., 2020, Bahrick & Lickliter, 2012; Bremner et al., 2012; Lewkowicz, 2000, 2014).

Multisensory Attention Skills in Typical Development

Young TD infants must learn to selectively attend to unified multisensory events (e.g., co-occurring visual and auditory information from a person speaking), while filtering out irrelevant, overlapping sensory information (e.g., sounds from TV). This requires flexibly shifting/disengaging attention from distracting stimulation, maintaining attention to events in the presence of distractors, and perceiving unified information by matching appropriate visual and auditory stimulation. Research has demonstrated that infants' selective attention to multisensory events facilitates processing of information that is common across more than one sense modality (e.g., temporal synchrony, rhythm, tempo, intensity; for reviews see, Bahrick et al., 2020; Bahrick & Lickliter, 2012), and is critical for enabling infants to organize their perceptual world (for reviews, see Bahrick, 2010; Bremner et al., 2012). For example, research using infant looking time methods (e.g., habituation, intermodal preference paradigms) has demonstrated that the presentation of temporally synchronous audible and visual stimulation is highly salient and leads to enhanced attention and perceptual processing of redundantly specified aspects of social events such as prosody and affect (e.g., Bahrick et al., 2019; Flom & Bahrick, 2007), as well as nonsocial events such as object substance (rigid versus elastic) and composition (single versus compound objects; e.g., Bahrick, 1987, 1988). Further, it also facilitates mapping linguistic labels to objects in early infancy (e.g., Gogate & Bahrick, 1998; for a review, see Gogate & Hollich, 2010). However, much of this research is based on averages from group-level data, across a small number of trials, and with typically just one measure of attention (e.g., habituation and/or visual recovery of looking time). In the absence of a fine-grained individual difference measure, relations among skills, links with language and social outcomes, and which skills are most predictive of atypical language and social functioning in children with ASD remain unclear.

Multisensory Attention Skills in Children with ASD

ASD is characterized by atypical social-communicative functioning, including atypical social attention, as well as restricted and repetitive interests and behaviors (American Psychiatric Association, DSM-V, 2013). Though atypical attention and intersensory processing are not considered core symptoms of ASD, a wide array of research findings demonstrates that children with ASD show differences from TD children in the three attention skills we call MASks: (1) disengaging attention from competing stimulation, (2) maintaining attention in the presence of competing stimulation, and (3), intersensory processing of visual and auditory information. Further, atypical patterns of attention are most evident for social events (e.g., social orienting impairments; Dawson et al., 1998, 2004), but less so for nonsocial events (for reviews, see Bahrick & Todd, 2012; Davis & Carter, 2014). Findings are drawn from a diverse and rich body of research. However, the specific nature of attentional strengths and weaknesses in ASD are difficult to evaluate, given that studies have used a wide range of methods, measures, and stimuli, and assessed just one or two MASks in a single study, making findings across studies difficult to compare.

Attention maintenance: Compared to controls, children with ASD show atypical attention to social events, as well as atypical social interactions (for a review, see Davis & Carter, 2014). They show social orienting impairments, including looking less and making fewer attempts to orient to a live experimenter compared to TD controls (e.g., Dawson et al., 1998, 2004; Swettenham et al., 1998). Reduced attention to faces is also evident in high-risk infant siblings of children with ASD compared to low-risk siblings (e.g., Ibanez et al., 2008). Further, compared to TD controls, toddlers with ASD show reduced attention to social events (as indexed by inhibition of eye blinks when viewing social scenes) but not nonsocial events (Shultz et al., 2011). Also, toddlers with ASD show reduced attention to the internal features of faces (e.g., eyes, mouth) and atypical scanning patterns of speaking faces (Chawarska & Shic, 2009; Klin et al., 2002; for a review, see Guillon et al., 2014). In contrast, few studies report differences in attention to nonsocial objects and events between children with ASD and TD children (e.g., Dawson et al., 1998; Shultz et al., 2011; but see Swettenham et al., 1998). Further, individual differences in social attention predict language and social outcomes in ASD, with greater attention to faces predicting better language outcomes (Dawson et al., 2004) and greater attention to the eyes predicting lower Social Affect scores on the Autism Diagnostic Observation Schedule (ADOS: Jones et al., 2008; though see Chawarska et al., 2022).

Shifting/disengaging: Compared to TD children, children with ASD show difficulties shifting attention between objects and events, especially when required to disengage from a competing visual event (for reviews, see Sacrey et al., 2014). For example, they showed significantly slower reaction times (RTs) than TD controls to disengage from a central visual distractor event to shift attention to a lateral event (Landry & Bryson, 2004). However, in this same study, children with ASD showed no differences in shifting attention in the absence of a distractor event compared to controls, indicating difficulties in disengaging attention rather than general difficulties in shifting attention (but see Fischer et al., 2013). Similarly, high-risk infant siblings of children with ASD show atypical patterns of disengagement, particularly those later diagnosed with ASD (Elsabbagh et al., 2013; Zwaigenbaum et al., 2005). Thus. early appearing difficulties in disengaging attention are seen in infants at risk for ASD and may cascade into atypical language and social outcomes (for a review, see Sacrey et al., 2014).

Intersensory matching: Compared to TD controls, children with ASD also show atypical detection of temporal synchrony and intersensory processing of multisensory events, especially for social events (for reviews, see Bahrick & Todd, 2012; Wallace et al., 2020). For example, children with ASD show no evidence of preferential looking towards a sound-synchronous face of a person speaking (e.g., Bebko et al., 2006; Righi et al., 2018), or sound-synchronous biological point light displays (Falck-Ytter et al., 2013; but see Klin et al., 2009), but do show preferential looking towards sound-synchronous nonsocial events (e.g., Bebko et al., 2006; Walker-Andrews et al., 1994). Further, greater preferences to a sound-synchronous face of a person speaking predict greater language skills in ASD (Righi et al., 2018). Children with ASD also show a wider temporal binding window for integrating auditory and visual stimuli, (suggesting more extensive but less temporally precise audiovisual integration than TD children; Foss-Feig et al., 2010; Stevenson et al., 2014), as well as difficulties in performance in audiovisual speech-in-noise tasks (e.g., Foxe et al., 2013; Smith & Bennetto, 2007). Further, children with ASD who show stronger neural (ERP) signals in responding to audiovisual events show lower symptom severity on the ADOS (Brandwein et al., 2015). Finally, children with ASD are also less susceptible to perceiving audiovisual illusions, which depend on audiovisual integration skills (e.g., McGurk effect; Mongillo et al., 2008; Woynaroski et al., 2013).

In summary, a large body of research demonstrates that children with ASD show atypical patterns of shifting and maintaining attention, as well as intersensory processing of sights and sounds, which are particularly evident in the presence of competing stimulation and social events.

Selective Attention in ASD: The Intersensory Processing Disturbance Hypothesis

Children with ASD show atypical selective attention, including slower speed and lower accuracy of target selection in the presence of distractors (Burack, 1994), difficulties in discriminating changes in set size (Remington et al., 2009), and difficulties in ignoring task-irrelevant auditory and visual stimulation (Murphy et al., 2014). Further, as discussed above, they show atypical patterns of unitizing audiovisual stimulation. The *intersensory processing disturbance hypothesis* (IDH; Bahrick & Todd, 2012) was developed as a framework for characterizing how early disturbances in intersensory processing and selective attention to audiovisual events could lead to later atypical language and social functioning. Although atypical MASks are not considered core symptoms/deficits of ASD, intersensory processing and selective attention to audiovisual social events provide critical foundations for language and social functioning (for a review, see Bahrick et al., 2020). According to the IDH, decreased efficiency in processing and attending to redundant audiovisual information may lead to altered unitization of audible and visible stimulation from single events, as well as reduced saliency and difficulties in processing these events. Because of the extraordinary amount of variability and multisensory stimulation provided by social relative to nonsocial events, we would expect difficulties in intersensory processing efficiency to be more evident in the context of social than nonsocial events. Thus, for social events, this decreased intersensory processing efficiency could lead to reduced attention to faces and voices and decreased perception of the intersensory information provided by social events, including affect, prosody, and intensity patterns common across audible and visible speech. Further, even small differences in intersensory processing and selective attention to social events in early development could lead to substantial and increasingly greater differences in social attention, language, and social functioning in later development. Thus, we would predict that in children with ASD, difficulties in intersensory processing should be associated with decreased social attention (i.e., social orienting impairments; Dawson et al., 1998, 2004), atypical language functioning, and increased symptom severity.

The Multisensory Attention Assessment Protocol (MAAP)

Though a sizeable body of research has characterized the role of attention development in children with ASD and TD children, the lack of fine-grained individual difference measures assessing multiple attention skills has made it difficult to characterize developmental trajectories of attention to audiovisual events and which attention skills are most predictive of language and social outcomes. To overcome this obstacle, we developed the first fine-grained measures of individual differences in attention to dynamic, audiovisual events, appropriate for infants and young children, the Intersensory Processing Efficiency Protocol (IPEP; Bahrick et al., 2018a) and the Multisensory Attention Assessment Protocol (MAAP; Bahrick et al., 2018b). The IPEP focuses exclusively on intersensory matching. In contrast, the MAAP assesses all three attention skills—maintenance, shifting/disengaging, and intersensory matching—and is the focus of this manuscript. The MAAP was specifically designed to characterize, within a single protocol, attention skills that prior research has found to best differentiate children with ASD from controls. This includes intersensory matching of sights and sounds from single events, disengagement, and attention maintenance to social (as compared with nonsocial) events, in the presence versus absence of attentional load from competing stimulation (e.g., Landry & Bryson, 2004). Prior research has rarely examined intersensory processing skills (e.g., matching sights and sounds from single events based on detecting common temporal information such as synchrony, rhythm or tempo) along with basic attention skills (maintaining and shifting/disengaging attention). These basic skills have typically been studied in the context of static events or silent dynamic visual events, whereas intersensory processing skills are assessed in the context of dynamic audiovisual events. The MAAP presents multiple, short trials (to provide stable means) and blocks of both social and nonsocial events. It also indexes the cost of competing stimulation from a visual distractor event on each attention skill. The MAAP presents audiovisual events, reflecting the natural, multisensory learning environment of the child. Further, the MAAP requires no verbal responses, and is thus able to provide a common measure for assessing development across infancy and early childhood, the period during which developmental disorders are diagnosed and interventions are most effective.

In our first study using this new protocol with TD children (Bahrick et al., 2018b), we reported a number of noteworthy findings relevant to the current study. Two–5-year-old TD children showed evidence of (1) effects of competing stimulation on MASks (reduced attention maintenance, slower speed of shifting/disengaging, lower intersensory matching) in the presence of a visual distractor compared to its absence, (2) moderate relations among the three MASks, and (3) developmental improvements in MASks, with increasingly longer attention maintenance and greater intersensory matching of social (but not nonsocial) events across age. Finally, (4) structural equation modeling (SEM) revealed that individual differences in MASks for social events predicted language outcomes. Longer attention maintenance to faces predicted greater intersensory matching of faces and voices, which in turn predicted higher receptive and expressive language. In a recent study, we replicated and extended our findings to younger TD infants (Edgar et al., 2022). Intersensory matching of social events on the MAAP at 6 months was a significant predictor of child speech production and vocabulary size at 18, 24, and 36 months, even after controlling for well-established predictors of language, including parent language input and socioeconomic status. Together, these recent findings highlight the foundational role of individual differences in MASks for language and social outcomes.

The Present Study

The present study is organized around three main research questions. First, compared to TD controls, we assessed strengths and weaknesses in MASks in children with ASD and whether weaknesses were more evident in the presence of competing stimulation and for social events. Second, we examined which MASks are most predictive of language and social outcomes in children with ASD and TD children, as well as symptom severity in children with ASD. Finally, we asked whether there are unique pathways from basic MASks to language outcomes for the two groups. Given atypical MASks for children with ASD, we expected a unique pattern of relations among MASks and language outcomes relative to TD children. We addressed these questions using a combination of group differences and individual difference approaches.

Group differences: First, we sought to determine whether the MAAP would provide a sensitive tool for characterizing meaningful differences in MASks between children with ASD and TD children. We administered the MAAP to a group of 2–5-year-old children diagnosed with ASD and a group of TD children matched on nonverbal mental age (MA-matched TDs). We assessed which MASks (attention maintenance, shifting/disengaging, intersensory matching) and which conditions (social versus nonsocial, presence versus absence of competing stimulation) differentiated ASD and TD groups. Based on the literature, we predicted that, compared to TD children, children with ASD would show differences in all three MASks (shorter attention maintenance, slower shifting/disengaging, lower intersensory matching), particularly for social events. Also, based on prior findings using the MAAP with TD children (Bahrick et al., 2018b), we predicted that difficulties in attention in the presence of an irrelevant visual distractor event (providing competing stimulation) would be particularly evident in children with ASD.

Individual differences: We next explored individual differences in MASks and relations between MASks and language and social outcomes for children with ASD and TD children, as well as symptom severity in children with ASD. Specifically, we assessed whether individual differences in MASks predicted cognitive and language functioning, as well as symptom severity. Based on our prior findings (Bahrick et al., 2018b; Edgar et al., 2022), we predicted that better MASks (longer attention maintenance, faster speed of shifting/disengaging, better intersensory matching) particularly for social events, would be associated with better language outcomes for both groups, as well as lower symptom severity for children with ASD.

Pathways to language outcomes: Finally, using structural equation modeling, we tested conceptual models to characterize pathways between MASks for social events and language outcomes for children with ASD and TD children. We predicted that faster speed of shifting/disengaging to faces would predict greater attention maintenance to faces and better intersensory matching of faces and voices, which in turn would predict better language outcomes. We assessed whether pathways were similar or different for children with ASD relative to TD children.

Method

Participants

Twenty-one 2–5-year-old children with a diagnosis of ASD (1 female, 20 males) and 21 TD children (7 females, 14 males) matched on nonverbal mental age (MA-matched TDs) participated.[2] The 21 TD children recruited for this study were also part of a larger study establishing the viability of the MAAP protocol (Bahrick et al., 2018b). Demographic information for the two groups is presented in Table 1.

TD children were matched with children with ASD based on their nonverbal Age Equivalence Score (average of Visual Reception and Fine Motor Age Equivalence Scores) from the Mullen Scales of Early Learning (MSEL; Mullen, 1995), $t(40)=0.20$, $p=0.85$ (for descriptive statistics, see Table 2). Consistent with prior research demonstrating lower verbal than nonverbal functioning in ASD (e.g., Akshoomoff, 2006), children with ASD in our study showed lower Verbal Age Equivalence scores (average of Receptive and Expressive Language scores) than TD children, $t(40)=2.46$, $p=0.02$. Children with ASD were recruited via advertisements at a local Center for Autism and Related Disorders clinic, and TD children were recruited using country birth records and contacted by phone via public phone records. Families received \$10–\$30 in gift cards for participating. The data from one child with ASD was excluded for experimenter error on the MAAP. The data from four TD children were excluded: one for experimenter error on the MAAP, one for fussiness (e.g., crying, eyes closed, and/or squirming or turning away from the screen for multiple trials) during the MAAP, and two for MSEL Early Learning Composite scores greater than two *SD*s below the standardized norm ($M=100$, $SD=15$; $n=2$).

Children with ASD had received a diagnosis of either autism or ASD from a community practitioner or neurologist. Diagnoses were confirmed with the ADOS (Lord et al., 2002). Each ADOS was administered by a certified research-reliable administrator or by an administrator in training who was supervised by a research-reliable administrator. The appropriate ADOS module was administered based on the child's age and language ability (Module 1, No Words: $n=10$; Module 1, Some Words: $n=8$; Module 2, Younger Than 5: $n=2$; Module 2, 5 or Older: $n=1$). All children with a diagnosis of autism or ASD met or exceeded the ADOS-defined cutoff for autism ($n=18$) or ASD (n = 3; using the revised algorithms detailed in Gotham et al., 2007). In addition, an ADOS was administered to all 21 TD children and none scored above the cutoff for ASD (for ADOS descriptive statistics, see Table 2).

Table 1 Demographic information for children with ASD and TD children (Matched on Nonverbal Mental Age)

	Diagnostic group			
	ASD		TD	
	n	Percentage	*n*	Percentage
Sex				
Female	1	4.76%	7	33.33%
Male	20	95.24%	14	66.67%
Ethnicity				
Hispanic	21	100%	16	76.19%
Non-Hispanic	0	0%	5	23.81%
Race				
Asian American	0	0%	1	4.76%
Black and/or African American	0	0%	1	4.76%
White and/or European American	17	81.95%	19	90.48%
Other/Not Reported	4	19.05%	0	0%
Home Language				
Primarily English	6	28.57%	8	38.09%
Primarily Spanish	5	23.81%	4	19.05%
Both English and Spanish	9	42.86%	9	42.86%
Other/Not Reported	1	4.76%	0	0%

Similar to prior research, children with ASD ($M=51.10$ months, $SD=10.31$) were significantly older on average than nonverbal MA-matched TD children ($M=30.18$ months, $SD=7.09$), $t(40)=7.66$, $p<0.001$. Thus, as a secondary aim, we included a smaller group of twelve TD children (4 females, 8 males) who served as a chronological age (CA) matched control group (CA-matched TD children) also drawn from participants in Bahrick et al., (2018b). These CA-matched TD children were included to determine if any differences between children with ASD and TD controls could be attributed to CA and experience rather than ASD diagnosis per se. However, few meaningful differences emerged as results of the CA-matched group mirrored those of the nonverbal MA-matched controls, and

[2] Assuming a β of .80 and a two-tailed *p* value of .05, a sample size of $n=21$ per group (ASD, TD; total $N=42$) has sufficient power to detect effect sizes of 1) $d=.88$ for between-group differences, 2) $d=.76$ for within-group differences (assuming a .30 correlation of scores between groups), and 3) $r=.54$ for correlation analyses.

Table 2 Means and standard deviations (in parentheses) for chronological age, symptom severity measures (ADOS, SCQ, SRS), and cognitive and language measures (MSEL Age Equivalence scores) for children with ASD and TD children (Matched on Nonverbal Mental Age)

Measure	Diagnostic Group		Group differences
	ASD ($n = 21$)	TD ($n = 21$)	Significance
Chronological age (in months)	51.10 (10.30)	30.18 (7.09)	***
ADOS total standardized scores	6.90 (1.48)	1.67 (0.86)	***
SCQ	17.32 (5.61)	5.57 (3.60)	***
SRS T score	79.47 (9.89)	51.45 (9.08)	***
MSEL age equivalence scores (in months)			
Overall	28.46 (14.28)	33.40 (9.16)	*ns*
Nonverbal	32.50 (15.24)	33.26 (9.21)	*ns*
Visual reception	31.33 (15.49)	35.81 (10.16)	*ns*
Fine motor	33.67 (15.80)	30.71 (9.00)	*ns*
Verbal	24.43 (14.44)	33.55 (9.38)	*
Receptive language	27.10 (16.00)	35.43 (9.44)	*
Expressive language	21.76 (12.96)	31.67 (10.30)	**
MSEL early learning composite	62.05 (22.66)	105.95 (11.99)	***

ASD Autism spectrum disorders, *TD* Typically developing, *ADOS* Autism Diagnostic Observation Schedule, *SCQ* Social Communication Questionnaire, *SRS* Social Responsiveness Scale, *MSEL* Mullen Scales of Early Learning. Independent sample *t*-tests: *ns* not significant

*$p < .05$, **$p < .01$, ***$p < .001$

thus, the results of analyses with CA-matched TD children are reported in the Supplemental Material, pp. 1–4.

Assessments

Multisensory Attention Assessment Protocol (MAAP)

The MAAP characterizes individual differences in three foundational MASks—attention maintenance, shifting/disengaging, intersensory matching—to audiovisual social and nonsocial events in the presence and absence of competing stimulation in a single protocol. MASks are assessed across multiple trials to provide a stable mean for assessing relations with developmental outcomes, including language and social functioning, as well as symptom severity. During the MAAP, children were seated in front of a widescreen (102 × 57 cm) monitor. Each trial began with a central distractor event (geometric animation), followed 3 s later by two concurrent 10 s lateral events on the left and right sides. On each trial, one lateral event was synchronous with its natural soundtrack and the other was asynchronous (approximately 3 s out of phase with the soundtrack). There were two types of lateral events, social and nonsocial. Social events consisted of women telling stories with positive, infant-directed speech (social positive) and with neutral affect (social neutral). Nonsocial events consisted of objects impacting a surface in an erratic rhythmic pattern (nonsocial), including wooden blocks and spools and metal nuts and washers (see Fig. 1 for static images of selected examples). Example stimulus videos can be seen on Databrary (https://nyu.databrary.org/volume/326). Three blocks of 8 trials of each event (social positive, social neutral, nonsocial) were presented (24 trials total). Half of the trials in each block (4 trials) were high competition trials, and the remaining trials were low competition trials. On high competition trials, the central event remained on during the audiovisual lateral events, serving as a visual distractor. On low competition trials, the central event was turned off as soon as the lateral events began. Each child received all possible pairs of actresses and objects. Further, we counterbalanced presentation order of event (social neutral, social positive, nonsocial first) and competition type (high or low competition trials first) across participants, resulting in six possible presentation orders.[3] While children viewed the videos, trained observers (previously trained to reach a reliability criterion of 90% or greater with another trained observer) recorded children's look durations to the screen by depressing buttons on a game pad. They were blind to condition and unable to see the videos. Durations of individual looks to the left and right lateral events and latencies (reaction time; RT) to shift attention from the central event to one of the two lateral events were calculated by a custom MATLAB-based computer program.

[3] Preliminary analyses revealed little evidence of effects of presentation order on the three MASks. Thus, we chose not to include presentation order as a factor in our main analyses. For details, see Supplemental Material, p. 4.

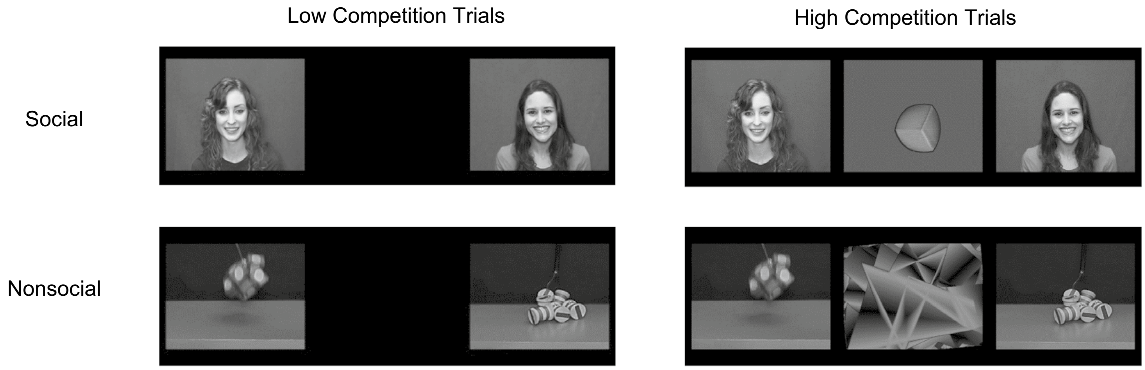

Fig. 1 Static images of the dynamic audiovisual events from the MAAP. On all trials, a three-second central distractor event (computerized geometric shape) was followed by two side-by-side lateral events (social, nonsocial), one of which was synchronous with its appropriate soundtrack. On low competition trials (images on the left), the central event was turned off during the lateral events, whereas on high competition trials (images on the right), the central event remained on during the lateral events

Measures of attention maintenance (duration), shifting/disengaging (speed), and intersensory matching (accuracy) were calculated (see Table 3 for descriptive statistics). They were obtained for each trial and then averaged across trials. Attention maintenance (duration) was indexed by the proportion of available looking time (PALT) to the lateral events and calculated for each trial by dividing the total looking time to both lateral events by the length of the trial (10 s). It reflects overall interest, or the "attention holding" value, of the social and nonsocial events. Shifting/disengaging (speed) was indexed by the latency to shift attention (reaction time, RT) in seconds from the central distractor event to one of the two lateral events (calculated by subtracting the onset time of the lateral event from the onset time of the first look to the lateral event according to the observer's button press). It reflects the "attention getting" value of the social and nonsocial events. Intersensory matching (accuracy) was indexed by the proportion of total looking time (PTLT) to the sound-synchronous lateral event by dividing the looking time to the audiovisual synchronous event by the total looking time to both the synchronous and asynchronous events. PTLTs greater than 50% indicate a preference for the synchronous event (e.g., Bahrick, 1983; Spelke, 1976). We excluded PTLTs greater than 0.90 or less than 0.10 so that our PTLT measure was less influenced by strong side biases. Given there were no significant differences between performance on the social neutral and social positive events for any MAAP measure for either group of children (ASD: $ps > 0.47$, TD: $ps > 0.44$), data were collapsed into an overall social score and contrasted with nonsocial scores for each measure (for details, see Supplemental Material, p. 4).[4] The percentage of trials with useable data was quite high for both groups of children (ASD: $M = 92\%$, $SD = 10\%$; TD: 94%, $SD = 7\%$), indicating that children were highly engaged with the displays. For details on the percentage of useable trials for each MASk, see Supplemental Material, p. 4, Table S4.

Language and Cognitive Outcomes and Indices of Symptom Severity

To assess verbal and nonverbal functioning, the four subscales of the MSEL (Mullen, 1995) were administered to children in English or Spanish, depending on the child's primary home language (per parental report) by a reliable administrator (see Table 2 for descriptive statistics). ADOS raw scores were converted into standardized scores (for details, see Gotham et al., 2009). In addition, parents completed both the Social Communication Questionnaire (SCQ; Rutter et al., 2003), and the Social Responsiveness Scale (SRS; Constantino & Gruber, 2005). Raw scores were calculated for the SCQ, and T scores were calculated for the SRS. Compared to TD children, children with ASD showed higher symptom severity scores on the ADOS, $t(40) = 14.04$, $p < 0.001$, SCQ, $t(38) = 7.96$, $p < 0.001$, and SRS, $t(37) = 9.22$, $p < 0.001$.

[4] Given that social neutral ($n = 8$) and social positive ($n = 8$) trials were collapsed into a single condition, there were a greater number of social ($n = 16$ trials), than nonsocial trials ($n = 8$) overall. It is thus possible that estimates of MASks for social events may have been more stable, and less impacted by outliers, than MASks for nonsocial events. However, analyses revealed little difference in individual variability in MASks for social or nonsocial events. We calculated the coefficient of variation (CV; a scale independent index of variability) for each MASk. On average, for TD children, the CVs for social and nonsocial events were equivalent (social: 32.41, range: 14.29–82.08; nonsocial: 36.44, range: 13.46–108.49) suggesting no difference in MAAP performance on social and nonsocial trials. In contrast, for children with ASD, the average CV for social events was larger than for nonsocial events (social: 43.93, range: 21.57–71.82; nonsocial: 36.39, range: 12.00–97.18).

Table 3 Means and standard deviations (in parentheses for the three multisensory attention skills (duration, speed, accuracy) as a function of type of competition (low, high) and event (social, nonsocial) for children with ASD and TD children (matched on nonverbal mental age)

Measure	Condition	Diagnostic group		Group differences significance
		ASD ($n=21$)	TD ($n=21$)	
Duration: attention maintenance (PALT)	Low competition	.71 (.16)	.80 (.12)	*ns*
	Social	.63 (.21)	.77 (.17)	*
	Nonsocial	.82 (.15)	.84 (.10)	*ns*
	High competition	.54 (.15)	.68 (.15)	*
	Social	.45 (.20)	.68 (.15)	***
	Nonsocial	.69 (.18)	.68 (.21)	*ns*
	Overall	.63 (.14)	.74 (.11)	*
	Social	.55 (.19)	.72 (.14)	**
	Nonsocial	.76 (.13)	.77 (.11)	*ns*
Speed: Attention Shifting (RT)	Low competition	0.75 (0.34)	0.69 (0.19)	*ns*
	Social	0.92 (0.62)	0.72 (0.28)	*ns*
	Nonsocial	0.55 (0.24)	0.65 (0.24)	*ns*
	High competition	1.88 (1.22)	1.12 (0.90)	**
	Social	2.20 (1.58)	1.06 (0.87)	**
	Nonsocial	1.42 (1.38)	1.06 (1.15)	*ns*
	Overall	1.34 (0.75)	0.90 (0.48)	*
	Social	1.65 (1.12)	0.87 (0.49)	**
	Nonsocial	0.92 (0.63)	0.84 (0.56)	*ns*
Accuracy: Intersensory Matching (PTLT)	Low competition	.51 (.07)	.54 (.06)**	*
	Social	.51 (.11)	.56 (.08)**	*ns*
	Nonsocial	.50 (.06)	.52 (.07)	*ns*
	High competition	.50 (.10)	.52 (.06)	*ns*
	Social	.48 (.12)	.53 (.08)	*ns*
	Nonsocial	.52 (.11)	.52 (.08)	*ns*
	Overall	.50 (.07)	.53 (.05)**	*ns*
	Social	.51 (.10)	.55 (.08)*	*
	Nonsocial	.51 (.06)	.52 (.06)	*ns*

ASD autism spectrum disorders, *TD* typically developing, *PALT* proportion of available looking time, *PTLT* proportion of total looking time, *RT* reaction time (in seconds). For PTLT, single sample *t*-tests: *$p<.05$, **$p<.01$. For planned comparisons (ASD versus TD comparisons): *ns*: not significant

*$p<.05$, **$p<.01$, ***$p<.001$

Results

Results are organized around three research questions. *Group differences*: First, which of the three MASks (attention maintenance, shifting/disengaging, intersensory matching) differentiated children with ASD from nonverbal mental age matched (MA-matched) TD children, and were these differences more evident for social than nonsocial events and in the presence versus absence of competing stimulation? To address this research question, we conducted ANOVAs and planned comparisons for each of the three MASks. *Individual differences*: Second, are individual differences in MASks related to language outcomes in children with ASD and MA-matched TD children, as well as symptom severity in children with ASD? For this research question, we conducted bivariate correlations between MASks and measures of language and symptoms severity. *Pathways to language outcomes*: Finally, are there different pathways from basic MASks to language outcomes in children with ASD compared to MA-matched TD children? To address this, we constructed structural equation models (SEMs) to characterize the pathways from MASks for social events to language outcomes in the two groups.

Group Differences: Multisensory Attention Skills in Children with ASD and TD Children

To assess differences in MASks between children with ASD and TD children, a 2 group (ASD, TD; between subjects) × 2 competition (high, low; within subjects) × 2 event

Table 4 Main effects and interaction from the 2 Group (ASD, TD) × 2 type of competition (High, Low) × 3 event (Social, Nonsocial) ANOVA for the three multisensory attention skills (duration, speed, accuracy)

Measure	Source	*F*	*df*	*p*	η^2_p
Duration: Attention Maintenance (PALT)	Group (G)	6.09	1, 40	.02	.13
	Competition (C)	34.75	1, 40	< .001	.47
	Event (E)	26.45	1, 40	< .001	.40
	G × C	.42	1, 40	.52	.01
	G × E	12.85	1, 40	.001	.24
	C × E	.05	1, 40	.83	.00
	G × C x E	3.34	1, 40	.08	.08
Speed: Attention Shifting (RT)	Group (G)	5.69	1, 40	.02	.13
	Competition (C)	20.51	1, 40	< .001	.34
	Event (E)	5.08	1, 40	.03	.11
	G × C	4.76	1, 40	.04	.11
	G × E	3.97	1, 40	.05	.09
	C v E	.57	1, 40	.45	.01
	G × C × E	1.14	1, 40	.29	.03
Accuracy: Intersensory Matching (PTLT)	Group (G)	3.31	1, 40	.08	.08
	Competition (C)	.65	1, 40	.42	.02
	Event (E)	.14	1, 40	.71	.00
	G × C	.23	1, 40	.63	.01
	G × E	1.87	1, 40	.18	.05
	C × E	2.42	1, 40	.13	.06
	G × C × E	.30	1, 40	.59	.01

ASD autism spectrum disorders, *TD* typically developing, *PALT* proportion of available looking time, *PTLT* proportion of total looking time, *RT* reaction time (in seconds)

(social, nonsocial; within subjects) factorial analysis of variance (ANOVA) was conducted for each of the three measures (attention maintenance, shifting/disengaging, intersensory matching).[5] ANOVA main effects and interactions are presented in Table 4. Main effects of group indicated that, collapsed across all trials, children with ASD showed significantly shorter attention maintenance, $F(1, 40) = 6.09$, $p = 0.02$, partial eta squared: $\eta^2_p = 0.13$, and slower shifting/disengaging $F(1, 40) = 5.69$, $p = 0.02$, $\eta^2_p = 0.13$, and marginally lower levels of intersensory matching, $F(1,40) = 3.31$, $p = 0.08$, $\eta^2_p = 0.08$, than TD children (see Table 3). Inspection of η^2_p values revealed that the main effect of group (ASD, TD) accounted for a moderate/medium amount (8% to 13%) of the variance in MASks (e.g., Cohen et al., 2003). Further, to assess evidence of intersensory matching (preference for the synchronous event greater than 50%) within each group, single sample *t*-tests were conducted (see Table 3). TD children showed significant intersensory matching across all trials, $t(20) = 3.06$, $p = 0.01$, whereas children with ASD did not, $t(20) = 0.17$, $p = 0.87$. None of the findings of intersensory matching were qualified by analyses of side biases (preference for left- or right-hand event; see Supplemental Material, p. 5).

Cost of Competing Stimulation on Multisensory Attention Skills

We assessed the cost of competing stimulation (lower performance on high than low competition trials, collapsed across social and nonsocial events) for both groups of children, as well as group differences (ASD versus TD) in the presence versus absence of competing stimulation. Recall that on high competition trials, the central visual event remains on during the lateral events and provides competing stimulation. In contrast, on low competition trials, the central event is turned off at the onset of the lateral events and thus no competing stimulation is present.

Attention maintenance: Both groups of children showed lower attention maintenance to the audiovisual lateral events in the presence than absence of competing stimulation. A main effect of competition indicated shorter attention maintenance on high than on low competition trials, $F(1, 40) = 34.75$, $p < 0.001$, $\eta^2_p = 0.47$ (accounting for a large amount of variance), and planned comparisons revealed

[5] Prior to conducting our main analyses, we conducted preliminary analyses to determine if home language, sex, and ethnicity would be important covariates to include in our analyses. Also, given that MA-matched TD children had higher MSEL Verbal Age Equivalence scores than children with ASD, we also assessed whether Verbal scores should be included as covariates in our analyses. Results of these preliminary analyses indicate that the inclusion of home language, sex, ethnicity, and Verbal Age Equivalence Scores did not qualify our results, and thus they were not included as covariates in the main analyses (for details, see Supplemental Material, pp. 4–5).

this was evident for both children with ASD, $t(40) = 4.62$, $p < 0.001$, and TD children, $t(40) = 3.71$, $p = 0.001$ (see Table 3). Further, planned comparisons of group differences indicated that children with ASD showed significantly shorter attention maintenance than TD children on high competition trials, $t(40) = 2.30$, $p = 0.03$, and marginally shorter attention maintenance on low competition trials, $t(40) = 1.88$, $p = 0.07$ (see Table 3).

Shifting/disengaging: Shifting attention on high competition trials indexes disengagement speed: the time it takes children to disengage from the central event and shift attention to the audiovisual lateral events. In contrast, shifting attention on low competition trials indexes shift speed: the time it takes children to shift attention from the central event to the audiovisual lateral events. It does not reflect disengagement because the central event is turned off as soon as the lateral events begin. Children with ASD, but not TD children, showed slower shifting in the presence than in the absence of the central event. Though a main effect of competition indicated slower speeds to shift attention on high competition (i.e., disengagement) than on low competition (i.e., shift) trials, $F(1,40) = 20.51$, $p < 0.001$, $\eta^2_p = 0.34$, this was qualified by a significant group x competition interaction, $F(1, 40) = 4.76$, $p = 0.04$, $\eta^2_p = 0.11$ (accounting for a moderate amount of variance). Children with ASD showed slower shifting on high than low competition trials, $t(40) = 4.75$, $p < 0.001$, whereas TD children did not, $t(40) = 1.66$, $p = 0.11$ (see Table 3).[6] Further, children with ASD showed slower disengagement than TD children on high competition trials, $t(40) = 2.36$, $p = 0.02$, but no difference in shifting on low competition trials, $t(40) = 0.67$, $p = 0.51$ (see Table 3).

Intersensory matching: There was no significant main effect of competition for intersensory matching, $F(1, 40) = 0.65$, $p = 0.42$, $\eta^2_p = 0.02$ (accounting for a small amount of variance), and no evidence of differences in intersensory matching in the presence versus absence of competing stimulation for either children with ASD, $t(40) = 0.22$, $p = 0.82$, or TD children, $t(40) = 0.89$, $p = 0.37$ (see Table 3). However, children with ASD showed lower levels of intersensory matching than TD children on low competition trials, $t(40) = 2.06$, $p = 0.05$, but not on high competition trials, $t(40) = 1.04$, $p = 0.30$, where both groups showed low levels of intersensory matching. Single-sample t-tests revealed that TD children showed significant intersensory matching on low competition trials, $t(20) = 3.27$, $p = 0.004$, and marginally significant matching on high competition trials, $t(20) = 1.91$, $p = 0.07$. In contrast, children with ASD showed no evidence of intersensory matching on either low, $t(20) = 0.36$, $p = 0.72$, or high competition trials, $t(20) = 0.06$, $p = 0.95$.

In sum, both groups of children showed shorter attention maintenance and slower speeds of shifting/disengaging on high competition trials (distractor event present) than on low competition trials (distractor event absent). Further, children with ASD showed distinctive patterns of attention strengths and weaknesses relative to TD children. Weaknesses were most evident in the context of competing stimulation. On high competition trials, children with ASD showed slower disengagement from the distractor events and shorter attention maintenance to the audiovisual lateral events than TD children. In contrast, strengths were evident on low competition trials. Children with ASD showed no difference in terms of shifting to look to the audiovisual events and only marginally shorter attention maintenance. Finally, children with ASD showed lower levels of intersensory matching with respect to TD children on low, but not on high, competition trials.

Atypical Multisensory Attention Skills for Social but not Nonsocial Events

We assessed differences in attending to social compared to nonsocial events for both groups of children, as well as group differences in attending to social versus nonsocial events.

Attention maintenance: Children with ASD showed shorter attention maintenance to social than nonsocial events relative to TD children. Though a significant main effect of event indicated that children overall showed shorter attention maintenance for social than nonsocial events, $F(1, 40) = 26.45$, $p < 0.001$, $\eta^2_p = 0.40$, this was qualified by a significant group x event interaction, $F(1, 40) = 12.85$, $p = 0.001$, $\eta^2_p = 0.24$ (accounting for a large amount of variance). Planned comparisons revealed that children with ASD showed shorter attention maintenance for social than nonsocial events, $t(40) = 6.11$, $p < 0.001$, whereas TD children did not, $t(40) = 1.09$, $p = 0.28$ (see Table 3). Further, planned comparisons of group differences also indicated that children with ASD showed shorter attention maintenance for social events than TD children, $t(40) = 3.53$, $p = 0.001$, but no difference for nonsocial events, $t(40) = 0.11$, $p = 0.91$ (see Table 3).

Shifting/disengaging: Children with ASD also showed slower shifting/disengaging attention to social than nonsocial events relative to TD children. A significant main effect of event indicated slower shifting/disengaging to social than nonsocial events, $F(1, 40) = 20.51$, $p = 0.03$, $\eta^2_p = 0.11$; however, this was qualified by a significant group x event interaction, $F(1, 40) = 3.97$, $p = 0.05$, $\eta^2_p = 0.09$ (accounting

[6] Significantly slower shift speeds on high than low competition trials were evident for TD children when both MA- and CA-matched TD children were analyzed as a single TD group ($n = 33$), $F(1, 32) = 8.13$, $p = .01$. Thus, the lack of a significant difference for MA-matched TD children was likely due to decreased statistical power due to a small sample size.

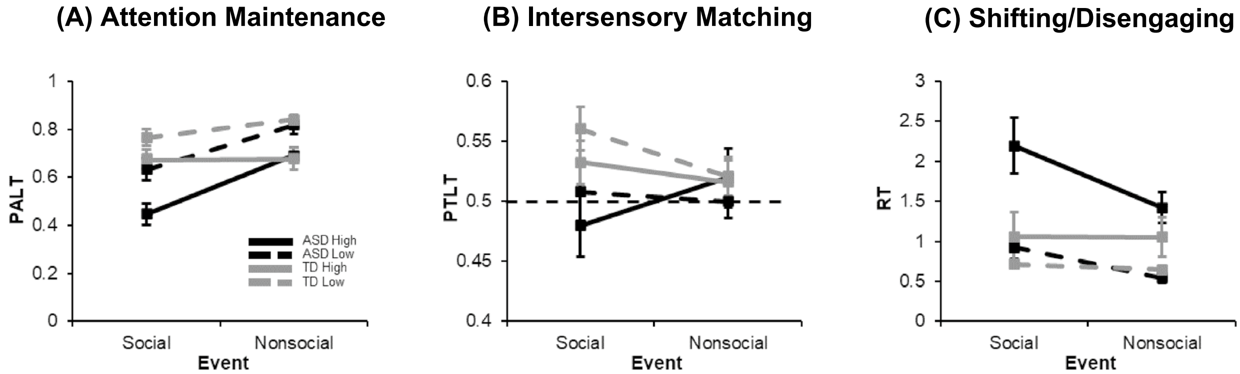

Fig. 2 The three multisensory attention skills—**A** Attention Maintenance (Duration), **B** Intersensory Matching (Accuracy), and **C** Shifting/Disengaging (Speed)—as a Function of Group (ASD, TD), Type of Competition (Low, High) And Event (Social, Nonsocial). *ASD* autism spectrum disorders, *TD* typically developing, *PALT* proportion of available looking time, *PTLT* proportion of total looking time, *RT* reaction time (in seconds). Black solid lines; children with ASD on high competition trials, black dashed lines: children with ASD on low competition trials, grey solid lines: TD children on high competition trials, grey dashed lines: TD children on low competition trials. Error bars are standard errors of the means

for a moderate amount of variance). Children with ASD showed slower shifting/disengaging to social than nonsocial events, $t(40) = 2.99$, $p = 0.01$, whereas TD children did not, $t(40) = 0.04$, $p = 0.85$ (see Table 3). Children with ASD were slower to shift/disengage to social events than TD children, $t(40) = 2.93$, $p = 0.01$, but showed no difference for nonsocial events, $t(40) = 0.66$, $p = 0.51$.

Intersensory matching: There was no significant main effect of event for intersensory matching, $F(1, 40) = 0.14$, $p = 0.71$, $\eta^2_p = 0.00$, and no evidence of differences in intersensory matching for social than nonsocial events for children with ASD, $t(40) = 0.70$, $p = 0.49$, or TD children, $t(40) = 1.22$, $p = 0.22$. However, children with ASD showed lower levels of intersensory matching for social events than TD children, $t(40) = 2.04$, $p = 0.05$, but showed no difference for nonsocial events, $t(40) = 0.45$, $p = 0.67$ (see Table 3). Further, single-sample *t*-tests revealed that TD children showed significant evidence of intersensory matching for social events, $t(20) = 2.69$, $p = 0.01$, whereas children with ASD did not, $t(20) = -0.12$, $p = 0.91$. In contrast, neither group showed evidence of matching for nonsocial events, $ps > 0.16$.[7]

In sum, children with ASD showed patterns of attentional strengths and weaknesses in attending to social compared to nonsocial events relative to TD children. Weaknesses were evident only for social events. Only children with ASD showed shorter attention maintenance and slower shifting/disengaging attention to social than nonsocial events. Further, children with ASD showed shorter attention maintenance, slower speed of shifting, and lower levels of intersensory matching for social events than TD children. In contrast, strengths were evident in the context of nonsocial events as there were no differences in attending to nonsocial events between children with ASD and TD controls.

The Cost of Competing Stimulation on Multisensory Attention Skills for Social and Nonsocial Events

We assessed the cost of competing stimulation for social and for nonsocial events. Planned comparisons revealed that for social and nonsocial events, children with ASD showed shorter attention maintenance, $ts > 2.67$, $ps < 0.01$, and slower shifting/disengaging, $ts > 3.17$, $ps < 0.003$ on high than low competition trials, whereas TD children only showed shorter attention maintenance, $ts > 2.47$, $ps < 0.02$, (but not shifting/disengaging, $ts < 1.48$, $ps > 0.15$) on high than low competition trials (see Table 3 and Fig. 2). Neither group showed differences in intersensory matching for high compared with low competition trials for social or for nonsocial events, $ts < 1.22$, $ps > 0.24$.

We also assessed differences in attending to social compared to nonsocial events for high and for low competition trials. On high and low competition trials, children with ASD showed shorter attention maintenance, $ts > 4.63$, $ps < 0.001$, and longer shifting/disengaging, $ts > 2.38$, $ps < 0.02$, to social than nonsocial events (see Table 3 and Fig. 2). In contrast, TD children showed marginally shorter attention maintenance to social than nonsocial events on low competition trials, $t(40) = 1.88$, $p = 0.07$ (but no difference on high competition trials, $t(40) = 0.04$, $p = 0.97$), and no differences in shifting/disengaging to social versus nonsocial events on low or high competition trials, $ts > 0.57$, $p = 0.57$.

[7] Our prior research using the MAAP demonstrates significant intersensory matching of nonsocial events in both 2–5-year-old TD children, $p = .05$, and 12-month-old TD infants, $p = .01$ (Bahrick et al., 2018a, 2018b; pp. 2213, 2218).

Neither group showed differences in intersensory matching of social versus nonsocial events for high or for low competition trials, $ts > 1.48$, $ps > 0.15$. Finally, for high competition trials, children with ASD showed shorter attention maintenance, $t(40) = 4.16$, $p < 0.001$, and longer disengagement, $t(40) = 2.88$, $p = 0.01$, than TD children for social events (but no difference in intersensory matching, $t(40) = 1.66$, $p = 0.10$). For low competition trials, children with ASD showed shorter attention maintenance, $t(40) = 4.55$, $p = 0.03$, and marginally lower levels of intersensory matching, $t(40) = 1.73$, $p = 0.09$, than TD children for social events (but no difference in shifting, $t(40) = 1.40$, $p = 0.17$). In contrast, there were no group differences for nonsocial events on high, $ts < 0.22$, $ps > 0.83$, or low competition trials, $ts < 1.00$, $ps > 0.33$.

In sum, children with ASD and TD controls showed comparable effects of competing stimulation on MASks on social and nonsocial trials. Both groups of children showed shorter attention maintenance and slower disengaging on high compared to low competition trials for social and nonsocial events. In contrast, differences between children with ASD and TD controls were particularly evident for social events on high competition trials. Compared to TD controls, children with ASD showed shorter attention maintenance, slower shifting/disengaging, and lower levels of intersensory matching for social (but not nonsocial) events on high competition trials. Smaller differences were evident on low competition trials.

Internal Consistency of Multisensory Attention Skills in Children with ASD and TD Children

To assess reliability of our measures in children with ASD and TD children, we also examined the internal consistency of the three MASks. This allowed us to assess whether reliability of MAAP measures in children with ASD was equal to or less reliable than in TD children. For each MASk, we calculated the absolute difference between scores on odd and even trials for each child. Differences between scores on odd and even trials should be close to zero to the extent that measures are relatively free of random error (i.e., reliable). This method is superior to traditional split-half correlations (which are subject to artifacts; Goodwin & Leech, 2006; Jaccard & Becker, 2009). Results indicated small median absolute differences relative to the range of scores for each MASk and little difference between children with ASD and TD children (attention maintenance: ASD = 0.09, TD = 0.08; range: 0.00 to 1.00; shifting/disengaging: ASD = 0.26, TD = 0.24, range: 0 to 10 s; intersensory matching: ASD = 0.04, TD = 0.06; range: 0.00 to 1.00). Thus, MASks show evidence of strong internal consistency for both children with ASD and TD children.

Individual Differences: Multisensory Attention Skills and Relations with Language Outcomes and Symptom Severity

We next assessed relations between MASks and language outcomes for children with ASD and nonverbal mental age matched (MA-matched) TD children, as well as relations between MASks and symptom severity outcomes for just children with ASD. We predicted that better MASks (longer attention maintenance, faster speed of shifting, and better intersensory matching) for social events, particularly in the presence of competing stimulation (high competition trials), would predict better language outcomes (MSEL Receptive and Expressive Language scores) and lower symptom severity scores (ADOS, SCQ, SRS). Also, these correlations served as a basis for the development of the conceptual models used to assess pathways between MASks and language outcomes (see section, Different Pathways from Multisensory Attention Skills to Language Functioning in Children with ASD and TD Children).

Multisensory Attention Skills Predict Language Functioning in Children with ASD and TD Children

We calculated bivariate Pearson correlation coefficients between MASks and MSEL Verbal Age Adjusted Scores (Receptive Language, Expressive Language) for children with ASD and TD children (see Table 5). Correlations between MASks and MSEL nonverbal scores for children with ASD are available in Table S7 (for nonverbal MA-matched TD children, see Table S5). We controlled for family-wise error using a modified Bonferroni procedure (Holm, 1979).[8] Results revealed significant correlations between MAAP performance on trials with social events, particularly in the presence of competing stimulation (high competition trials), and MSEL verbal scores (see Table 5).

For children with ASD, longer attention maintenance and faster speed of disengaging to social events on high competition trials predicted higher Receptive, $rs > 0.52$, $ps < 0.02$, and Expressive Language scores, $rs > 0.50$, $ps < 0.02$ (see Table 5). Also, better intersensory matching of social events on low competition trials predicted higher Receptive, $r = 0.48$, $p = 0.03$, and Expressive Language scores, $r = 0.48$, $p = 0.03$.

[8] We created a family of four for each language outcome (Receptive, Expressive Language) given there are four possible correlations in our 2 event (social, nonsocial) × 2 competition (high, low) design (see Table 5), Thus, the correlation with the smallest p value is compared against a critical value of $p < .0125$ (.05/4). If the correlation with the smallest p value is less than .0125, it is declared significant. Then, the correlation with the next smallest p value is compared against a critical value of $p < .0167$ (.05/3), and so on.

Table 5 For children with ASD, Pearson correlation coefficients between MSEL age equivalent verbal scores (receptive language, expressive language), and each of the three multisensory attention skills (duration, accuracy, speed) as a function of type of competition (low, high) and event (social, nonsocial) for children with ASD and TD children (matched on nonverbal mental age)

Measure	Condition	MSEL age equivalent scores			
		Receptive language		Expressive language	
		ASD	TD	ASD	TD
Duration: Attention Maintenance (PALT)	Low Competition				
	Social	.30	.63**	.30	.46*[f]
	Nonsocial	.04	.17	.16	.14
	High Competition				
	Social	.58**	.59**	.63**	.63**
	Nonsocial	− .09	.32	− .02	.49*
Speed: Attention Shifting (RT)	Low Competition				
	Social	− .35	− .39	− .34	− .10
	Nonsocial	− .25	.14	− .29	.16
	High Competition				
	Social	− .52*	− .10	− .50*	− .06
	Nonsocial	− .06	− .07	− .10	− .12
Accuracy: Intersensory Matching (PTLT)	Low Competition				
	Social	.48*	.74***	.48*	.70***
	Nonsocial	− .24	− .18	− .42	− .21
	High Competition				
	Social	.36	.72***	.46*[f]	.52*[f]
	Nonsocial	− .07	− .07	− .09	− .05

ASD autism spectrum disorders, *TD* typically developing, *MSEL* Mullen Scales of Early Learning, *PALT* proportion of available looking time, *PTLT* proportion of total looking time, *RT* reaction time (in seconds)

*$p<.05$, **$p<.01$ ***$p<.001$.; [f] did not meet significance cutoff when controlling for familywise error

In contrast, there were no significant correlations between MASks for nonsocial events and language outcomes, $ps>0.06$.

For TD children, longer attention maintenance to social events and better intersensory matching of social events on high and low competition trials predicted higher Receptive Language scores, $rs>0.63$ $ps<0.005$ (see Table 5). Further, longer attention maintenance to social events on high competition trials and better intersensory matching for social events on low competition trials predicted higher Expressive Language Scores, $rs>0.63$, $ps<0.002$. In contrast, no significant correlations between shifting/disengaging and language outcomes were evident, and there were no significant correlations between MASks for nonsocial events and language outcomes (except that longer attention maintenance for nonsocial events on high competition trials predicted higher Expressive Language, $r=0.49$, $p=0.02$).

Thus, for both children with ASD and TD children, better MASks (longer attention maintenance, faster shifting/disengaging, better intersensory matching) for social (but not nonsocial) events predicted better language functioning.

Multisensory Attention Skills Predict Symptom Severity in Children with ASD

We calculated bivariate correlations between each MASk and symptom severity scores on the ADOS, SCQ, and SRS (see Table 6). We used total ADOS standard scores (see Table 2; Gotham et al., 2009), as well as social affect (SA) and restrictive and repetitive behavior (RRB) standard scores (Hus et al., 2014). Given the limited range of variability in ADOS standard scores (range: 4–10 in our sample of children with ASD), we calculated Spearman's rank correlations (which apply the Pearson correlation formula to ranked values rather than raw scores; Wilcox, 2017) between MASks and ADOS scores.

ADOS: On high competition trials, longer attention maintenance to social events predicted lower total ADOS scores, Spearman's rho: $r_s=-0.52$, $p=0.02$ (see Table 6). Also, on high competition trials, longer attention maintenance and faster shifting/disengaging to social events predicted lower SA scores, $r_s s>0.49$, $ps<0.03$. Further, on low competition

Table 6 For children with ASD, Spearman's rank correlation coefficients between total and social affect (SA) ADOS standard scores and the three multisensory attention skills (duration, accuracy, speed), as a function of type of competition (low, high) and event (social, nonsocial)

Measure	Condition	ADOS standard scores	
		Total	SA
Duration: attention maintenance (PALT)	Low competition		
	Social	−.36	−.50*
	Nonsocial	−.52*	−.43*
	High Competition		
	Social	−.52*	−.66**
	Nonsocial	−.34	−.38
Speed: attention shifting (RT)	Low Competition		
	Social	.31	.37
	Nonsocial	−.21	−.02
	High Competition		
	Social	.28	.49*
	Nonsocial	.30	.26
Accuracy: intersensory matching (PTLT)	Low Competition		
	Social	−.12	−.15
	Nonsocial	.27	.31
	High Competition		
	Social	−.19	.00
	Nonsocial	−.02	.01

ASD autism spectrum disorders, *ADOS* Autism diagnostic observation schedule, *SA* Social Affect, *PALT* proportion of available looking time, *PTLT* proportion of total looking time, *RT* reaction time (in seconds)

$^*p < .05$, $^{**}p < .01$

trials, longer attention maintenance to social events predicted lower SA scores, $r_s = -0.50$, $p = 0.02$, and longer attention maintenance to nonsocial events predicted lower total ADOS and SA scores, r_s s > 0.43, *p*s < 0.05. In contrast, there were no significant correlations between intersensory matching and total ADOS or SA scores, or between any of the three MASks and RRB scores (for a summary of correlations with RRB scores, see Supplemental Material, Table S7).

SCQ and SRS: We also calculated Pearson correlation coefficients among MASks, SCQ, and SRS scores. Only attention maintenance to social events predicted SRS (but not SCQ) scores. Longer attention maintenance to social events on high (but not low) competition trials predicted lower SRS scores, $r = -0.50$, $p = 0.03$. There were no significant correlations among SRS scores and shifting/disengaging or intersensory matching, or among MASks and SCQ scores (see Supplemental Material, Table S8).

In sum, better performance on two of the three MASks (attention maintenance, shifting) for social events, particularly in the presence of competing stimulation, predicted lower symptom severity on the ADOS, and longer attention maintenance to social events predicted lower SRS scores.

Relations Among Multisensory Attention Skills in Children with ASD and TD Children

We also conducted bivariate correlations to assess relations among the three MASks for children with ASD and non-verbal MA-matched TD children, given prior findings of correlations among MASks (Bahrick et al. 2018b). *Children with ASD*: On high competition trials, faster disengaging to social events was significantly correlated with longer attention maintenance to social events, $r = -0.65$, $p = 0.001$ (see Table 7). No other significant correlations among MASks for social or nonsocial events were evident. *TD children*: On both low and high competition trials, longer attention maintenance to social events was significantly correlated with both intersensory matching, *r*s > 0.45, *p*s < 0.04, and shifting/disengaging for social events, *r*s > 0.43, *p*s < 0.04. Also, on high competition trials, faster disengaging to nonsocial events was significantly correlated with longer attention maintenance, $r = -0.45$, $p = 0.04$. Thus, whereas shifting/disengaging attention was associated with attention maintenance for both groups of children for social events, relations between intersensory matching and attention maintenance

Table 7 Correlations among the three multisensory attention skills (duration, accuracy, speed), as a function of type of competition (low, high) and event (social, nonsocial) for children with ASD and nonverbal MA-matched TD children

Condition/Measure	ASD						TD					
	Low competition			High competition			Low competition			High competition		
	PALT	PTLT	RT	PALT	PTLT	RT	PALT	PTLT	RT	PALT	PTLT	RT
Social												
Duration (PALT)	–	.08	−.32	–	.24	−.65***	–	.45*	−.43*	–	.55**	−.51*
Accuracy (PTLT)		–	−.01		–	−.26		–	−.16		–	−.06
Speed (RT)			–			–			–			–
Nonsocial												
Duration (PALT)	–	.11	−.27	–	.04	−.14	–	−.41	−.33	–	−.12	−.45*
Accuracy (PTLT)		–	.26		–	−.11		–	−.14		–	−.38
Speed (RT)			–			–			–			–

ASD autism spectrum disorder, *TD* typically developing, *PALT* proportion of available looking time, *PTLT* proportion of total looking time, *RT* reaction time

*$p<.05$, *$p<.01$, ***$p<.001$

were only evident for TD children. Taken together with the results of our analyses of relations between MASks and language outcomes, these findings suggest that unique pathways from basic MASks to language outcomes may be evident for children with ASD and TD children. We further explored this possibility with structural equation models.

Unique Pathways from Multisensory Attention Skills to Social Event to Language Functioning in Children with ASD and TD Children

Next, we conducted SEMs to explore unique pathways from MASks for social events to language outcomes in children with ASD and nonverbal mental age matched (MA-matched) TD children. Models were guided by predictions of the IRH and research findings indicating that intersensory processing and selective attention to audiovisual events are foundations for language outcomes (Bahrick et al., 2020; Bahrick & Lickliter, 2012). Based on the patterns of correlations among MASKs and between MASks and language outcomes for the two groups of children, we constructed and tested separate SEMs for children with ASD and TD children. SEMs were conducted using a limited-information estimation (LIE) framework (Bollen, 1996, 2001; Jaccard et al., 2006). Unlike maximum likelihood estimation (MLE) frameworks that are based on asymptotic theory and thus are dependent on large sample sizes, LIE frameworks are appropriate for smaller sample sizes. Also, unlike MLE frameworks in which pathway estimates are derived simultaneously, in LIE frameworks, pathway estimates are derived on an equation-by-equation basis, without calculating indices of model fit. Our outcome measure, Language, was an average of MSEL Age Equivalent Receptive and Expressive Language scores, and MASk measures for social events were averaged across low and high competition trials.

Children with ASD: Pathways from Multisensory Attention Skills to Language Outcomes

For children with ASD, only shifting/disengaging and attention maintenance for social events were significantly correlated (but were not correlated with intersensory matching; see Table 7), but both attention maintenance and intersensory matching for social events predicted language outcomes (see Table 5). Thus, we constructed and tested a model with two pathways from MASKs to language outcomes: (1) Shifting/Disengaging predicted Maintenance, which in turn predicted Language, and (2) Intersensory Matching predicted Language (see Fig. 3A).[9] These two pathways predicted a significant 46% of the variance in language outcomes in children with ASD. Shifting/Disengaging to social events accounted for 41%, $p=0.002$, of the total variance explained in Maintenance to social events. A 1 s decrease in Shifting/Disengaging

[9] Alternative models were tested but ultimately rejected. For children with ASD, a model including Shift/Disengage as a separate predictor of Language, in addition to Maintenance and Intersensory, failed to predict variance in language over and above that predicted by just Maintenance and Intersensory, R^2 change = .06, $p=.17$. Also, for children with ASD we conducted the model depicted in Fig. 3A but using nonsocial trials. Unlike the model with social events, this nonsocial model only predicted a nonsignificant 6% of the variance on Language, $p=.59$, and neither Maintenance nor Intersensory Matching of nonsocial events were significant predictors of Language, $ps>.32$. Thus, none of the alternative models using attention to social or nonsocial events improved our ability to predict language outcomes in children with ASD.

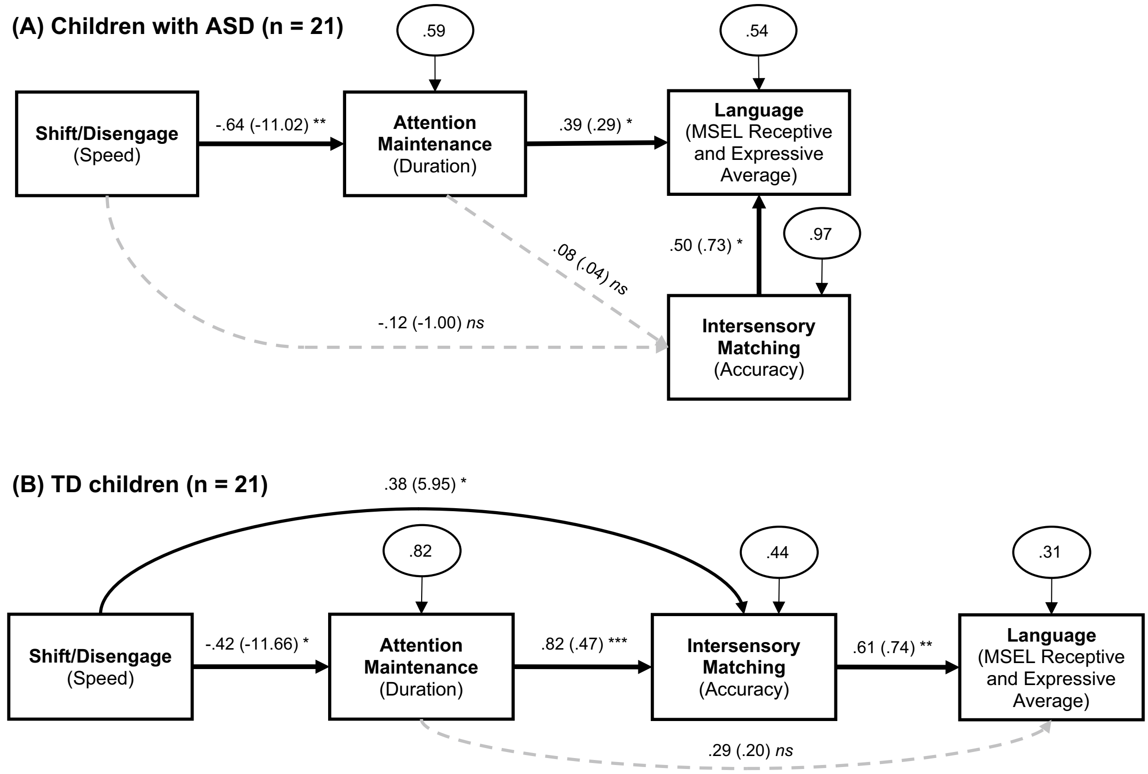

Fig. 3 Structural equation models depicting relations among multisensory attention skills to social events and language outcomes in **A** Children with ASD and **B** TD Children (Matched on Nonverbal Mental Age). *ASD* autism spectrum disorders, *TD* typically developing. Standardized regression coefficients are presented outside the parentheses and unstandardized coefficients are presented inside parentheses. The proportions of variance unaccounted for by predictor variables (error variance) are presented in circles above the outcome variables (Attention Maintenance, Intersensory Matching, Language). Solid black lines depict significant path coefficients, and dashed grey lines depict non-significant path coefficients. *$p<.05$, **$p<.01$, ***$p<.001$

was associated with a 11.02% ($SE=3.00\%$) increase in Maintenance to social events, $p=0.002$. Together, Maintenance and Intersensory Matching accounted for 46%, $p=0.004$, of the total variance explained in Language (Maintenance: R^2 change = 15%, $p=0.04$; Intersensory Matching: R^2 change = 24%, $p=0.01$). A 1% increase in Maintenance for social events predicted a 0.29-month ($SE=0.13$) increase in Language, $p=0.04$ (holding Intersensory Matching constant), and a 1% increase in Intersensory Matching for social events predicted a 0.73-month ($SE=0.26$) increase in Language, $p=0.01$ (holding Maintenance constant). There was no evidence of significant pathways from Shifting/Disengaging or Maintenance to Intersensory Matching of social events (see grey dashed lines in Fig. 3A). Our analyses thus indicate two distinct pathways from MASks for social events to language outcomes in children with ASD: 1) greater attention to faces predicts better language outcomes, and 2) better intersensory matching of faces and voices predicts better language outcomes.

TD Children: Pathways from Multisensory Attention Skills to Language Outcomes

Unlike children with ASD, for TD children, attention maintenance was significantly correlated with shifting/disengaging attention and intersensory matching for social events (see Table 7), but only intersensory matching for social events (not attention maintenance) predicted language outcomes. Thus, for TD children, we tested a model with a single pathway to language outcomes: Shifting/Disengaging predicts Maintenance to social events, which in turn predicted Intersensory Matching, which in turn predicted Language (see Fig. 3B).[10] This pathway predicted

[10] For TD children, we also tested the model depicted in Fig. 3B but using nonsocial events. Results indicated no significant pathways among MASks for nonsocial events. Further, unlike our model with social events, intersensory matching for nonsocial events was not a significant predictor of language, $p=.57$. Thus, the alternative model using attention to nonsocial events did not improve our ability to predict language outcomes in TD children.

a surprisingly large and significant 68% variance in language outcomes in TD children. Shifting/Disengaging to social events accounted for 18%, $p = 0.05$, of the variance in Maintenance for social events, and a 1 s decrease in Shifting/Disengaging was associated with a 11.66% ($SE = 5.79\%$) increase in Maintenance, $p = 0.05$. Together, Maintenance and Shifting/Disengaging for social events accounted for 56%, $p = 0.001$, of the variance in Intersensory Matching for social events. A 1% increase in Maintenance was associated with a 0.47% ($SE = 0.10$) increase in Intersensory Matching, $p = 0.04$ (holding Shifting/Disengaging constant), and a 1 s increase in Shifting/Disengaging was associated with a 5.95% ($SE = 2.71\%$) increase in Intersensory Matching, $p < 0.001$ (holding Maintenance constant). Finally, Intersensory Matching and Maintenance for social events accounted for 68%, $p < 0.001$, of the variance in Language; however, this was primarily due to variance attributable to Intersensory Matching (Intersensory Matching: R^2 change = 0.21, $p = 0.003$; Maintenance: R^2 change = 5%, $p = 0.12$). A 1% increase in Intersensory Matching was associated with a 0.74-month ($SE = 0.21$) increase in Language, $p = 0.003$ (holding Maintenance constant). In contrast, Maintenance was not a significant predictor of Language (holding Intersensory Matching constant; grey dashed line, Fig. 3B).

These novel findings suggest distinct pathways from MASks to language abilities in children with ASD versus TD children. For TD children, findings demonstrate mediational pathways from shifting/disengaging to faces → attention maintenance to faces → intersensory matching of faces and voices → language. In contrast, in children with ASD, intersensory matching of faces and voices does not mediate the relation between attention maintenance to faces and language. Attention to faces and intersensory matching of faces and voices are unrelated and are thus separate predictors of language outcomes in children with ASD. Our finding that longer attention maintenance to faces does not appear to cascade to enhanced intersensory processing of faces and voices and in turn, language, in children with ASD has important implications and deserves further research.

Discussion

Efficient shifting, disengaging, and maintaining attention to social events, as well as detecting audiovisual synchrony across faces and voices, are fundamental skills for typical language, social and cognitive development (for reviews, see Bahrick et al., 2020; Bremner et al., 2012; Colombo, 2001; Ruff & Rothbart, 1996). Children with ASD show altered patterns of selective attention skills. However, developmental pathways from these basic skills to social-communicative functioning remain unclear (Bahrick & Todd, 2012; Mundy & Burnette, 2005). This has been the case in part, because, until recently, there have been no fine-grained individual difference measures of these foundational "multisensory attention skills" (MASks). As a result, which skills become are most affected in atypical development, and which are most predictive of developmental outcomes is not known. Here, we extend our previous findings characterizing MASks in 2–5-year-old TD children using the MAAP, to characterizing MASks in children with ASD. Our findings illustrate important patterns of similarities and differences in MASks between children with ASD and TD children matched on nonverbal mental age (MA-matched TD children). First, they show important group differences in MASks in some areas but not others. Second, our findings reveal that individual differences in MASks in children with ASD predict cognitive and language functioning as well as symptom severity. Finally, these patterns of similarities and differences are elucidated by two new structural equation models depicting unique pathways between basic MASks and language outcomes that differ for children with ASD versus TD controls. These novel findings are interpreted in light of our theoretical frameworks, the *intersensory redundancy hypothesis* (IRH; Bahrick et al., 2020; Bahrick & Lickliter, 2012) and the *intersensory disturbance hypothesis* (IDH; Bahrick & Todd, 2012). A summary of the study's main findings can be found in Table 8.

Group Differences: Multisensory Attention Skills in Children with ASD and TD Children

Our prior findings using the MAAP revealed that 2–5-year-old TD children show (1) atypical MASks (reduced duration of attention, slowed speed of disengaging, low accuracy of intersensory matching) in the presence of an irrelevant visual distractor event, (2) correlations among MASks, (3) improved attention skills with age, and (4) that individual differences in attention to social events predict language outcomes. We replicated many of these findings in a sample of children with ASD, but also found several important differences.

Differences in Attention are Evident for Social Events but not Nonsocial Events

Differences in MASks between children with ASD and TD controls were evident for social (women speaking) but not nonsocial events (objects impacting a surface). Specifically, compared to TD controls, children with ASD looked less at faces, were slower to disengage attention from a distractor event to look to faces, and showed reduced detection of audiovisual synchrony across faces and voices (intersensory matching). In contrast, we found no evidence of these group

Table 8 Summary of study results

Findings	Type of competition		Event type	
	High	Low	Social	Nonsocial
Group differences (ASD, TD)				
Evidence for differences in MASks				
Attention maintenance	Yes	No	Yes	No
Shifting/disengaging	Yes	No	Yes	No
Intersensory matching	No	Yes	Yes	No
Individual Differences in ASD				
Predicting language outcomes				
Attention maintenance	Yes	No	Yes	No
Shifting/disengaging	Yes	No	Yes	No
Intersensory matching	No	Yes	Yes	No
Predicting symptom severity				
Attention maintenance	Yes	Yes	Yes	Yes
Shifting/disengaging	Yes	No	Yes	No
Intersensory matching	No	No	No	No

Unique pathways from social MASks to language	*ASD*	*TD*
Predictors of Language	Attention maintenance Intersensory matching	Intersensory matching
Relations among MASks	Shifting/disengaging predicts attention maintenance	Shifting/disengaging predicts attention maintenance Attention maintenance predicts intersensory matching

differences in MASks for nonsocial events. These findings are consistent with prior studies demonstrating differences in attending to social events in children with ASD as compared with children of typical development (i.e., social orienting impairments; Dawson et al., 1998). For example, children with ASD show difficulties in attention maintenance (e.g., Dawson et al., 1998; Swettenham et al., 1998), speed of disengaging (e.g., Dawson et al., 2004), and intersensory perception (e.g., Bebko et al., 2006; Stevenson et al., 2014) for social events, but little differences in attending to nonsocial events (but see Swettenham et al., 1998). However, some studies have reported that children with ASD show difficulties in intersensory perception for nonsocial events, though these were for simple audiovisual events (e.g., flashes and tones) and tasks that require verbal comprehension or responses (e.g., Foss-Feig et al., 2010). Thus, our results converge with those studies demonstrating patterns of attentional weaknesses for social events, but patterns of attentional strengths for nonsocial events, in children with ASD relative to TD controls.

Differences in Attention are Evident in the Context of Competing Stimulation but not in Its Absence

Children with ASD and TD children showed atypical MASks in the presence (versus absence) of a competing visual distractor event, with shorter attention maintenance and slower speed of disengaging on high than low competition trials. Though children with ASD showed lower overall levels of attention (lower attention maintenance, slower shifting/disengaging across all trials), attentional weaknesses were most evident under conditions of high attention load (trials on which the distractor was present). Further, whereas TD children showed evidence of intersensory matching when the distractor was absent (but not present), children with ASD showed no evidence of intersensory matching in either the presence or absence of the distractor event. Consistent with prior research, our findings demonstrate patterns of strengths in the absence of a distracting event and of weaknesses in the presence of a distracting event in children with ASD (e.g., Landry & Bryson, 2004; but see Fischer et al., 2013).

Attentional Demands of Social Events and Competing Stimulation: Implications for Mechanisms of Atypical Development

For TD children, the distractor event appeared to impact attention and processing of social and nonsocial events in equal manner. However, for children with ASD who have more difficulty attending and processing information provided by social events, the presence of the distractor event appeared to disproportionally impact attention to the social events, leading to the greatest group differences between ASD and TD groups. How should these differences be interpreted?

In general, relative to nonsocial events, social events are more complex, have more variable movement patterns, and provide an extraordinary amount of rapidly changing patterns of intersensory redundancy across face, voice, and gesture (Bahrick & Todd, 2012). Thus, attending and processing information provided by social events should generally be more difficult and challenging for perceivers than attending and processing information provided by nonsocial events. Similarly, the task of shifting and sustaining attention to a target in the presence of competing stimulation creates additional attention load, making the task more difficult and requiring greater attentional control. This is consistent with our finding that the combination of attending to social events, in the context of competing visual stimulation, is particularly difficult for children with ASD. It suggests that attentional difficulties in children with ASD are most evident under conditions of high task difficulty and attentional load. These attentional difficulties are thus likely a reflection of general processing skills rather than specific difficulties in attending to social events. Future studies should assess the role of complexity and variability of social and nonsocial events when characterizing attentional strengths and weaknesses in children with ASD. Though some have argued that the typical salience of social events becomes altered in ASD (e.g., Gergely, 2001; Hoehl et al., 2009), or that social events are less attentionally rewarding (e.g., Dawson et al., 2004), we would argue that the parallel patterns of decreased attention to social events, and to events with competing stimulation, suggest that general processing difficulties cascade disproportionately to social events due to their greater demand on attentional resources. This, in turn, may lead to decreased preferences for social events in ASD.

Individual Differences: Multisensory Attention Skills Predict Language Functioning and Symptom Severity in Children with ASD

Although children with ASD showed altered MASks relative to nonverbal MA-matched TD controls, both groups showed substantial variability in their performance on the MAAP. MASks for social events predicted meaningful variability in language outcomes for both groups of children, as well as symptom severity in children with ASD.

Individual Differences in Multisensory Attention Skills for Social Events Predict Language Functioning

Children with ASD showed relations between MASks and language outcomes that mirrored those of MA-matched TD children, as well as some unique patterns of relations. For children with ASD and TD children, longer attention maintenance and better intersensory matching for social events predicted higher receptive and expressive language scores on the MSEL. In contrast, for just children with ASD, speed of attention disengaging from a central distractor to social events also predicted receptive and expressive language. Further, there was no evidence that MASks for nonsocial events predicted language outcomes in children with ASD, whereas only one MASk (attention maintenance) for nonsocial events predicted expressive language in TD children. These results extend our prior findings with 2–5-year-old TD children by demonstrating meaningful relations between MASks and language outcomes in children with ASD. They highlight the foundational role of MASks in promoting language, social, and cognitive development (Bahrick et al., 2020) and the potential of the MAAP as a novel tool for early detection of children at risk for delayed language functioning.

Individual Differences in Multisensory Attention Skills for Social Events Predict Symptom Severity

Children with ASD also showed meaningful relations between indices of symptom severity and MASks, particularly for social events. Greater attention to faces and faster speed of disengaging to faces predicted lower symptom severity scores on the ADOS and, to a lesser extent, the SRS (but no relations with SCQ scores were evident). Further, relations between MASks and symptoms indexed by ADOS standard scores were particularly evident for social events when a distractor event was present: attention maintenance and speed of disengaging attention to faces in the presence of competing stimulation predicted ADOS total and social affect scores. In contrast, intersensory matching was not a significant predictor of ADOS scores. For nonsocial events, few relations between MASks and ADOS scores were evident. These findings suggest that atypical patterns in some of the MASks, particularly attention maintenance and disengaging attention to social events may underlie atypical social-communicative functioning in ASD.

Unique Pathways Between Multisensory Attention Skills and Language Outcomes for Children with ASD and TD Children

Findings from SEM analyses revealed unique pathways from basic MASks for social events to language outcomes. We found both striking similarities and differences between models for children with ASD and TD children. For TD children, a four-part mediational causal chain was evident: faster shifting/disengaging to faces predicted longer looking at faces, which in turn predicted better matching of faces and voices, and in turn language outcomes. This mediational chain predicted a surprisingly high 68% of the variance in language outcomes (for similar findings, see Bahrick et al., 2018b). Intersensory matching of face-voice relations was the only significant predictor of language outcomes in TD children, and it mediated the relation between more basic MASks (shifting/disengaging and attention maintenance) and language outcomes. In contrast, for children with ASD, this four-part causal chain was interrupted and two separate pathways to language outcomes emerged: (1) faster shifting/disengaging to faces predicted longer looking at faces, and in turn better language outcomes, and (2) better matching of face-voice relations (which was unrelated to shifting/disengaging and looking to faces) predicted better language outcomes. Attention to faces and face-voice matching together predicted 46% of the variance in language outcomes. Thus, the link between looking to faces and matching of face-voice relations found in TD controls was not evident in children with ASD. Instead, for children with ASD, maintaining attention to faces and matching faces with voices based on common temporal patterns are separate predictors of language outcomes. If replicated with larger samples, these findings may provide important insights into the relations among social attention, intersensory processing of social events, and language functioning in early development. They also have important implications for interventions to improve language.

Our models were guided by and are consistent with the predictions of IRH that intersensory processing and selective attention to certain properties of audiovisual events are a foundation for language outcomes (Bahrick et al., 2020; Bahrick & Lickliter, 2012). According to the IRH, intersensory processing and efficient unitization of auditory and visual stimulation should improve across age as a result of experience with audiovisual events, particularly experience with faces and voices. Research indicates that in TD children, developmental improvements in intersensory processing go hand in hand with enhanced attention to faces and voices, and together increase language learning opportunities provided by caregivers (Edgar et al., 2022). Thus, the co-development of social attention and intersensory processing of faces and voices may result from a selective attention "feedback loop" (for a similar perspective on child vocalizations and adult language input, see Warlaumont et al., 2014). Children are more likely to attend to the faces of people speaking if they are efficiently unitizing auditory and visual speech information. In turn, they should have more opportunities for processing the auditory and visual stimulation provided by social events the more they attend to the faces of people speaking. This selective attention feedback loop provides children with a powerful opportunity to take advantage of language learning opportunities that occur in the context of parent–child interactions. Improvements in social attention and intersensory processing of faces and voices should lead to increased engagement in behaviors that are well-established predictors of language, including engagement in joint attention (e.g., Mundy & Burnette, 2005) and word mapping (e.g., Gogate & Hollich, 2010).

However, these critical developmental linkages between intersensory processing of auditory and visual stimulation, social attention, and language development appear to be substantially altered in children with ASD. According to the predictions of the IDH, small disturbances in unitizing faces and voices may lead to the reduced salience and processing of faces and voices, which in turn may cascade into worsening language outcomes. Thus, decreased efficiency in the unitization of faces and voices may attenuate the selective attention feedback loop, such that attention to faces and face-voice matching may develop relatively independently of one another in children with ASD. Consistent with the predictions of the IDH, findings from our model demonstrate that attention to faces and face-voice matching were unrelated to one another yet were separate predictors of language outcomes. This attenuated feedback loop may also result in different developmental trajectories for attention to faces and matching faces and voices.

Given the lack of connection we found between attention to faces and face-voice matching in ASD, the development of these attention skills may follow relatively independent developmental trajectories such that developmental improvements in one behavior (e.g., attention to faces) may not accompany improvements in the other (e.g., face-voice matching) yet each may still facilitate language learning. For example, enhanced attention to faces in the context of parent–child interactions, regardless of the ability to match faces and voices, may lead to increased parent language input (a well-established predictor of child language outcomes; (e.g., Hart & Risley, 1995). Further, efficient matching of faces and voices, regardless of attention to faces, may also lead to language learning opportunities, even with reduced parent language input. For example, temporally synchronous auditory and visual stimulation (labeling and gesture) may facilitate object-label mapping (e.g., Gogate & Hollich, 2010). Thus, both the attenuated feedback loop and the possibility of relatively independent developmental trajectories for attention to faces and face-voice matching may in part explain why enhanced attention to faces does not lead to greater face-voice matching in children with ASD.

Another reason why enhanced attention to faces may not result in greater face-voice matching in ASD may be gleaned from recent eye-tracking studies in children with ASD. They suggest that children with ASD and TD children may show selective attention to different parts of the face. Some studies indicate that TD children show greater attention to the mouth as compared with the eyes at the ages during which they are first learning a language (e.g., Lewkowicz & Hansen-Tift, 2012; Tenenbaum et al., 2015). It is thought that greater attention to the mouth reflects a greater preference for face-voice synchrony, although this has not yet been directly tested. In contrast, some studies show that children with ASD show preferences for the eyes over the mouth of a speaking face (e.g., Campbell et al., 2014; Chawarska et al., 2012). However, other studies have found contradictory results, with preferences for the mouth over the eyes in ASD (e.g., Jones et al., 2008; Jones & Klin, 2013; for a reviews, see Klin et al., 2015). Thus, future research is needed to characterize the conditions under which attention is directed to the mouth versus the eyes, and relations with language outcomes in children with ASD (e.g., Chawarska et al., 2022), as well as to test the mechanisms underlying these different attention patterns.

In sum, findings from our models, guided by and consistent with predictions of the IRH and IDH, provide novel insights into the developmental pathways from basic attention and processing of audiovisual events to language outcomes in children with ASD and TD children. They highlight the importance of the developmental links between social attention and intersensory processing in typical development, as well as the consequences of a breakdown between these behaviors in children with ASD. Characterizing the developmental trajectories and relations between social attention and intersensory processing using longitudinal designs is an important goal for future research.

Multisensory Attention Skills in Children with ASD

We developed the term “multisensory attention skills” (MASks) to examine under a single conceptual framework, three attention skills typically studied separately (duration of attention, speed of shifting/disengaging, accuracy of intersensory processing; Bahrick et al., 2020). We consider these basic MASks to be foundational for more complex language, social, and cognitive outcomes (Bahrick et al., 2018a, 2018b, 2020). Further, children with ASD show patterns of attentional strengths and weaknesses in these skills. The MAAP can, for the first time, characterize variability in these MASks at a fine enough level of detail so that developmental trajectories of attention and links with developmental outcomes can be meaningfully assessed. Further, unlike most methods, it provides a common measure that can be used across infancy and early childhood. This allows researchers to assess individual differences and developmental trajectories across a transition period when developmental disorders are often diagnosed, and interventions are most effective.

Limitations and Future Directions

Some future directions and limitations should be noted. First, we did not include a control group of children with general developmental delays (e.g., Downs Syndrome) matched with children with ASD on both mental and chronological age. Comparisons between these two groups of children would allow researchers to characterize which attentional strengths and weaknesses are specific to ASD and which are due to general developmental delays. Second, though our sample size was sufficient for identifying medium or larger effects, we were likely underpowered for identifying smaller effects. Thus, in important goal for future research will be to replicate these findings (particularly those from our SEM findings) with sufficient sample sizes for identifying smaller effects. Third, in the present study, there were more social than nonsocial trials (16 versus 8 trials). One might thus speculate that estimates of MASks on social trials were more stable and less impacted by outliers than on nonsocial trials. However, analyses indicate equivalent amounts of individual variability on social and nonsocial trials for children with ASD and TD children, suggesting no difference in stability of the measures. An updated version of the MAAP provides an equal number of social and nonsocial trials (12 each; see Edgar et al., 2022). Fourth, there were too few females with ASD enrolled into the study to assess potential sex differences. However, our prior research with TD infants and children found no evidence of sex differences on MASks or relations between MASks and language outcomes (Bahrick et al., 2018b; Edgar et al., 2022). Fifth, we were unable to derive a model of pathways between MASks and atypical social functioning/symptom severity in children with ASD. A more fine-grained measure of intersensory processing, such as the IPEP (Bahrick et al., 2018a), may be capable of revealing relations between intersensory processing of faces and voices and symptom severity in children with ASD. Sixth, future research should also assess whether the MAAP may be useful for detecting atypical attention patterns in infants and children at-risk for later developmental problems. Given that the MAAP can be administered to infants as young as 3 months of age, the MAAP shows promise as a clinical tool for the detection of atypical patterns of attention indicating risk for later atypical language functioning.

This would enable interventions to be administered while these patterns of attention are still developing and before attention disturbances cascade into worsening language and social functioning.

Finally, despite a greater understanding of the importance for MASks in early development, developmental trajectories of MASks and relations with cognitive, language, and social outcomes in remain relatively unexplored. Developmental norms characterizing the early development of MASks in TD children are needed as a foundation for identifying the atypical development of these skills. To address this need, we are building a large database to characterize trajectories of the typical development of MASks from 3 through 60 months of age, along with relations with cognitive, language, and social outcomes. This can then serve as a baseline for identifying atypical patterns of attention.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10803-022-05752-3.

Acknowledgments We would like to acknowledge the following for their assistance in data collection: Melissa A. Argumosa, Irina Castellanos, Rosalynn Garcia, Lissette Robles, Christoph H. Ronacher, Barbara M. Sorondo, Mariana Vaillant-Molina, and Janet Vasquez. We would also like to thank the Center for Autism and Related Disorders in Miami, FL, for their assistance in participant recruitment.

Author Contributions JTT and LEB contributed to the study conception and design, material preparation, data collection and analysis. Funding was obtained by LEB and the first draft of the manuscript was written by JTT. JTT and LEB commented on previous versions of the manuscript, as well as read and approved the final manuscript.

Funding This research was supported by National Institute of Child Health and Human Development grants K02 HD064943 and RO1 HD053776, a grant from Autism Speaks 1906, and a grant from the Marino Autism Research Institute, the University of Miami, awarded to Lorraine E. Bahrick.

Declarations

Conflict of interest The authors have no conflicts of interest or competing interests to declare that are relevant to the content of this article.

References

Akshoomoff, N. (2006). Use of the Mullen scales of early learning for the assessment of young children with autism spectrum disorders. *Child Neuropsychology, 12*(4–5), 269–277.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Bahrick, L. E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior & Development, 6*, 429–451. https://doi.org/10.1016/S0163-6383(83)90241-2

Bahrick, L. E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development, 10*(4), 387–416. https://doi.org/10.1016/0163-6383(87)90039-7

Bahrick, L. E. (1988). Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development, 59*(1), 197–209. https://doi.org/10.2307/1130402

Bahrick, L. E. (2001). Increasing specificity in perceptual development: Infants' detection of nested levels of multimodal stimulation. *Journal of Experimental Child Psychology, 79*, 253–270. https://doi.org/10.1006/jecp.2000.2588

Bahrick, L. E. (2002). Generalization of learning in three-and-a-half-month-old infants on the basis of amodal relations. *Child Development, 73*, 667–681. https://doi.org/10.1111/1467-8624.00431

Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. In J. G. Bremner & T. D. Wachs (Eds.), *The Wiley-Blackwell handbook of infant development: Vol 1. Basic research* (2nd ed., pp. 120–165). Wiley-Blackwell. https://doi.org/10.1002/9781444327564.ch4

Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 183–206). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586059.003.0008

Bahrick, L. E., Lickliter, R., & Castellanos, I. (2013). The development of face perception in infancy: Intersensory interference and unimodal visual facilitation. *Developmental Psychology, 49*, 1919–1930. https://doi.org/10.1037/a0031238

Bahrick, L. E., Lickliter, R., & Todd, J. T. (2020). The development of multisensory attention skills: Individual differences, developmental outcomes, and applications. In J. J. Lockman & C. S. Tamis-LeMonda (Eds.), *The Cambridge handbook of infant development* (pp. 303–338). Cambridge University Press. https://doi.org/10.1017/9781108351959.011

Bahrick, L. E., Lickliter, R., Castellanos, I., & Vaillant-Molina, M. (2010). Increasing task difficulty enhances effects of intersensory redundancy: Testing a new prediction of the intersensory redundancy hypothesis. *Developmental Science, 13*, 731–737. https://doi.org/10.1111/j.1467-7687.2009.00928.x

Bahrick, L. E., Lickliter, R., & Flom, R. (2006). Up versus down: The role of intersensory redundancy in the development of infants' sensitivity to the orientation of moving objects. *Infancy, 9*, 73–96. https://doi.org/10.1207/s15327078in0901_4

Bahrick, L. E., McNew, M. E., Pruden, S. M., & Castellanos, I. (2019). Intersensory redundancy promotes infant detection of prosody in infant-directed speech. *Journal of Experimental Child Psychology, 183*, 295–309. https://doi.org/10.1016/j.jecp.2019.02.008

Bahrick, L. E., Soska, K. C., & Todd, J. T. (2018a). Assessing individual differences in the speed and accuracy of intersensory processing in young children: The intersensory processing efficiency protocol. *Developmental Psychology, 54*(12), 2226–2239. https://doi.org/10.1037/dev0000575

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. 1453–1508). MIT Press.

Bahrick, L. E., Todd, J. T., & Soska, K. C. (2018b). 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. https://doi.org/10.1037/dev0000594

Bebko, J. M., Weiss, J. A., Demark, J. L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *Journal of Child Psychology and Psychiatry, 47*(1), 88–98. https://doi.org/10.1111/j.1469-7610.2005.01443.x

Bollen, K. A. (1996). A limited information estimator for LISREL models with and without heteroscedasticity. In G. A. Marcoulides (Ed.), *Advanced structural equation modeling: Issues and techniques* (pp. 227–241). Lawrence Erlbaum.

Bollen, K. A. (2001). Two-stage least squares and latent variable models: Simultaneous estimation and robustness to misspecifications. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 119–138). Scientific Software International.

Brandwein, A. B., Foxe, J. J., Butler, J. S., Frey, H. P., Bates, J. C., Shulman, L. H., & Molholm, S. (2015). Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *Journal of Autism and Developmental Disorders, 45*(1), 230–344. https://doi.org/10.1007/s10803-014-2212-9

Bremner, A. J., Lewkowicz, D. J., & Spence, C. (2012). *Multisensory development*. Oxford University Press.

Burack, J. A. (1994). Selective attention deficits in persons with autism: Preliminary evidence of an inefficient attentional lens. *Journal of Abnormal Psychology, 103*(3), 535–543.

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

Chawarska, K., Lewkowicz, D., Feiner, H., Macari, S., & Vernetti, A. (2022). Attention to audiovisual speech does not facilitate language acquisition in infants with familial history of autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.13595

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

Chawarska, K., & Shic, F. (2009). Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 39*(12), 1663–1672. https://doi.org/10.1007/s10803-009-0803-7

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203774441

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 51*(1), 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Constantino, J. N., & Gruber, C. P. (2005). *The social responsiveness scale*. Western Psychological Services.

Davis, N. O., & Carter, A. S. (2014). Social development in autism. In F. R. V. Volkmar, R. Paul, S. J. Rogers, & K. A. Pelphrey (Eds.), *Handbook of autism and pervasive developmental disorders, vol 1: Diagnosis, development, and brain mechanisms* (4th ed., pp. 311–338). Wiley. https://doi.org/10.1002/9781118911389

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders, 28*(6), 479–485. https://doi.org/10.1023/A:1026043926488

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*, 271–283. https://doi.org/10.1037/0012-1649.40.2.271

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*. https://doi.org/10.1037/dev0001375

Elsabbagh, M., Fernandes, J., Jane Webb, S., Dawson, G., Charman, T., & Johnson, M. H. (2013). Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. *Biological Psychiatry, 74*(3), 189–194. https://doi.org/10.1016/j.biopsych.2012.11.030

Falck-Ytter, T., Rehnberg, E., & Bölte, S. (2013). Lack of visual orienting to biological motion and audiovisual synchrony in 3-year-olds with autism. *PLoS ONE, 8*(7), e68816. https://doi.org/10.1371/journal.pone.0068816

Fischer, J., Koldewyn, K., Jiang, Y. V., & Kanwisher, N. (2013). Unimpaired attentional disengagement and social orienting in children with autism. *Clinical Psychological Science, 2*(2), 214–223. https://doi.org/10.1177/2167702613496242

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology, 43*(1), 238–252. https://doi.org/10.1037/0012-1649.43.1.238

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window for autism spectrum disorders. *Experimental Brain Research, 203*(2), 381–389. https://doi.org/10.1007/s00221-010-2240-4

Foxe, J. J., Molholm, S., Del Bene, V. A., Frey, H.-P., Russo, N. N., Blanco, D., Saint-Amour, D., & Ross, L. A. (2013). Severe multisensory speech integration deficits in high-functioning school-aged children with autism spectrum disorder (ASD) and their resolution during early adolescence. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bht213

Gergely, G. (2001). The obscure object of desire: "Nearly, but clearly not, like me'': Contingency preference in normal children versus children with autism". *Bulletin of the Menninger Clinic, 65*, 411–426.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

Gogate, L. J., & Bahrick, L. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology, 69*(2), 133–149. https://doi.org/10.1006/jecp.1998.2438

Gogate, L. J., & Hollich, G. (2010). Invariance detection within an interactive system: A perceptual gateway to language development. *Psychological Review, 117*(2), 496–516. https://doi.org/10.1037/a0019049

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

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37*(4), 613–627.

Gottlieb, G., Wahlsten, D., & Lickliter, R. (1998). The significance of biology for human development: A developmental psychobiological systems view. In R. Lerner (Ed.), *Handbook of child psychology, volume 1: Theoretical models of human development* (pp. 233–274). John Wiley.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews, 42*, 279–297. https://doi.org/10.1016/j.neubiorev.2014.03.013

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Brookes Publishing.

Hoehl, S., Reid, V. M., Parise, E., Handl, A., Palumbo, L., & Striano, T. (2009). Looking at eye gaze processing and its neural correlates in infancy—Implications for social development and autism spectrum disorder. *Child Development, 80*, 968–985.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

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

Ibanez, L. V., Messinger, D. S., Newell, L., Lambert, B., & Sheskin, M. (2008). Visual disengagement in the infant siblings of children with an autism spectrum disorder (ASD). *Autism, 12*(5), 473–485. https://doi.org/10.1177/1362361308094504

Jaccard, J. J., & Becker, M. A. (2009). *Statistics for the behavioral sciences* (5th ed.). Wadsworth Publishing.

Jaccard, J. J., Guilamo-Ramos, V., Johansson, M., & Bouris, A. (2006). Multiple regression analyses in clinical child and adolescent psychology. *Journal of Clinical Child and Adolescent Psychology, 35*, 456–479. https://doi.org/10.1207/s15374424jccp3503_11

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry, 65*(8), 946–954. https://doi.org/10.1001/archpsyc.65.8.946

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature, 504*(7480), 427–431. https://doi.org/10.1038/nature12715

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809–816.

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature, 459*(7244), 257–261.

Klin, A., Shultz, S., & Jones, W. (2015). Social visual engagement in infants and toddlers with autism: Early developmental transitions and a model of pathogenesis. *Neuroscience and biobehavioral reviews* (Vol. 50, pp. 189–203). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2014.10.006

Landry, R., & Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry, 45*(6), 1115–1122. https://doi.org/10.1111/j.1469-7610.2004.00304.x

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126*, 281–308. https://doi.org/10.1037//0033-2909.126.2.281

Lewkowicz, D. J. (2014). Early experience and multisensory perceptual narrowing. *Developmental Psychobiology, 56*, 292–315. https://doi.org/10.1002/dev.21197

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences, 109*(5), 1431–1436. https://doi.org/10.1073/pnas.1114783109

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (2002). *Autism diagnostic observation schedule: Manual*. Western Psychological Services.

Mongillo, E. A., Irwin, J. R., Whalen, D. H., Klaiman, C., Carter, A. S., & Schultz, R. T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(7), 1349–1358. https://doi.org/10.1007/s10803-007-0521-y

Mullen, E. M. (1995). *Mullen scales of early learning* (AGS). American Guidance Service.

Mundy, P., & Burnette, C. (2005). Joint attention and neurodevelopmental models of autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders, Vol. 1: Diagnosis, development, neurobiology, and behavior* (3rd ed., pp. 650–681). Wiley. https://doi.org/10.1002/9780470939345.ch25

Murphy, J. W., Foxe, J. J., Peters, J. B., & Molholm, S. (2014). Susceptibility to distraction in autism spectrum disorder: Probing the integrity of oscillatory Alpha-Band Suppression Mechanisms. *Autism Research, 7*(4), 442–458. https://doi.org/10.1002/aur.1374

Remington, A., Swettenham, J., Campbell, R., & Coleman, M. (2009). Selective attention and perceptual load in autism spectrum disorder. *Psychological Science, 20*, 1388–1393.

Righi, G., Tenenbaum, E. J., McCormick, C., Blossom, M., Amso, D., & Sheinkopf, S. J. (2018). Sensitivity to audio-visual synchrony and its relation to language abilities in children with and without ASD. *Autism Research, 11*(4), 645–653. https://doi.org/10.1002/aur.1918

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: Themes and variations*. Oxford University Press.

Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire*. Western Psychological Services.

Sacrey, L.-A.R., Armstrong, V. L., Bryson, S. E., & Zwaigenbaum, L. (2014). Impairments to visual disengagement in autism spectrum disorder : A review of experimental studies from infancy to adulthood. *Neuroscience and Biobehavioral Reviews, 47*, 559–577. https://doi.org/10.1016/j.neubiorev.2014.10.011

Shultz, S., Klin, A., & Jones, W. (2011). Inhibition of eye blinking reveals subjective perceptions of stimulus salience. *Proceedings of the National Academy of Sciences of the United States of America, 108*(52), 21270–21275. https://doi.org/10.1073/pnas.1109304108

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry, 48*(8), 813–821. https://doi.org/10.1111/j.1469-7610.2007.01766.x

Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8*, 553–560. https://doi.org/10.1016/0010-0285(76)90018-9

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience, 34*, 691–697. https://doi.org/10.1523/JNEUROSCI.3615-13.2014

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., Rees, L., & Wheelwright, S. (1998). The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants. *Journal of Child Psychology and Psychiatry, 39*(5), 747–753. https://doi.org/10.1111/1469-7610.00373

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42*(6), 1173–1190. https://doi.org/10.1017/S0305000914000725

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT Press.

Vaillant-Molina, M., & Bahrick, L. E. (2012). The role of intersensory redundancy in the emergence of social referencing in 5½-month-old infants. *Developmental Psychology, 48*, 1–9. https://doi.org/10.1037/a0025263

Walker-Andrews, A. S., Haviland, J. M., Huffman, L., & Toci, L. (1994). Brief report: Preferential looking in intermodal perception by children with autism. *Journal of Autism and Developmental Disorders, 24*(1), 99–107. https://doi.org/10.1007/BF02172216

Wallace, M. T., Woynaroski, T. G., & Stevenson, R. A. (2020). Multisensory integration as a window into orderly and disrupted cognition and communication. *Annual Review of Psychology, 71*, 193–219. https://doi.org/10.1146/annurev-psych-010419-051112

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science, 25*, 1314–1324. https://doi.org/10.1177/0956797614531023

Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing* (4th ed.). Academic Press.

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

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23*(2–3), 143–152. https://doi.org/10.1016/j.ijdevneu.2004.05.001