The Multisensory Attention Assessment Protocol (MAAP): Characterizing Individual Differences in Multisensory Attention Skills in Infants and Children and Relations with Language and Cognition

Lorraine E. Bahrick, James Torrence Todd, and Kasey C. Soska

Florida International University

Author Note

Lorraine E. Bahrick, Department of Psychology, Florida International University, Miami, FL; James Torrence Todd, Department of Psychology, Florida International University, Miami, FL; Kasey C. Soska, Department of Psychology, Florida International University, Miami, FL.
Portions of these data were presented at the International Congress on Infant Studies, 2016; the International Meeting for Autism Research, 2009, 2010, 2011, 2012, 2013, and 2014; the International Society of Developmental Psychobiology, 2016; and the Society for Research in Child Development, 2011 and 2017.

This research was supported by NICHD grants K02 HD064943 and RO1 HD053776, a grant from Autism Speaks #1906, from the Marino Autism Research Institute, the University of Miami, awarded to the first author.

Correspondence concerning this article should be addressed to Lorraine E. Bahrick, Department of Psychology, Florida International University, bahrick@fiu.edu.

Abstract

Multisensory attention skills provide a crucial foundation for early cognitive, social, and language development, yet there are no fine-grained, individual difference measures of these skills appropriate for preverbal children. The Multisensory Attention Assessment Protocol (MAAP) fills this need. In a single video-based protocol requiring no language skills, the MAAP assesses individual differences in three fundamental building blocks of attention to multisensory events—the duration of attention maintenance, the accuracy of intersensory (audiovisual) matching, and the speed of shifting—for both social and nonsocial events, in the context of high and low competing visual stimulation. In Experiment 1, 2- to 5-year-old children (*N* = 36) received the MAAP and assessments of language and cognitive functioning. In Experiment 2 the procedure was streamlined and presented to 12-month-olds (*N* = 48). Both infants and children showed high levels of attention maintenance to social and nonsocial events, impaired attention maintenance and speed of shifting when competing stimulation was high, and significant intersensory matching. Children showed longer maintenance, faster shifting, and less impairment from competing stimulation than infants. In 2- to 5-year-old children, duration and accuracy were inter-correlated, showed increases with age, and predicted cognitive and language functioning. The MAAP opens the door to assessing developmental pathways between early attention patterns to audiovisual events and language, cognitive, and social development.

Keywords: individual difference measure, multisensory attention development, social and nonsocial events, intersensory processing, disengagement, attention maintenance

The Multisensory Attention Assessment Protocol (MAAP): Characterizing Individual Differences in Multisensory Attention Skills in Infants and Children and Relations with Language and Cognition

Selective attention is the gateway for information pickup and processing and the basis for all we perceive, learn, and remember (Bahrick, 2010; Bahrick & Lickliter, 2012, 2014; Gibson, 1969, 1988; Gibson & Pick, 2000; Ruff & Rothbart, 1996; Scerif, 2010). In turn, what we perceive, learn, and remember influences what we attend to next, creating a cycle of attention → perception → learning → memory → attention (Bahrick, 2010; Bahrick & Lickliter, 2012). Attention entails exploratory behaviors integrated across multiple sensory (e.g. vision, audition) and action systems (e.g., eye movements, head turning; Adolph & Berger, 2006; Gibson, 1988; Gibson & Pick, 2000). These processes undergo dramatic development between infancy and childhood (Bahrick, 2010; Colombo, 2001; Gibson, 1988; Ruff & Rothbart, 1996). Researchers are just beginning to characterize the critical role of attention in the typical emergence of perceptual, language, cognitive, and social development (Bornstein, Hahn, & Wolke, 2013; Rose, Feldman, Jankowski, & Van Rossem, 2005; Scerif, 2010; Steele, Karmiloff-Smith, Cornish, & Scerif, 2012).

Moreover, developmental disorders affecting attention (e.g., autism spectrum disorders, ASD; Attention-Deficit/Hyperactivity Disorder, ADHD; dyslexia; William's and Down's syndromes) pose a significant public health concern (Christensen et al., 2016; Visser et al., 2014), and are characterized by language, social, and cognitive impairments (Karmiloff-Smith, 1998; Nigg, 2013; Volkmar, Paul, Klin, & Cohen, 2005). For example, children with ASD show reduced attention to unified multisensory events (Bebko, Weiss, Demark, & Gomez, 2006; Foss-Feig et al., 2010; for a review, see Bahrick & Todd, 2012) and impaired attention shifting and maintenance to social events (Ames & Fletcher-Watson, 2010; Dawson et al., 2004; Landry & Bryson, 2004). These early impairments are thought to cascade to later language and social impairments (Bahrick & Todd, 2012; Dawson et al., 2004; Mundy & Burnette, 2005). Thus, the availability of a standard method for characterizing the typical development of multiple multisensory attention skills and their relations with cognitive, social, and language development will be critical for addressing how these processes go awry and for guiding interventions. We developed the MAAP to address the need.

The Multisensory Attention Assessment Protocol (MAAP)

The MAAP assesses three multisensory attention skills, and the impact of competing visual stimulation on each, in a single protocol: *duration* of looking, *accuracy* of matching audio and visual stimulation, and *speed* of shifting, to social and nonsocial events. It is the first method designed to characterize individual differences in attention to audiovisual events and is appropriate for infants and children alike. Each trial begins with a 3 s dynamic central visual stimulus (silent, colorful moving shapes) followed by two side-by-side lateral (both social or nonsocial) 10 s events with a natural soundtrack synchronous with one of the two events. On half of the trials, the central stimulus remains on during the lateral events, serving as the visual distractor (high competition trials), and on the other half, it is turned off when the lateral events appear (low competition trials; for details, see Supplemental Material, p. 1). The MAAP has many advantages over existing methods. Six goals guided its development.

Include multiple measures in a single protocol. Although attention is viewed as multifaceted (Colombo, 2001; Posner & Petersen, 1990; Ruff & Rothbart, 1996), it has typically been studied piecemeal, with various measures assessed in separate studies using different methods and stimuli, making comparisons across age and studies challenging. The MAAP provides a basis for assessing interrelations among three attention skills and fosters comparisons across studies and ages.

Index the cost of competing stimulation. Attention becomes more efficient, flexible, and selective across development (Colombo, 2001; Ruff & Rothbart, 1996), and increased attentional control is associated with better cognitive and language outcomes (Bornstein & Colombo, 2012; Rose et al., 2005). The MAAP is designed to compare attention on trials with high versus low competing stimulation (central stimulus present vs. absent) for each measure of attention. When competition is high, task difficulty is increased, likely amplifying individual differences. In addition, distracting events are typical in the natural environment and relative to other methods that present no competing stimulation, the MAAP can enhance generalizability to real world learning contexts.

Characterize atypical attention development. Children with ASD show impairments in sustained visual attention, attention shifting, and intersensory matching, especially when disengaging from a concurrent stimulus (Bahrick & Todd, 2012; Dawson et al., 2004; Foss-Feig et al., 2010; Landry & Bryson, 2004). These impairments are most pronounced for social events (Bebko et al., 2006; Dawson et al., 2004; Stevenson et al., 2014) in part, because social events challenge attentional resources by providing high levels of variability and complexity (Bahrick & Todd, 2012; Kennedy & Adolphs, 2012). The MAAP characterizes attention during high vs. low levels of competing visual stimulation and for social and nonsocial events. It thus provides a tool for identifying early impairments in attentional control and identifying asymmetries in attention to social versus nonsocial events characteristic of atypical development.

High ecological validity. Prior studies have primarily assessed attention to static images, devoid of sound or movement (e.g., Colombo, Shaddy, Richman, Maikranz, & Blaga, 2004;

Fagan, Holland, & Wheeler, 2007; Rose, Feldman, & Jankowski, 2003), limiting generalization to the natural, dynamic, multimodal learning environment. Moreover, intersensory processing plays a fundamental role in directing typical attention allocation (e.g., synchrony detection directs looking and listening to the same event). Nevertheless, standard measures of attention (e.g., duration, speed) have rarely been investigated together in the context of multimodal events. The MAAP addresses this need by using dynamic audiovisual events.

Nonverbal measure. The MAAP requires no verbal instructions or responses and is suitable for nonverbal and verbal participants. Typically, nonverbal methods are used with infants (e.g., Colombo et al., 2004; Fagan et al., 2007), whereas methods for children require verbal responses or following instructions. The MAAP provides a single, common protocol for assessing development across infancy and early childhood, the period during which symptoms of developmental disorders (e.g., ASD, ADHD) emerge and are most responsive to intervention.

Stable individual difference measures. The MAAP incorporates multiple relatively short trials—rather than a few longer trials as is typical in much infant research—allowing stable attention patterns to emerge across trials for individual participants. Psychometric properties (e.g., reliability, predictive validity) can be assessed for each measure.

The MAAP provides the first individual difference protocol for nonverbal children assessing multiple multisensory attention skills. This will make it possible, for the first time, to assess relations between the development of attention duration, intersensory accuracy, and speed of shifting to audiovisual events and their relations with specific language, social, and cognitive outcomes. This tool provides a basis for future studies to chart the range of normal variability and to define atypical developmental patterns.

Development of Three Basic Indices of Attention

The attention skills assessed by the MAAP emerge in early infancy and develop rapidly, with improvements in maintaining and shifting attention and in integrating stimulation across the senses (Bahrick, 2010; Colombo, 2001; Ruff & Rothbart, 1996). Research has shown that these attention skills underlie more complex cognitive, social, and language outcomes (e.g., Bornstein & Colombo, 2012; Rose et al., 2005; Salley, Panneton, & Colombo, 2012). Although there is no single method assessing duration of attention, intersensory accuracy, and speed of shifting to dynamic audiovisual events together, there is a rich literature assessing development of each of these skills using various methods.

Duration of attention (maintenance). Infants show faster and more efficient processing across the first year, with decreases in length of visual fixations and overall looking time (Bornstein, Pecheux, & Lecuyer, 1988; Colombo et al., 2004), as well as longer times spent in active processing, according to heart-rate defined phases of attention (Courage, Reynolds, & Richards, 2006). Over the first year of life infants shift from reactive, passive information seeking to more volitional control of attention (see Colombo, 2001; Ruff & Rothbart, 1996), including the ability to sustain attention and inhibit shifting to distracting or competing stimulation (Colombo & Cheatham, 2006; Ruff & Capozzoli, 2003). Greater sustained attention predicts better social and cognitive outcomes (Andrade, Brodeur, Waschbusch, Stewart, & McGee, 2009; Colombo et al., 2004; Murphy, Laurie-Rose, Brinkman, & McNamara, 2007). For example, “short lookers’ are thought to process information more efficiently and show better developmental outcomes than “long lookers” (for a review, see Bornstein & Colombo, 2012).

However, definitions and methods for assessing sustained attention have varied widely. They have included total look length, length of longest look, number of shifts to a distractor, as well as sustained decreases in heart rate (e.g., Colombo et al., 2004; Graziano, Calkins, & Keane, 2011; Richards & Casey, 1991). The MAAP provides an index of “sustained attention” or maintenance, defined as total duration of looking summed across successive looks within a trial, providing a single, common metric that can be used across the lifespan.

Accuracy (intersensory matching). Detecting temporal synchrony (the co-occurrence and temporal alignment of patterns of auditory and visual stimulation) is thought to be a “gateway” to subsequent perceptual processing and is evident even in neonates (Lewkowicz, Leo, & Simion, 2010; Sai, 2005; Slater, Quinn, Brown, & Hayes, 1999). By attending to synchrony (e. g., the sights and sounds of a person speaking) unified events are perceived and meaningful processing can proceed. Research using a group-differences approach demonstrates that, in young infants, sensitivity to temporal synchrony guides audiovisual perception of speech (Kuhl & Meltzoff, 1982; Lewkowicz & Flom, 2013; Patterson & Werker, 1999), the substance and composition of objects (Bahrick, 1987, 1988), emotional expressions (Walker-Andrews, 1997), and mapping linguistic labels onto objects (Gogate & Bahrick, 1998; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Sensitivity to temporal synchrony continues to improve through adolescence (Hillock-Dunn & Wallace, 2012). Although synchrony detection also involves filtering out irrelevant asynchronous stimulation, there is little research on this topic. The MAAP includes trials with high vs. low competing stimulation to index this skill. By using an individual difference measure of intersensory processing (rather than a group-differences approach), the MAAP allows researchers for the first time to assess the specific links between intersensory processing skills (both selecting and filtering), other basic attention skills, and language and social functioning.

Speed of shifting (attention orienting and disengaging). Developmental improvements in the speed of attention shifting occur across the first year of life (Butcher, Kalverboer, & Geuze, 2000; Canfield, Smith, Brezsnyak, & Snow, 1997; Hood & Atkinson, 1993) with possible improvements into adulthood (Munoz, Broughton, Goldring, & Armstrong, 1998). Visual attention shifting, (*orienting*; Butcher et al., 2000; Hood & Atkinson, 1993), emerges early and involves shifting to fixate different locations, objects, or events in the absence of competing stimulation (Colombo, 2001; Ruff & Rothbart, 1996). Shifting away from competing stimulation (*disengaging*) develops later, around 4 months of age, and requires greater attentional flexibility and inhibitory control because it entails termination of attention from an interesting object or event (Johnson, Posner, & Rothbart, 1991; Posner & Petersen, 1990).

Attention shifting is often studied using the “gap-overlap” task (e.g., Landry & Bryson, 2004; Ross-Sheehy, Schneegans, & Spencer, 2015; Saslow, 1967), which presents trials with a static peripheral image with or without a static central stimulus. Infants, children, and adults take longer to shift attention to a peripheral image in the presence of competing stimulation (requiring disengagement) than in its absence (Johnson et al., 1991; Landry & Bryson, 2004; Saslow, 1967). Further, children with ASD show impairments in disengagement compared to typically-developing (TD) controls (Landry & Bryson, 2004), especially for social events (Dawson et al., 2004; though see Fischer, Koldewyn, Jiang, & Kanwisher, 2013). Similar to the “gap-overlap” task, the MAAP presents lateral/peripheral events along with a competing central stimulus (high competition trials) to index disengagement speed, and without a competing central stimulus (low competition trials) to index orienting speed. Rather than one silent static image at a time, however, children view two side-by-side dynamic audiovisual events.

Development of Attention in Childhood

The development of attention from infancy through childhood can be conceptualized as a transition from more external (exogenous) to increased internal (endogenous) attention control (for a review, see Ruff & Rothbart, 1996). Accordingly, studies of attention in childhood have focused on endogenous attention skills such as sustaining attention (maintaining focus despite competing stimulation), selective attention (finding “targets” while ignoring “distractors”), and attentional control (flexible rule set switching while inhibiting reflexive responses). These protocols typically involve understanding directions or making verbal responses (Breckenridge, Braddick, & Atkinson, 2013; Manly, Anderson, Nimmo-Smith, Turner, & Robertson, 2001; Steele et al., 2012). Research with adults and older children has revealed three distinct attention factors: sustained attention, selective attention, and attentional control (Breckenridge et al., 2013; Manly et al., 2001; for additional details, see Supplemental Material, p. 1).

The three attention skills assessed by the MAAP are comparable to those described above. Attention maintenance (duration) is similar to sustained attention. When the central visual distractor is present (high competition), it indexes the ability to focus on the audiovisual events while ignoring irrelevant distractors. Speed of attention shifting in the presence of the distractor indexes disengagement. This requires attentional control—inhibiting looking to the central distractor and shifting attention to one of the lateral events. Shifting on trials with no central stimulus (low competition) serves as a baseline for assessing the cost of disengagement on speed of attention shifting. Intersensory accuracy requires selective attention to one of two concurrent visual events—the one synchronous with the soundtrack—and is assessed in the presence versus absence of the visual distractor (high versus low competition). Performance on the MAAP thus likely reflects all three factors typically assessed in studies with children: sustained attention, attentional control, and selective attention.

Multisensory Attention and Relations with Language

Thus far most studies have taken a group-differences approach to assessing links between multisensory attention and language skills (e.g., Gogate & Bahrick, 1998; Kuhl, Tsao, & Liu, 2003; Lewkowicz & Hansen-Tift, 2012; Patterson & Werker, 1999), limiting research of how basic skills cascade into more complex abilities. Only a few studies, primarily focusing on children with ASD, have assessed relations between individual differences on continuous measures of multisensory attention and language (Patten, Watson, & Baranek, 2014; Righi et al., 2018; Woynaroski et al., 2013; though see Altvater-Mackensen & Grossmann, 2015). Eye-tracking studies with TD children have demonstrated links between attention to the mouth during audiovisual speech and language outcomes (Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan, 2015; Tsang, Atagi, & Johnson, 2018). Using the MAAP, we can now directly assess fine-grained individual differences in three multisensory attention skills and relations with language outcomes, opening the door to understanding developmental pathways.

The Present Studies

In Experiment 1, we evaluate performance on the MAAP in TD 2- to 5-year-olds. Based on prior research, we expected 1) shorter duration (attention maintenance) to the lateral events during high than low competition trials, 2) slower speed (attention shifting) to the lateral events on high competition (requiring disengagement from the central distractor) than low competition trials, and 3) accurate intersensory matching (preference for the synchronous audiovisual event). We also evaluate performance as a function of event: social neutral (with neutral affect), social positive (with positive affect and infant-directed speech, IDS) and nonsocial (dynamic object) events. Finally, we assess concurrent relations between performance on the MAAP and language and cognitive function on the Mullen Scales of Early Learning (MSEL, Mullen, 1995), providing one of the first tests of individual differences in multisensory attention skills as a predictor of language and cognitive skills in TD children.

In Experiment 2, we streamlined the protocol and extended it to 12-month-olds to determine its appropriateness as an individual difference measure for TD infants. Predictions were the same as those for children. We also characterize developmental differences between infants and children along each measure and use results to further shape and refine the protocol.

Experiment 1

Method

Participants. Thirty-six 2- to 5-year-old children (11 females, 25 males) participated (*M* = 3.00 years, *SD* = 1.03; range: 2.00 to 5.75 years) and completed the MAAP and the MSEL.[1] For details on the age distribution, see Supplemental Material, p. 1. Mean-adjusted mental age, from the MSEL, was *M* = 3.17 years (*SD* = 1.14; range: 2.06 to 5.63 years). Participants were recruited using county birth records and contacted via public phone records. Families received a $10 gift card for participating. Twenty-nine children were Hispanic, three were non-Hispanic Caucasian, two were African-American, and two were Asian. Participants came from homes in which the primary language was English (*n* = 11), Spanish (*n* = 11), both English and Spanish (*n* = 13), or Korean (*n* = 1), per parent report. The data of 8 additional children were excluded: 3 for equipment failure, 2 for fussiness, and 3 for MSEL Composite Scores greater than two standard deviations below the standardized norm (*M* = 100, *SD* = 15)—given our goal of characterizing typical development. The research protocol (project: XXXX, protocol number: 15-0010) was approved by the Institutional Review Board at XXXX.

[1] Assuming a β of .80 and a two-tailed *p* value of .05, a sample size of *N* = 36 has sufficient power to detect a Cohen-defined medium effect size of *d* = .57 and *r* = .44.

The Multisensory Attention Assessment Protocol (MAAP).

Stimuli and apparatus. The stimulus events consisted of videotaped displays of dynamic, audiovisual social and nonsocial events (see Figure 1). Films depicted three types of events: 1) women speaking (telling stories) with neutral affect (social neutral, SN), 2) women speaking (telling stories) in affectively positive infant-directed speech (social positive, SP), and 3) objects, including wooden blocks and spools and metal nuts and washers, impacting a surface in an erratic rhythmic pattern (nonsocial, NS). Two pairs of actresses were used for each of the two types of social events, and two pairs of objects were used for the nonsocial events, resulting in four actresses or objects for each event (12 pairs total). Each video was 1-2.5 mins long and then looped. Shorter trials were created from these videos. All videos were bright, colorful, and designed to be interesting to young children. For additional details on stimuli, see Supplemental Material, p. 2. Further, digital films of three geometric animations (floating shape morphing from ball to square, multiple looming spirals, expanding and contracting lines and angles) oriented attention at the beginning of each trial and served as the distracting central stimulus during high competition trials. These images were counterbalanced across the different events, and a new central stimulus was presented on each block of trials.

For each trial, three videos were presented side-by-side: the geometric animations in the center (central stimulus) and the two actresses or two objects on the left and right (lateral events). The three video clips (31 x 23 cm each, framed by 3 cm black space) and their respective audio tracks were edited into a single film using Adobe Premiere CS3 and were presented on a widescreen (102 x 57 cm) monitor. The display was 99 x 23 cm and with the child seated approximately 100 cm away, it subtended a visual angle of approximately 52°. Audio soundtracks were presented via a centrally-located speaker positioned behind the monitor at a mean dB(A) of 61.94 (*SD* = 5.33).

Procedure. Children (accompanied by their parents, who sat behind them or with the child on their lap) sat in front of the widescreen monitor. Each trial began with the central stimulus (geometric animation), followed 3 s later by two concurrent 10 s lateral events on the left and right. 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). Three blocks of 8 trials of each event (SN, SP, NS) were presented (24 trials total). On half the trials within each block (4 trials) the central stimulus remained on while the lateral events were presented (high competition trials), and on the other half, the central stimulus was turned off as soon as the lateral events began (low competition trials). Presentation order of event (SN, SP, or NS first) and competition type (high or low competition trials first) was counterbalanced across participants (6 total presentation orders) so that each child received all possible pairs of actresses and objects (see Supplemental Material, p. 2).

Trained observers, blind to condition and unable to see the videos, recorded visual fixations to the lateral events by depressing a button on a game pad while children viewed the videos. A custom computer program calculated latencies (reaction times, RT) to shift attention from the central stimulus to one of the two lateral events (onset of the lateral event minus onset of the first look to the lateral event indexed by observer's button press), and the duration of individual visual fixations to left and right events for each trial. For trials on which a child was not fixating the central stimulus before the onset of the lateral events, no RT was included. To assess inter-observer reliability, a second observer recorded looking for 11 of the children (31% of the sample). Total fixation time and latency to shift attention from the central stimulus to the lateral event for each trial was calculated independently for each observer. Pearson product moment correlations for judgments of the primary and secondary observers were: duration: .90, accuracy: .91, speed: .90.

Mullen Scales of Early Learning (MSEL). Four subscales of the MSEL (Mullen, 1995), a standardized test of cognitive verbal and nonverbal functioning, were administered to children in English or Spanish, depending on the child's primary home language (per parental report) by a reliable administrator: Visual Reception and Fine Motor to assess nonverbal functioning, and Receptive and Expressive language to assess verbal functioning. The following raw scores (based on all children, $N = 36$) were used for analysis: Visual Reception: $M = 37.03$, $SD = 8.53$; Fine Motor: $M = 32.53$, $SD = 8.89$; Receptive Language: $M = 33.64$, $SD = 7.13$; Expressive Language: $M = 30.56$, $SD = 9.42$. The Early Learning Composite score (an overall score of cognitive functioning) was used to exclude participants with low cognitive functioning ($M = 103.47$, $SD = 15.25$; see Participants).

Results and Discussion

Measures of duration, accuracy, and speed were obtained from each trial and then averaged across trials (see Table 1). There was sufficient variability across children on each of the three measures of attention (see Figure 2) for exploring individual differences in MAAP performance and relations with other variables. The percentage of trials with useable data was quite high ($M = 96\%$, $SD = 5\%$), indicating that children were highly engaged with the displays. For details on exploratory analyses (e.g., outliers, missing data), controls for multiple comparisons, and preliminary analyses on demographics (e.g., home language, ethnicity, gender), see Supplemental Material, p. 3.

Duration, accuracy, and speed as a function of competition and event.

Duration. Duration of attention (maintenance) was indexed by the proportion of available looking time (PALT) on each trial by dividing the total looking time to both lateral events by the length of the trial (10 s; see Table 1 for descriptive statistics and Figure 2a for variability). A 2 competition (high, low) x 3 event (SN, SP, NS) within-participants ANCOVA was conducted, with chronological age in years (mean-centered) entered as a covariate (to control for potential effects of age given the wide range in the sample). The main effect of competition was significant, $F(1, 34) = 21.87$, $p < .001$, *partial eta squared*, $\eta_p^2 = .39$, with shorter PALTs on high than low competition trials, for all events (*p*s < .03; see Figure 3A). This difference was still significant when holding RT differences between high and low competition trials constant, $F(1, 33) = 7.55$, $p = .01$, $\eta_p^2 = .19$ (RT covariate: $p = .11$). Thus, reduced attention maintenance during high compared to low competition trials was not simply due to the slower latencies to disengage on these trials. Both the main effect of event, $F(2, 68) = .45$, $p = .64$, $\eta_p^2 = .01$, and interaction between competition and event failed to reach significance, $F(2, 68) = 1.34$, $p = .27$, $\eta_p^2 = .04$.

Overall, children were highly engaged in the MAAP, spending an average of 74% of the available time fixating the events, with no differences for social versus nonsocial events. Consistent with our predictions, maintaining attention was substantially impaired by the distractor event.

Accuracy. Intersensory accuracy (audiovisual matching) was indexed by the proportion of total looking time (PTLT) to the sound-synchronous lateral events and calculated for each trial by dividing the looking time to the audiovisual synchronous event by the total looking time to both the synchronous and asynchronous event (see Table 1 for descriptive statistics and Figure 2b for variability). To assess evidence of intersensory matching (PTLTs to the synchronous event significantly greater than chance of .50), single sample *t*-tests were conducted. Results indicate significant matching overall, $t(35) = 3.78$, $p = .001$, *Cohen's d* = .63, as well as on both low, *t*(35) = 2.75, *p* = .01, *d* = .50, and high competition trials, *t*(35) = 2.99, *p* = .005, *d* = .46. Significant matching was evident for both SN, *t*(35) = 2.53, *p* = .02, *d* = .42, and SP events, *t*(35) = 3.24, *p* = .003, *d* = .54, but not NS events, *t*(35) = 1.47, *p* = .15, *d* = .24 (except on low competition trials, *t*(35) = 2.04, *p* = .049). None of these findings were qualified by analyses of side biases (preference for left- or right-hand event; see Supplemental Material, p. 4).

A 2 competition (high, low) x 3 event (SN, SP, NS) within-participants ANCOVA (with mean-centered age in years entered as a covariate) was conducted. Although the main effects of competition, *F*(1, 34) = .07, *p* = .79, η_p^2 = .00, and event, *F*(2, 68) = 2.37, *p* = .10, η_p^2 = .07, failed to reach significance, the competition x event interaction was significant, *F*(2, 68) = 3.99, *p* = .02, η_p^2 = .10 (see Figure 3b). Greater intersensory matching for SP than NS events was evident on high competition trials (*p* = .004) but not on low competition trials (*p* = .09). All other comparisons failed to reach significance (*p*s > .12).

In sum, children showed clear evidence of intersensory matching overall and on both high and low competition trials as well as for social events. In contrast with findings on attention maintenance, there was no evidence that the distractor event impaired performance.

Speed. Speed of attention shifting was indexed on each trial by the latency to shift attention (reaction time, RT) in seconds from the central stimulus to one of the two lateral events (see Table 1 for descriptive statistics and Figure 2c for variability). A 2 competition (high, low) x 3 event (SN, SP, NS) within-participants ANCOVA (with mean-centered age in years entered as a covariate) was conducted. The main effect of competition was significant, *F*(1, 34) = 13.39, *p* = .001, η_p^2 = .28, revealing longer RTs on high than low competition trials. The main effect of event failed to reach significance, *F*(2, 68) = 1.29, *p* = .28, η_p^2 = .04. The competition x event interaction was significant, *F*(2, 68) = 4.81, *p* = .01, η_p^2 = .12, and follow-up comparisons revealed longer RTs on high than low competition trials for SN events ($p < .001$, marginally longer for NS events ($p =.06$), but no differences for SP events ($p = .56$; see Figure 3c). No differences between events emerged on high or low competition trials (ps > .04; .at $\alpha = .0167$). For an alternative approach to assessing the cost of competing stimulation on RT, see Supplemental Material, p. 4.

Consistent with our predictions and prior findings (Butcher et al., 2000; Fischer et al., 2013; Landry & Bryson, 2004), speed of shifting was significantly impaired by the distractor event, particularly for affectively neutral speech, with little impairment for nonsocial events.

Reliability and internal consistency. A subgroup ($n = 15$) of the 36 children returned to assess test-retest reliability and revealed high reliabilities for duration, and speed, and moderate reliability for accuracy (see Supplemental Material, p. 5). We assessed internal consistency by calculating the absolute difference between scores on odd and even trials for each child. To the extent that measures are free of random error (i.e., reliable), scores on odd and even trials should be comparable (difference close to zero). This method is superior to split-half correlations, which are subject to artifact (Goodwin & Leech, 2006; Jaccard & Becker, 2009). Inspection of the median absolute differences relative to the range of possible scores for each measure indicates little difference between odd-even trials and thus excellent reliabilities: PALT: .08, range: .00 to 1.00; PTLT: .07, range: .00 to 1.00; RT: .15 s, range: 0 to 10 s).

Developmental change in duration, accuracy, and speed. To assess whether multisensory attention skills increased, decreased, or remained stable across age within Experiment 1, we assessed parameter estimates of the chronological age covariates from the ANCOVA models (see Table 2 for estimates). Duration increased with age, $p = .005$; however, the increase was most evident on low competition trials, $p = .003$ (high competition, $p = .08$).

Increased duration across age was most evident for SP events, $p = .002$, particularly on low competition trials, $p = .005$. Increased accuracy across age was evident for SP events, $p = .02$, and was significant on high, $p = .02$, (but not low) competition trials. No change across age was evident for duration and accuracy for NS events, ps > .13. Thus, between 2 and 5 years of age, children showed improvements in duration and accuracy for social events. No change across age was evident for speed of attention shifting, ps > .27. Shifting speed is reported to improve across the first year of life before leveling off (Canfield et al., 1997; but see Munoz et al., 1998). Thus, we expected changes in shift speed to be evident earlier in development (see Experiment 2).

Individual differences among duration, accuracy, and speed. Given significant relations between attention measures and age, we controlled for effects of chronological age in correlational analyses between measures (using partial correlation coefficients; see Table 3). Overall, children with longer duration showed greater matching, $p = .003$, even when holding age constant. In contrast, there were no significant relations between duration and speed, or between speed and accuracy (ps > .37). Significant correlations were most evident during high competition trials. On high competition trials, children with longer durations showed greater accuracy, $p = .005$ (particularly for social events, ps < .02), as well as faster speed for SN and NS events, ps < .01. On low competition trials, only one significant correlation between duration of attention and speed of shifting was evident, $p = .01$. In sum, individual differences among measures (controlling for chronological age) were most evident when competing stimulation was high and attentional resources were challenged.

Predicting cognitive/language functioning from duration, accuracy, and speed. We assessed relations among the three MAAP measures and cognitive and language scores on the MSEL (see Table 4)—controlling for effects of chronological age, given that raw MSEL scores increase with age (see Supplemental Material, p. 6). Across all trials, longer duration predicted higher scores on all subscales of the MSEL (Receptive and Expressive Language, Visual Reception, and Fine Motor), *p*s < .03, and higher accuracy predicted higher scores on three of the four subscales (except Fine Motor), *p*s < .01. In contrast, no significant relations between any of the MSEL subscales and shift speed emerged (*p*s > .32). Correlations between MSEL subscales with duration and accuracy were carried by performance on low competition trials, with longer duration predicting higher Receptive Language scores, $r(33) = .43$, $p = .01$, and better accuracy predicting higher scores on all four subscales, *r*s(33) > .38, *p*s < .03 (particularly for SN events, see Table 4). No significant relations were evident on high competition trials (*p*s > .054). Correlations as a function of home language environment (English, Spanish), and between MAAP and vocabulary size, are presented in the Supplemental Material, p. 6. Thus, greater duration and accuracy on the MAAP (particularly for social events) predicted both verbal and non-verbal functioning on the MSEL. Shift speed and performance on NS trials were not significant predictors of MSEL scores.

Structural model: Duration and accuracy for social events predict language functioning. Based on the pattern of correlations among multisensory attention skills for social events and MSEL scores, we tested several structural equation models (see Supplemental Material, p. 7) and derived a model with excellent fit and strong relations with language (see Figure 4). Duration of attention (Maintenance) to social events (average across SN and SP events) predicts accuracy (Intersensory) for social events, which, in turn, predicts Receptive and Expressive Language (Language: averaged across the two subscales of the MSEL) with chronological age as a covariate to control for effects of age on language. Maintenance to social events accounted for 36% of the variance in Intersensory, $F(1, 34) = 18.81$, $p < .001$, and

Intersensory and Age together accounted for 81% of the variance in Language, $F(2, 33) = 71.23$, $p < .001$. Intersensory processing alone accounted for a significant 5% of the variance in Language, $p = .02$, above and beyond effects of age. All path coefficients were statistically significant, *p*s < .001, and this was also the case for performance on both high and low competition trials, *p*s < .05.

In summary, the current model is consistent with analyses from both the inter-correlations among the three measures of the MAAP and their correlations with the MSEL. Longer attention maintenance to social events predicts greater intersensory matching and processing of social events, and this in turn predicts greater language functioning. These findings provide new and exciting information about relations between basic attention skills and more complex, derivative skills of language and cognitive function.

Experiment 2

Experiment 1 demonstrated the feasibility of using the MAAP with 2- to 5-year old children and revealed individual differences in three multisensory attention skills and the cost of competing stimulation on each. Experiment 2 was conducted to test the feasibility of using the MAAP with infants and to provide converging evidence across age for our basic findings. We focused on 12-month-old infants because they are rapidly developing social-communicative behaviors at this age (Fenson et al., 2006; Flom, Lee, & Muir, 2007; Mundy et al., 2007; Walden & Ogan, 1988). We streamlined the procedure, using only positive social events (hereafter referred to as social; see MAAP stimuli and apparatus) and nonsocial events, and increased the number and duration of trials for each event while keeping the total number of trials at 24.

Method

Participants. Forty-eight 12-month-olds infants (26 females, 22 males) participated (*M* =

12.06 months, *SD* = .22; range: 11.60 to 12.60 months) and received the MAAP.[2] All infants were participants in an ongoing longitudinal study and the first 48 infants who completed data collection with sufficient useable trials (8 of 24) were included in the sample. As part of the longitudinal project, most participants had previously received the MAAP at 3 and 6 months of age (*n* = 45, and 44, respectively). Participants were recruited through the same means as Experiment 1, and families received $20 for participating. Thirty children were Hispanic, seven were non-Hispanic Caucasian, seven were African-American, two were Asian, one was of more than one race, and one infant's race was unknown. Participants came from homes in which the primary language was English only (*n* = 16), Spanish only (*n* = 5), both English and Spanish (*n* = 25), or other (*n* = 2), according to parent report. The data of 12 additional infants were excluded: 2 for equipment failure, 5 for fussiness, 1 for insufficient data (fewer than 8 trials), and 4 were unable to complete the procedure (e.g., parental interference, falling asleep). All infants were healthy and born full-term, weighing at least five pounds, had an APGAR score of at least 9, and had previous MSEL Composite Scores less than two standard deviations below the standardized norm (MSEL administered at 6 months of age). The research protocol (title: XXXX, protocol number: 13-0448) was approved by the Institutional Review Board at XXXX.

MAAP stimuli and apparatus. Several modifications were made to the MAAP stimuli. Only the positive social events (women speaking in affective positive IDS) were used from Experiment 1. They were judged most appropriate given IDS is highly salient to infants (Cooper & Aslin, 1990; Fernald, 1984; for a review, see Soderstrom, 2007) and it provides exaggerated intersensory information (Bahrick & Todd, 2012; Kubicek et al., 2014; Smith & Strader, 2014). Also, given our focus on assessing developmental change, positive social events were chosen as

[2] Assuming a β of .80 and a two-tailed *p* value of .05, a sample size of *N* = 48 has sufficient power to detect a Cohen defined medium effect size of *d* = .50 and *r* = .40.

performance on these trials showed the most improvement with age in Experiment 1 (see Table 2). Further, we found attention deficits in children with ASD compared with TD controls for the positive social events using the MAAP (Bahrick & Todd, 2012; Todd & Bahrick, manuscript in preparation). We also created new films of nonsocial events consisting of small wooden objects dropping into clear plastic containers of different shapes (see Figure 5; for details, see Supplemental Material, p. 8). These events depicted clear temporal synchrony between visual object impacts against one another and against the sides of the container and the impact sounds produced. Four of the actresses from Experiment 1 were used for the social events, and four objects (each with a different shaped container, and wooden objects of a different color and shape) were used for the nonsocial events, resulting in four actresses or objects (two pairs) for each event (8 total). Two of the geometric animations from Experiment 1 (morphing shape, looming spirals) were used to orient attention at the beginning of each trial and served as the central distractor. Animations were counterbalanced across social and nonsocial event trial blocks.

For each trial, three video clips (33 x 22 cm each, framed by 1.5 cm black space) were edited to be presented side-by side (along with the audio track for one of the clips) using a custom MATLAB-based software: the geometric animations in the center (central distractor) and the two actresses or two objects on the left and right (lateral events). Example stimulus videos are available on Databrary (https://nyu.databrary.org/volume/326). The display was 102 x 22 cm and was presented on a widescreen (102 x 57 cm) monitor. The child sat approximately 100 cm away, and the display subtended a visual angle of approximately 54°. Soundtracks were presented via a centrally-located speaker positioned behind the monitor at a mean dB(A) of 61.47 (*SD* = 2.53).

Procedure. Modifications were also made to the procedure. To protect against the greater possibility of data loss for younger children, we increased the number of trials from 8 to 12 for each event (social, nonsocial), keeping the same total 24 trials as Experiment 1. We also increased the trial duration from 10 to 12 s, allowing more time to process the events. Blocks of 12 social and 12 nonsocial trials (order counterbalanced across participants) were presented, with half (6 trials) in each block high competition and half low competition trials. Finally, the asynchronous events were played 750 ms out of phase with respect to the synchronous event and its soundtrack. All other details were identical to those of Experiment 1.

To assess inter-observer reliability, a second observer recorded looking for 16 of the infants (33.3% of the sample). The Pearson correlation coefficients for the judgments of the primary and secondary observer were: duration: .96, accuracy: .94, speed: .92. Reliability between the live and offline coding for RT was high (see Supplemental Material, p. 8).

Results and Discussion

Infants completed an average of 23.44 out of the 24 trials for attention maintenance, (M = 98%, SD = 5%), indicating the events were highly engaging. For details regarding outliers and missing data, see Supplemental Material, p. 9.

Duration, accuracy, and speed as a function of competition and event.

Duration. See Table 5 for descriptive statistics and Figure 6a for variability in PALT. A 2 competition (high, low) x 2 event (social, nonsocial) within-participants ANOVA was conducted on PALTs. The main effect of competition was significant, $F(1, 47) = 79.89$, $p < .001$, $\eta_p^2 = .63$, with shorter PALTs on high than low competition trials for both social and nonsocial events (ps < .001; see Figure 7A). This difference was still significant when holding RT differences between high and low competition trials constant, $F(1, 46) = 46.06$, $p < .001$, η_p^2 =

.50, similar to Experiment 1. Both the main effect of event, $F(1, 47) = .52$, $p = .48$, $\eta_p^2 = .01$, and the competition x event interaction failed to reach significance, $F(1, 47) = .00$, $p = .98$, $\eta_p^2 = .00$.

Overall, 12-month-olds spent approximately 57% of the available trial time fixating the social and nonsocial events. Similar to Experiment 1, attention maintenance showed a significant decrease (35%) when the distractor was present and did not differ as a function of event.

Accuracy. See Table 5 for descriptive statistics and Figure 6b for variability in PTLT. Single sample *t*-tests against chance (.50) on PTLTs to the sound synchronous events revealed significant matching overall, $t(47) = 3.39$, $p = .001$, $d = .50$, as well as on low, $t(47) = 3.64$, $p = .001$, $d = .57$, but not on high competition trials, $p = .30$. Significant matching was evident for nonsocial, $t(47) = 2.94$, $p = .01$, $d = .38$, but not for social events, $t(47) = .60$, $p = .55$, $d = .11$ (see Figure 7b). None of these findings were qualified by analyses of side biases (see Supplemental Material, p. 9).

A 2 competition (high, low) x 2 event (social, nonsocial) within-participants ANOVA on PTLTs revealed no significant effects of competition, event, or competition x event interaction, *p*s > .13. No planned comparisons reached significance (*p*s > .18; see Figure 7B).

Similar to children in Experiment 1, 12-month-olds showed evidence of intersensory matching overall, in the absence of competing stimulation, and for nonsocial events, with no decrease in matching due to the distractor. In contrast to children, they showed no matching for social events or on trials with competing visual stimulation, and no differences in matching between conditions.

Speed. See Table 5 for descriptive statistics and Figure 6c for variability in RTs. A 2 competition (high, low) x 2 event (social, nonsocial) within-participants ANOVA was conducted on RT scores. The main effect of competition was significant, $F(1, 47) = 16.25$, $p < .001$, $\eta_p^2 =$

.26, revealing longer RTs for high than low competition trials, for both social and nonsocial events, *p*s < .003 (Figure 7c). The main effect of event was marginally significant, $F(1, 47) = 3.78$, $p = .06$, $\eta_p^2 = .07$, indicating marginally greater RTs to nonsocial than social events. However, the competition x event interaction was not significant, $F(1, 70) = .41$, $p = .52$, $\eta_p^2 = .01$. For results using an offline measure of RT, as well as an RT disengagement difference scores, see Supplemental Material, p. 9.

Similar to children in Experiment 1, 12-month-olds showed longer latencies to shift attention to the audiovisual events when the distractor was present (requiring disengagement) than when it was absent, for both social and nonsocial events.

Internal consistency. Similar to Experiment 1, we assessed internal consistency via split-half absolute disparity scores. Results again indicated excellent reliabilities via small median absolute differences relative to the ranges of possible scores for each measure: PALT: .04, range: .00 to 1.00; PTLT: .06, range: .00 to 1.00; RT: .11 s, range: 0 to 12 s).

Individual differences among duration, accuracy, and speed. Pearson product-moment correlations were conducted to assess relations among three measures of attention (see Table 6). Across all trials, all three measures were significantly correlated. Infants with greater attention duration showed greater accuracy, $p = .01$, as well as faster shift speeds, $p = .002$. Further, infants with faster shift speeds showed greater matching, $p < .001$. These significant correlations were also evident on high competition trials, *p*s < .04, but few correlations emerged for low competition trials (see Table 6). Like Experiment 1, individual differences were most evident when competing stimulation was high and attentional resources were taxed.

Comparisons across Experiments 1 and 2: Children versus infants. Analyses were conducted to compare the performance of 12-month-olds (Experiment 2) with that of children

(Experiment 1). A 2 experiment (1: children, 2: infants) by competition (high vs. low) by 2 event (social, nonsocial) ANOVA was conducted for each of the measures (using only the positive social and nonsocial events from Experiment 1). Main effects of experiment emerged for duration and speed, with children showing significantly longer look durations, $F(1, 82) = 57.60$, $p < .001$, $\eta_p^2 = .41$, and faster shifting than infants, $F(1, 82) = 41.65$, $p < .001$, $\eta_p^2 = .34$, but comparable accuracy of intersensory matching, $p = .33$ (with both groups showing significant matching). There were main effects of competition for all measures indicating reduced duration and accuracy, and slower speeds, in the high compared to the low competition conditions, *F*s > 5.92, *p*s < .02, η_p^2s > .07. However, there were also significant experiment x competition interactions for duration, $F(1, 82) = 8.67$, $p = .004$, $\eta_p^2 = .10$, and for speed, $F(1, 82) = 3.81$, $p = .054$, $\eta_p^2 = .04$. Infants showed greater impairments (greater high minus low competition difference scores) from the distractor than children, both for duration (infants: $M = .24$, $SD = .19$; children: $M = .12$, $SD = .16$; $p = .004$) and for speed (infants: $M = .41$, $SD = .71$; children: $M = .28$, $SD = .46$; $p = .05$). No main effects of event emerged; however, there was a significant experiment x event interaction for accuracy, $F(1, 82) = 5.79$, $p = .02$, $\eta_p^2 = .07$. Children showed better matching than infants for social events ($p = .022$) but comparable matching for nonsocial events ($p = .26$). Together, these findings indicate significant improvements in attention maintenance and speed of shifting, as well as matching for social events, across infancy and childhood, and comparable intersensory matching for nonsocial events. Further, disengaging and maintaining attention in the presence of a distractor impaired attention in infants more than in children.

General Discussion

Although attention provides the foundation for language, social, and cognitive development, the early development of attention remains poorly understood in large part because of the lack of a fine grained, individual difference measure of basic attention skills appropriate for nonverbal children—critical for assessing continuity and change across infancy and childhood with common metrics. The Multisensory Attention Assessment Protocol (MAAP) addresses this need. It is a nonverbal, individual difference measure for assessing three fundamental indices of attention to audiovisual events during high versus low competing stimulation, in a single protocol, appropriate for both infants and children. Children view many short trials beginning with a dynamic central visual stimulus, followed by two side-by-side social or nonsocial events along with the synchronous soundtrack to one event. During the lateral events, on half of the trials the central stimulus (distractor) remains on (providing high competition) and on half of the trials, it is turned off (low competition). The MAAP can index three multisensory attention skills in the presence of high versus low competing stimulation, and their relations with language, social, and cognitive functioning. Further, by providing dynamic audiovisual events in the context of competing stimulation —in contrast to static images with no competing stimulation (typical of most attention studies)—findings of the MAAP can better generalize to natural, multimodal learning environments.

To establish the viability of the MAAP as an individual difference measure of multisensory attention skills appropriate across infancy and childhood, we presented the protocol to 2- to 5-year-old TD children (Experiment 1) and to 12-month-old TD infants (Experiment 2). In Experiment 1, we also explored whether the MAAP would predict language and cognitive skills in 2- to 5-year-olds. There were three important areas of discovery. First, we generated some of the first individual difference data indexing three fundamental attention skills to audiovisual social and nonsocial events in a single protocol for both infants and children.

Second, we documented several important and novel aspects of attention development. Third, we document relations between multisensory attention skills and language and cognitive development in 2- to 5-year-olds.

Multisensory Attention Skills and the Effects of Competing Stimulation

Infants (Experiment 2) and children (Experiment 1) showed strikingly similar effects of competing stimulation on duration, speed, and accuracy of multisensory attention skills. Although both infants and children showed high levels of *attention maintenance* (duration) to the audiovisual social and nonsocial events, they displayed significantly lower maintenance (shorter duration of looking) in the presence of the distractor (high competition trials) than in its absence (low competition trials). Similarly, at both ages, *speed of shifting* to the lateral events was significantly slower during competing stimulation (requiring disengaging attention from the central distractor) than in its absence (requiring only orienting attention). *Accuracy of intersensory matching* was significantly greater than chance for both children and infants—they reliably looked to the visual event synchronized with its natural soundtrack. Interestingly, for infants, intersensory matching was only consistently significant in the absence of the distractor (low competition trials). Moreover, across both groups attention was more sustained, faster, and more accurate in conditions of low competing stimulation. Thus, the MAAP provides 1) reliable evidence of three different multisensory attention skills (duration, accuracy, speed) for both infants and children, 2) common metrics for assessing these multisensory attention skills across infants and children, and 3) similar effects of a novel and fundamental variable—absence vs. presence of competing stimulation—on multisensory attention skills across infants and children.

Several notable developmental differences between infants and children were also found. Children showed significantly longer attention maintenance (74% of available looking time to the lateral events) than infants (57% looking) and attention maintenance was less impaired by competing stimulation (15% decrement in total looking duration) than for infants (35% decrement). Children also demonstrated significantly faster shifting (0.78 s overall) than infants (1.19 s overall), and this speed advantage was present in both high and low competing stimulation. Although both infants and children showed intersensory matching when competing stimulation was low (no distractor), only children showed matching when competing stimulation was high (distractor). Thus, children are significantly better at filtering out irrelevant, distracting stimulation than infants, with advantages for duration, accuracy, and speed of attention. Further, the largest developmental gains occurred between infancy and early childhood and were most evident when competing stimulation was high.

Multisensory Attention Skills and Social versus Nonsocial Events

In contrast to high vs. low competition, few overall differences in attention emerged for social vs. nonsocial events. Unlike infants (Experiment 2), children (Experiment 1) showed no decrement in speed of shifting to view positive social events as a result of high competing stimulation. This speed advantage for affectively positive social events may be due to their exaggerated intersensory information and attentional salience compared to nonsocial events (Bahrick & Todd, 2012). Further, intersensory matching for infants was only significant for nonsocial events and had not yet emerged for social events. In contrast, for children, intersensory matching was evident for both event types, but was greatest for social events. This may reflect developmental differences in relevant experience: 12-month-olds have typically been engaging with object manipulation (Eppler, 1995; Soska & Adolph, 2014), but 2- to 5-year-old children have had an extraordinary amount of experience interacting and communicating with social partners (Flom et al., 2007; Mundy et al., 2007). Given that information provided by social events tends to be more complex and variable compared to nonsocial events (Adolphs, 2009; Bahrick & Todd, 2012; Dawson et al., 2004), matching audiovisual stimulation from social events may require more experience with the social world.

In Experiment 1, duration of attention continued to increase across 2 to 5 years for positive social events (but not for nonsocial events). This is consistent with evidence that sustained attention improves and increases across early development, including both heart-rate defined measures of attention (Courage et al., 2006) and visual attention (Ruff & Capozzoli, 2003). We found a similar increase for intersensory accuracy for positive social (but not nonsocial) events. These findings suggest that intersensory processing of speaking faces continues to become more refined and precise across 2- to 5-years of age, likely as a result of experience with social events.

Interrelations Among Measures of Attention

The MAAP provides the first individual difference measure capable of indexing three fundamental dimensions of attention (duration, accuracy, speed), and their interrelations, within a single protocol. We found significant relations among all three measures of attention. Both children (Experiment 1) and infants (Experiment 2) showed significant correlations between the duration of attention maintenance and the accuracy of intersensory matching. Longer attention maintenance may lead to increased opportunities for detecting audiovisual synchrony and, in turn, greater intersensory matching may lead to longer attention maintenance. Further, for both infants and children, faster speed of shifting predicted significantly longer duration of attention. Finally, for infants, faster speed of shifting was associated with better intersensory matching. Infants who quickly disengaged from competing stimulation likely had greater attentional resources, promoting longer maintenance and better intersensory matching. In addition, there were significant relations between duration and accuracy for social events in children, but for nonsocial events in infants. These developmental differences may stem from the finding that children showed intersensory matching for social events but infants showed matching for nonsocial events.

Interestingly, for both infants and children, all of these relations were stronger in the presence of competing stimulation from the central distractor (5 of 6 correlations were significant) than in its absence (no significant correlations). Competing stimulation taxes the attentional resources. This reduces ceiling performance and exaggerates individual differences, revealing both strengths and weaknesses in attentional skills. The fact that a number of significant relations among attention skills emerged with the addition of competing stimulation suggests that there is a tighter coupling among attention skills when the attention system is challenged.

Multisensory Attention Skills and Relations with Language and Cognition in Children

Perhaps one of the most exciting empirical findings from Experiment 1 is that fine-grained individual differences in intersensory processing skills in TD children predict language and cognitive functioning. Prior studies using group-level analyses have suggested this may be the case (e.g., Gogate & Bahrick, 1998; Kuhl et al., 2003; Lewkowicz & Hansen-Tift, 2012); however, few have assessed links between multisensory attention skills and language/cognitive outcomes using an individual differences approach (but see Patten et al., 2014; Righi et al., 2018; Woynaroski et al., 2013, for studies with children with ASD). We found that both longer attention maintenance (duration of looking) and better intersensory matching (accuracy) predicted higher verbal (receptive and expressive language) and non-verbal functioning on the MSEL. Further, intersensory matching for social events—in particular—was highly correlated with both receptive and expressive language scores (*r*s > .50). These findings were characterized by a structural model providing some of the first evidence in TD children for how basic attention processes interact and contribute to language outcomes: *longer looking to social events predicts greater intersensory matching of social events, which, in turn, predicts receptive and expressive language functioning* (even after controlling for chronological age). In contrast, the model did not hold for nonsocial events, indicating the importance of attention to social events as a foundation for language, consistent with proposals in the literature (e.g., Bahrick & Todd, 2012; Feldman, 2007; Kuhl, 2007; Mundy & Burnette, 2005). For example, we have proposed that selective attention to amodal properties of audiovisual speech forms a critical foundation for cognitive and language skills (Bahrick, 2010; Bahrick & Lickliter, 2012; Bahrick & Todd, 2012) and Kuhl (2007) describes language as "gated by the social brain." Only with measures appropriate for assessing fine-grained individual differences in multiple multisensory attention skills, can researchers directly test these hypotheses and explore how basic skills lead to more complex language skills that rely on a multisensory foundation.

The Promise of the MAAP as an Individual Difference Measure of Attention

The present study confirms the viability of the MAAP as an individual difference measure of three basic indices of attention (duration, accuracy, speed). Children (Experiment 1) and infants (Experiment 2) appear to enjoy the procedure and complete a large number of trials (average of 97% of the 24 trials with useable data). Both age groups also show sufficient individual differences along each MAAP measure to provide a basis for correlations among measures and with outcomes. Further, the MAAP can be used in longitudinal studies to characterize how individual differences in basic attention skills cascade into more complex language, cognitive, or social skills that rely on this foundation.

In addition, the MAAP was designed to be appropriate for identifying atypical attention patterns in neurodevelopmental disorders such as ASD. It provides a nonverbal measure of multisensory attention skills, appropriate for assessing change between infancy and early childhood, the period when symptoms of ASD first emerge. Children with ASD show impaired language and communication skills as well as early deficits in attention maintenance (duration), intersensory processing (accuracy), and attention shifting (speed), particularly to social events (Bahrick & Todd, 2012; Bebko et al., 2006; Dawson et al., 2004; Landry & Bryson, 2004). The MAAP successfully indexes all these attention skills, both in the presence and absence of competing stimulation, characterizing performance when attention is challenged. Moreover, because it assesses attention to social and nonsocial events, the MAAP is designed to reveal any asymmetry in attention allocation and intersensory processing in favor of nonsocial events, a potential indicator of atypical development consistent with autism.

Limitations, Future Directions, and General Conclusions

Several steps are needed for further refining and establishing the MAAP as a reliable individual difference measure. First, test-retest reliability, although promising with the present subsample of children (N = 15; see Supplemental Material, p. 5), it must be established for a larger sample across a broader age range. Further, it will be important to validate the measures of the MAAP against already established measures of attention to more precisely link the measures with constructs already validated.

The MAAP will also need to be tested with a larger sample and across a broader age range. Although we characterized developmental differences between 12-month-old infants in Experiment 2 and 2- to 5-year-old children in Experiment 1, the sample of children was relatively smaller (N = 36). It also had a heavy representation of male (primarily Hispanic) participants, and relative to 2-year-olds, fewer 3- to 5-year-olds, limiting power for assessing developmental trends across 2 to 5 years. Future studies are needed (and are currently in progress in our lab) with a larger sample and assessing longitudinal developmental trajectories for measures assessed by the MAAP and relations with cognitive, social, and language outcomes. Finally, research exploring neural and psychophysiological correlates of these attention skills will also be critical for learning more about their nature and pathways of influence in typical and atypical developmental trajectories.

In sum, the MAAP provides the foundation for a new and unique individual difference measure for assessing multiple aspects of attention to audiovisual events in a single protocol appropriate for nonverbal children and infants. The present study demonstrates the feasibility of using this method with both infants and young children. The availability of the MAAP opens the door to assessing relations among multisensory attention skills (duration, accuracy, speed) and the effects of concurrent, distracting events on these skills. The MAAP will allow us to understand more about how basic multisensory attention skills create a foundation for higher level abilities by assessing developmental pathways between basic skills and social, cognitive, and language functioning in children of typical and atypical development.

References

Adolph, K. E., & Berger, S. E. (2006). Motor development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology, Vol. 2: Cognition, perception, and language* (pp. 161–213). New York: Wiley. doi: 10.1002/9780470147658.chpsy0204

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716. doi: 10.1146/annurev.psych.60.110707.163514

Altvater-Mackensen, N., & Grossmann, T. (2015). Learning to match auditory and visual speech cues: Social influences and acquisition of phonological categories. *Child Development*, *86*, 362–378. doi: 10.1111/cdev.12320

Ames, C., & Fletcher-Watson, S. (2010). A review of methods in the study of attention in autism. *Developmental Review*, *30*, 52–73. doi: 10.1016/j.dr.2009.12.003

Andrade, B. F., Brodeur, D. A., Waschbusch, D. A., Stewart, S. H., & McGee, R. (2009). Selective and sustained attention as predictors of social problems in children with typical and disordered attention abilities. *Journal of Attention Disorders*, *12*, 341–52. doi: 10.1177/1087054708320440

Bahrick, L. E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development*, *10*, 387–416. doi: 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*, 197–209. doi: 10.2307/1130402

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). Malden, MA: Wiley-Blackwell. doi: 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). New York: Oxford University Press. doi: 10.1093/acprof:oso/9780199586059.003.0008

Bahrick, L. E., & Lickliter, R. (2014). Learning to attend selectively: The dual role of intersensory redundancy. *Current Directions in Psychological Science*, *23*, 414–420. doi: 10.1177/0963721414549187

Bahrick, L. E., Soska, K. C., & Todd, J. T. (2018). *Assessing individual differences in intersensory processing in young children: The Intersensory Processing Efficiency Protocol*. Manuscript submitted for publication.

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). Cambridge, MA: MIT Press.

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*, 88–98. doi: 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). Mahwah, NJ: 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). Skokie, IL: Scientific Software International.

Bornstein, M. H., & Colombo, J. (2012). Infant cognitive functioning and mental development. In S. M. Pauen (Ed.), *Early childhood development and later achievement* (pp. 118–147). Cambridge, United Kingdon: Cambridge University Press.

Bornstein, M. H., Hahn, C.-S., & Wolke, D. (2013). Systems and cascades in cognitive development and academic achievement. *Child Development*, *84*, 154–162. doi: 10.1111/j.1467-8624.2012.01849.x

Bornstein, M. H., Pecheux, M.-G., & Lecuyer, R. (1988). Visual habituation in human infants: Development and rearing circumstances. *Psychological Research*, *50*, 130–133. doi: 10.1007/BF00309213

Breckenridge, K., Braddick, O., & Atkinson, J. (2013). The organization of attention in typical development: A new preschool attention test battery. *The British Journal of Developmental Psychology*, *31*, 271–288. doi: 10.1111/bjdp.12004

Butcher, P. R., Kalverboer, A. F., & Geuze, R. H. (2000). Infants' shifts of gaze from a central to a peripheral stimulus: A longitudinal study of development between 6 and 26 weeks. *Infant Behavior and Development*, *23*, 3–21. doi: 10.1016/S0163-6383(00)00031-X

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development*, *62*(2, Serial No. 250), 1–145. doi: 10.2307/1166196

Christensen, D. L., Baio, J., Braun, K. V. N., Bilder, D., Charles, J., Constantino, J. N., … Yeargin-Allsopp, M. (2016). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2012. *MMWR Surveillance Summaries*, *65*, 1–23. doi: 10.15585/mmwr.ss6503a1

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

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. In R. Kail (Ed.), *Advances in Child Development and Behavior* (Vol. 34, pp. 283–322). San Diego, CA: Academic Press. doi: 10.1016/S0065-2407(06)80010-8

Colombo, J., Shaddy, D. J., Richman, W. A., Maikranz, J. M., & Blaga, O. M. (2004). The developmental course of habituation in infancy and preschool outcome. *Infancy*, *5*, 1–38. doi: 10.1207/s15327078in0501_1

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*, 1584–1595. doi: 10.2307/1130766

Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development*, *77*, 680–695. doi: 10.1111/j.1467-8624.2006.00897.x

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. doi: 10.1037/0012-1649.40.2.271

Eppler, M. A. (1995). Development of manipulatory skills and the deployment of attention.

Infant Behavior & Development, *18*, 391–405.

Fagan, J. F., Holland, C. R., & Wheeler, K. (2007). The prediction, from infancy, of adult IQ and achievement. *Intelligence*, *35*, 225–231. doi: 10.1016/j.intell.2006.07.007

Feldman, R. (2007). Parent-infant synchrony: Biological foundations and developmental outcomes. *Current Directions in Psychological Science*, *16*, 340–345. doi: 10.1111/j.1467-8721.2007.00532.x

Fenson, L., Marchman, V., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2006). *The MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Brookes Publishing.

Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, D. Garvey, & R. Golinkoff (Eds.), *The origins and growth of communication* (pp. 5–29). Norwood, NJ: Ablex.

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

Flom, R., Lee, K., & Muir, D. (2007). *Gaze-following: It's development and significance*. Mahwah, NJ: Lawrence Erlbaum Associates.

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*, 381–389. doi: 10.1007/s00221-010-2240-4

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

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*, 1–41. doi: 10.1146/annurev.psych.39.1.1

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* New York, NY, US: Oxford University Press.

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*, 133–149. doi: 10.1006/jecp.1998.2438

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14,* 23–45. doi: 10.1017/S030500090001271X

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *The Journal of Experimental Education*, *74*, 249–266. doi: 10.3200/JEXE.74.3.249-266

Graziano, P. A., Calkins, S. D., & Keane, S. P. (2011). Sustained attention development during the toddlerhood to preschool period: Associations with toddlers' emotion regulation strategies and maternal behaviour. *Infant and Child Development*, *20*, 389–408. doi: 10.1002/icd

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*, *15*, 688–696. doi: 10.1111/j.1467-7687.2012.01171.x

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

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant*

Behavior and Development, *16*, 405–422. doi: 10.1016/0163-6383(93)80001-O

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

Jaccard, J. J., & Guilamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of Clinical Child and Adolescent Psychology*, *31*, 130–146. doi: 10.1207/S15374424JCCP3101_15

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. doi: 10.1207/s15374424jccp3503_11

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, *3*, 335–344. doi: 10.1162/jocn.1991.3.4.335

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2,* 389–398. doi: 10.1016/S1364-6613(98)01230-3

Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends in Cognitive Sciences*, *16*, 559–72. doi: 10.1016/j.tics.2012.09.006

Kubicek, C., Gervain, J., Hillairet de Boisferon, A., Pascalis, O., Lœvenbruck, H., & Schwarzer, G. (2014). The influence of infant-directed speech on 12-month-olds’ intersensory perception of fluent speech. *Infant Behavior and Development*, *37*, 644–651. doi: 10.1016/j.infbeh.2014.08.010

Kuhl, P. K. (2007). Is speech learning “gated” by the social brain? *Developmental Science, 10,* 110–120. doi: 10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*, 1138–1141. doi: 10.1126/science.7146899

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–101. doi: 10.1073/pnas.1532872100

Landa, R. J., Gross, A. L., Stuart, E. A., & Faherty, A. (2013). Developmental trajectories in children with and without autism spectrum disorders: The first 3 years. *Child Development*, *84*, 429–442. doi: 10.1111/j.1467-8624.2012.01870.x

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

Lewkowicz, D. J., & Flom, R. (2013). The audiovisual temporal binding window narrows in early childhood. *Child Development*. doi: 10.1111/cdev.12142

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*, 1431–1436. doi: 10.1073/pnas.1114783109

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, *15*, 46–60. doi: 10.1111/j.1532-7078.2009.00005.x

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., & Robertson, I. H. (2001). The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *Journal of Child Psychology and Psychiatry*, *42*, 1065–1081. doi: 10.1017/S0021963001007909

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

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

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). Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9780470939345.ch25

Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research*, *121*, 391–400. doi: 10.1007/s002210050473

Murphy, L. M. B., Laurie-Rose, C., Brinkman, T. M., & McNamara, K. A. (2007). Sustained attention and social competence in typically developing preschool-aged children. *Early Child Development and Care*, *177*, 133–149. doi: 10.1080/03004430500349559

Nigg, J. T. (2013). Attention deficits and hyperactivity-impulsivity: What have we learned, what next? *Development and Psychopathology*, *25*, 1489–1503. doi: 10.1017/S0954579413000734

Patten, E., Watson, L. R., & Baranek, G. T. (2014). Temporal synchrony detection and associations with language in young children with ASD. *Autism Research and Treatment*, *2014*. doi: 10.1155/2014/678346

Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, *22*, 237–247. doi:

10.1016/S0163-6383(99)00003-X

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. doi: 10.1146/annurev.neuro.13.1.25

Richards, J. E., & Casey, B. J. (1991). Heart-rate variability during attention phases in young infants. *Psychophysiology*, *28*, 43–53. doi: 10.1111/j.1469-8986.1991.tb03385.x

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*, 1–9. doi: 10.1002/aur.1918

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2003). Infant visual recognition memory: Independent contributions of speed and attention. *Developmental Psychology*, *39*, 563–571. doi: 10.1037/0012-1649.39.3.563

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2005). Pathways from prematurity and infant abilities to later cognition. *Child Development*, *76*, 1172–1184. doi: 10.1111/j.1467-8624.2005.00842.x-i1

Ross-Sheehy, S., Schneegans, S., & Spencer, J. P. (2015). The Infant Orienting With Attention Task: Assessing the neural basis of spatial attention in infancy. *Infancy*, *20*, 467–506. doi: 10.1111/infa.12087

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, *39*, 877–890. doi: 10.1037/0012-1649.39.5.877

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

Sai, F. Z. (2005). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development*, *14*, 29–50. doi:

10.1002/icd.376

Salley, B., Panneton, R. K., & Colombo, J. (2012). Separable attentional predictors of language outcome. *Infancy, 18*, 462–489. doi: 10.1111/j.1532-7078.2012.00138.x

Saslow, M. G. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *Journal of the Optical Society of America*, *57*, 1024–1029. doi: 10.1364/JOSA.57.001030

Scerif, G. (2010). Attention trajectories, mechanisms and outcomes: At the interface between developing cognition and environment. *Developmental Science*, *13*, 805–812. doi: 10.1111/j.1467-7687.2010.01013.x

Slater, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory-visual pairings. *Developmental Science*, *2*, 333–338. doi: 10.1111/1467-7687.00079

Smith, N. A., & Strader, H. L. (2014). Infant-directed visual prosody: Mothers' head movements and speech acoustics. *Interaction Studies*, *15*, 38–54. doi: 10.1075/is.15.1.02smi

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*, 501–532. doi: 10.1016/j.dr.2007.06.002

Soska, K. C., & Adolph, K. E. (2014). Postural position constrains multimodal object exploration in infants. *Infancy*, *19*, 138–161. doi: 10.1111/infa.12039

Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The multiple subfunctions of attention: Differential developmental gateways to literacy and numeracy. *Child Development*, *83*, 2028–2041. doi: 10.1111/j.1467-8624.2012.01809.x

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. doi: 10.1523/JNEUROSCI.3615-13.2014

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*, 1173–1190. doi: 10.1017/S0305000914000725

Todd, J. T., & Bahrick, L. E. (2018). *Impairments in social attention and intersensory processing in children with autism spectrum disorders compared to typically developing children: Evidence using the Multisensory Attention Assessment Protocol (MAAP)*. Manuscript in preparation.

Tsang, T., Atagi, N., & Johnson, S. P. (2018). Selective attention to the mouth is associated with expressive language skills in monolingual and bilingual infants. *Journal of Experimental Child Psychology*, *169*, 93–109. doi: 10.1016/j.jecp.2018.01.002

Visser, S. N., Danielson, M. L., Bitsko, R. H., Holbrook, J. R., Kogan, M. D., Ghandour, R. M., … Blumberg, S. J. (2014). Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/hyperactivity disorder: United States, 2003-2011. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*, 34–46. doi: 10.1016/j.jaac.2013.09.001

Volkmar, F. R., Paul, R., Klin, A., & Cohen, D. (2005). *Handbook of autism and pervasive developmental disorders, Vol. 1: Diagnosis, development, neurobiology, and behavior.* (3rd ed.). Hoboken, NJ, US: John Wiley & Sons Inc.

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development*, *59*, 1230–1240.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, *121*, 437–456. doi: 10.1037//0033-

2909.121.3.437

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*. doi: 10.1007/s10803-013-1836-5

Table 1

Experiment 1: Means (standard deviations in parentheses) for duration: attention maintenance (proportion of available looking time, PALT), accuracy: intersensory matching (proportion of total looking time to synchronous event, PTLT), and speed: attention shifting (reaction time, RT) as a function of type of competition (low, high) and event (social neutral, social positive, nonsocial).

		Event			
		Social Neutral	Social Positive	Nonsocial	Overall
Duration: Attention Maintenance (PALT)	Low Competition	.80 (.19)	.78 (.16)	.83 (.12)	.80 (.12)
	High Competition	.66 (.21)	.69 (.19)	.67 (.20)	.68 (.15)
	Overall	.73 (.17)	.73 (.14)	.75 (.11)	.74 (.11)
Accuracy: Intersensory Matching (PTLT)	Low Competition	.52 (.15)	.57 (.12)**	.53 (.10)**ƒ*	.54 (.08)**
	High Competition	.57 (.14)**	.54 (.11)**ƒ*	.50 (.11)	.54 (.07)**
	Overall	.55 (.11)*	.56 (.09)**	.52 (.07)	.54 (.06)***
Speed: Attention Shifting (RT)	Low Competition	.57 (.20)	.71 (.45)	.64 (.25)	.64 (.21)
	High Competition	1.15 (.91)	.78 (.49)	.84 (.61)	.93 (.45)
	Overall	.86 (.49)	.75 (.31)	.74 (.34)	.78 (.27)

Note: For PTLT, single sample *t*-tests; * $p < .05$, ** $p < .01$, *** $p < .001$; *ƒ* did not meet significance cutoff ($p = .0167$) when controlling for familywise error.

Table 2

Experiment 1: Pearson correlation coefficients between chronological age and each of the three indices of attention—duration: attention maintenance (proportion of available looking time, PALT), accuracy: intersensory matching (proportion of total looking time, PTLT), and speed: attention shifting (reaction time, RT)—as a function of type of competition (low, high) and event (social neutral, social positive, nonsocial).

	Duration: Attention Maintenance			
	Social Neutral	Social Positive	Nonsocial	Overall
Low Competition	.31†	.46**	.26	.47**
High Competition	.30†	.35**ƒ*	.01	.29†
Overall	.37**ƒ*	.50**	.15	.46**
	Accuracy: Intersensory Matching			
	Social Neutral	Social Positive	Nonsocial	Overall
Low Competition	.07	.23	.01	.15
High Competition	.38**ƒ*	.39*	-.22	.32†
Overall	.29†	.38*	-.17	.31†
	Speed: Attention Shifting			
	Social Neutral	Social Positive	Nonsocial	Overall
Low Competition	-.02	-.13	-.01	-.11
High Competition	-.19	-.05	.11	-.10
Overall	-.18	-.13	-.09	-.12

Note: † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$; *ƒ* did not meet significance cutoff ($p = .0167$) when controlling for familywise error.

Table 3

Experiment 1: Partial correlations between the three indices of attention—for duration of looking (Dur), accuracy of intersensory matching (Acc), and speed of attention shifting (Spd)—controlling for chronological age, as a function of type of competition (low, high) and event (social neutral, social positive, nonsocial).

	Low Competition Trials											
	Social Neutral			Social Positive			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Duration (Dur)	---	.34*ƒ	.18	---	.13	-.44**	---	.05	-.31	---	.29	-.12
Accuracy (Acc)		---	.08		---	.13		---	-.04		---	-.05
Speed (Spd)			---			---			---			---
	High Competition Trials											
	Social Neutral			Social Positive			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Duration (Dur)	---	.39**	-.45**	---	.46**	-.20	---	-.05	-.44**	---	.46**	-.37*ƒ
Accuracy (Acc)		---	.07		---	.24		---	-.36*ƒ		---	.01
Speed (Spd)			---			---			---			---
	Overall											
	Social Neutral			Social Positive			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Duration (Dur)	---	.50**	-.15	---	.26	-.38*ƒ	---	.01	-.16	---	.56***	-.16
Accuracy (Acc)		---	.15		---	.22		---	-.25		---	.08
Speed (Spd)			---			---			---			---

Note: * $p < .05$, ** $p < .01$, *** $p < .001$; ƒ did not meet significance cutoff ($p = .0167$) when controlling for familywise error

Table 4

Experiment 1: Partial correlation between the raw scores of the four MSEL scales (Visual Reception, Fine Motor, Receptive Language, Expressive Language) and each of the three indices of attention—duration of looking (Dur), accuracy of intersensory matching (Acc), and speed of attention shifting (Spd)—controlling for chronological age, as a function of type of competition (low, high) and event (social neutral, social positive, nonsocial).

	Low Competition Trials											
	Social Neutral			Social Positive			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Visual Reception	.23	.31†	-.11	.26	.40**f*	.04	.13	.07	.08	.31	.40*	.02
Fine Motor	.34**f*	.41*	-.13	.17	.30†	-.12	.22	-.03	.14	.36**f*	.38*	-.07
Receptive Language	.34**f*	.51**	-.13	.24	.35**f*	-.34	.30†	.24	.00	.43*	.57***	-.29
Expressive Language	.39**f*	.50**	.04	.11	.37**f*	.03	.26	.10	.11	.38	.51**	.08
	High Competition Trials											
	Social Neutral			Social Positive			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Visual Reception	.29†	.33**f*	-.21	.20	.24	.00	.24	-.18	-.16	.33	.22	-.21
Fine Motor	.09	.34**f*	-.03	.26	.12	.03	.17	-.37**f*	.12	.23	.07	.04
Receptive Language	.24	.33**f*	-.23	.25	.23	.14	.14	-.26	.13	.28	.17	-.05
Expressive Language	.19	.38**f*	-.02	.23	.19	.11	.22	-.30	.19	.29	.17	.11
	Overall											
	Social Neutral			Social Positive			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Visual Reception	.32†	.42*	-.22	.30†	.42**f*	.02	.30†	-.09	-.11	.41*	.45**	-.17
Fine Motor	.26	.49**	-.06	.29†	.28	-.07	.28	-.32†	.16	.36*	.33	.01
Receptive Language	.35**f*	.56***	-.25	.32	.38**f*	-.13	.30†	-.03	.12	.44**	.54**	-.15
Expressive Language	.34**f*	.59***	-.01	.23	.37**f*	.11	.35**f*	-.17	.21	.41*	.50**	.13

Note: † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$; ƒ did not meet significance cutoff ($p = .0125$) when controlling for familywise error.

Table 5

Experiment 2: Means (standard deviations in parentheses) for duration: attention maintenance (proportion of available looking time, PALT), accuracy: intersensory matching (proportion of total looking time to synchronous event, PTLT), and speed: attention shifting (reaction time, RT) as a function of type of competition (low, high) and event (social, nonsocial).

		Social	Nonsocial	Overall
Duration: Attention Maintenance (PALT)	Low Competition	.69 (.19)	.67 (.19)	.69 (.14)
	High Competition	.45 (.20)	.43 (.18)	.45 (.15)
	Overall	.58 (.15)	.56 (.13)	.57 (.10)
Accuracy: Intersensory Matching (PTLT)	Low Competition	.53 (.11)	.55 (.10)**	.54 (.07)**
	High Competition	.50 (.12)	.53 (.13)	.51 (.08)
	Overall	.51 (.09)	.54 (.08)**	.53 (.06)**
Speed: Attention Shifting (RT)	Low Competition	.96 (.09)	1.02 (.17)	.98 (.09)
	High Competition	1.34 (.87)	1.46 (.65)	1.39 (.67)
	Overall	1.15 (.51)	1.24 (.39)	1.19 (.42)

Note: For PTLT, single sample *t*-tests; ** $p < .01$.

Table 6

Experiment 2: Correlations between the three indices of attention—for duration of looking (Dur), accuracy of intersensory matching (Acc), and speed of attention shifting (Spd)—as a function of type of competition (low, high) and event (social neutral, social positive, nonsocial).

	Low Competition Trials								
	Social			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Duration (Dur)	---	-.20	-.45**	---	.26†	-.09	---	.03	-.28
Accuracy (Acc)		---	.24		---	.13		---	.24
Speed (Spd)			---			---			---
	High Competition Trials								
	Social			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Duration (Dur)	---	.22	-.52***	---	.23	-.30**ƒ*	---	.31*	-.51***
Accuracy (Acc)		---	.07		---	-.39**		---	-.38**
Speed (Spd)			---			---			---
	Overall								
	Social			Nonsocial			Overall		
	Dur	Acc	Spd	Dur	Acc	Spd	Dur	Acc	Spd
Duration (Dur)	---	.28	-.48**	---	.39**	-.24	---	.35*	-.50***
Accuracy (Acc)		---	.06		---	-.57***		---	-.51***
Speed (Spd)			---			---			---

Note: * $p < .05$, ** $p < .01$, *** $p < .001$; *ƒ* did not meet significance cutoff ($p = .0167$) when controlling for familywise error.

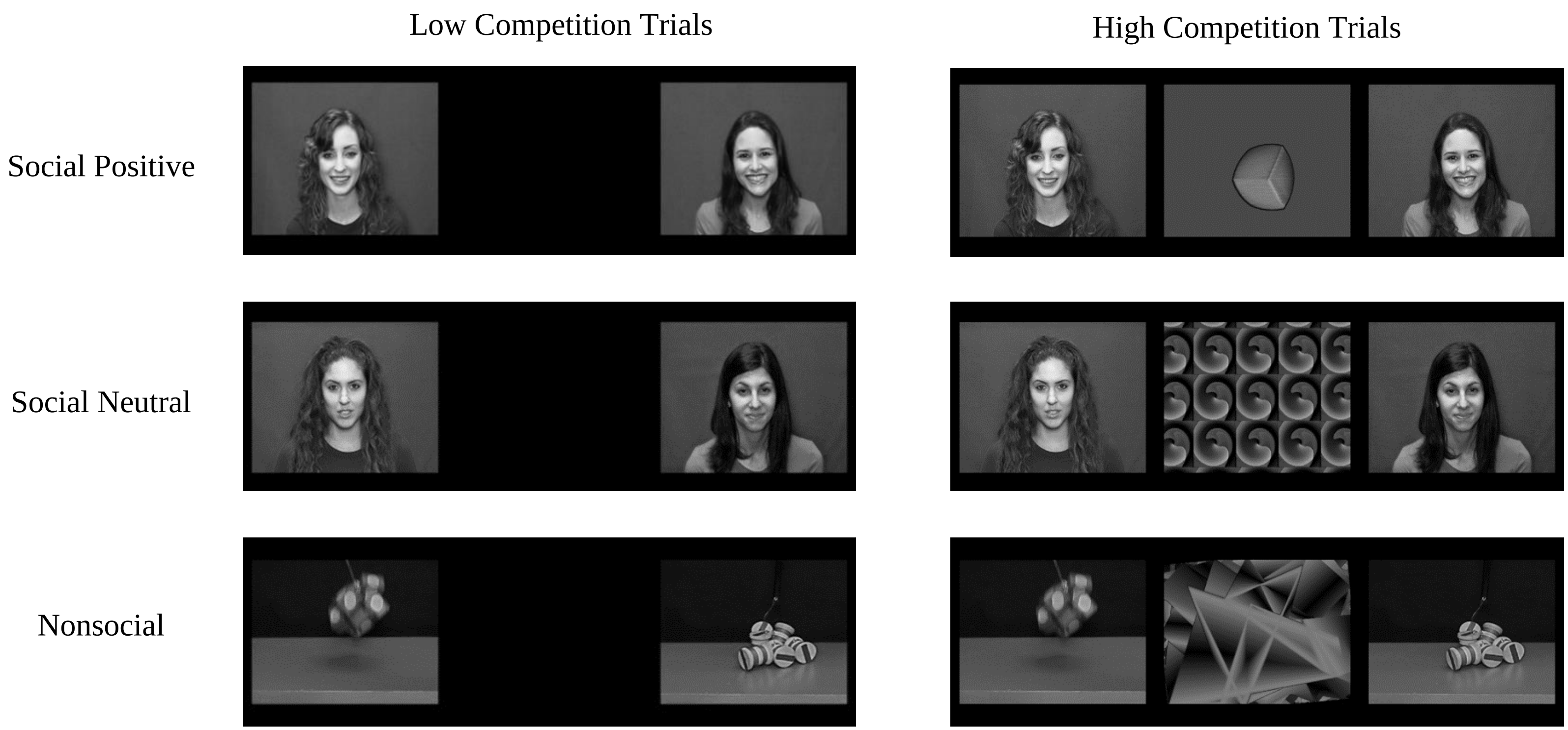

Figure 1. Experiment 1: Static images of the dynamic audiovisual events. Children were presented with a three second central stimulus (computerized geometric shape) followed by two side-by-side lateral events (social positive, social neutral, nonsocial), one of which was synchronous with its appropriate soundtrack. On low competition trials, the central stimulus was turned off during the lateral events, whereas on high competition trials, the central stimulus remained on during the lateral events.

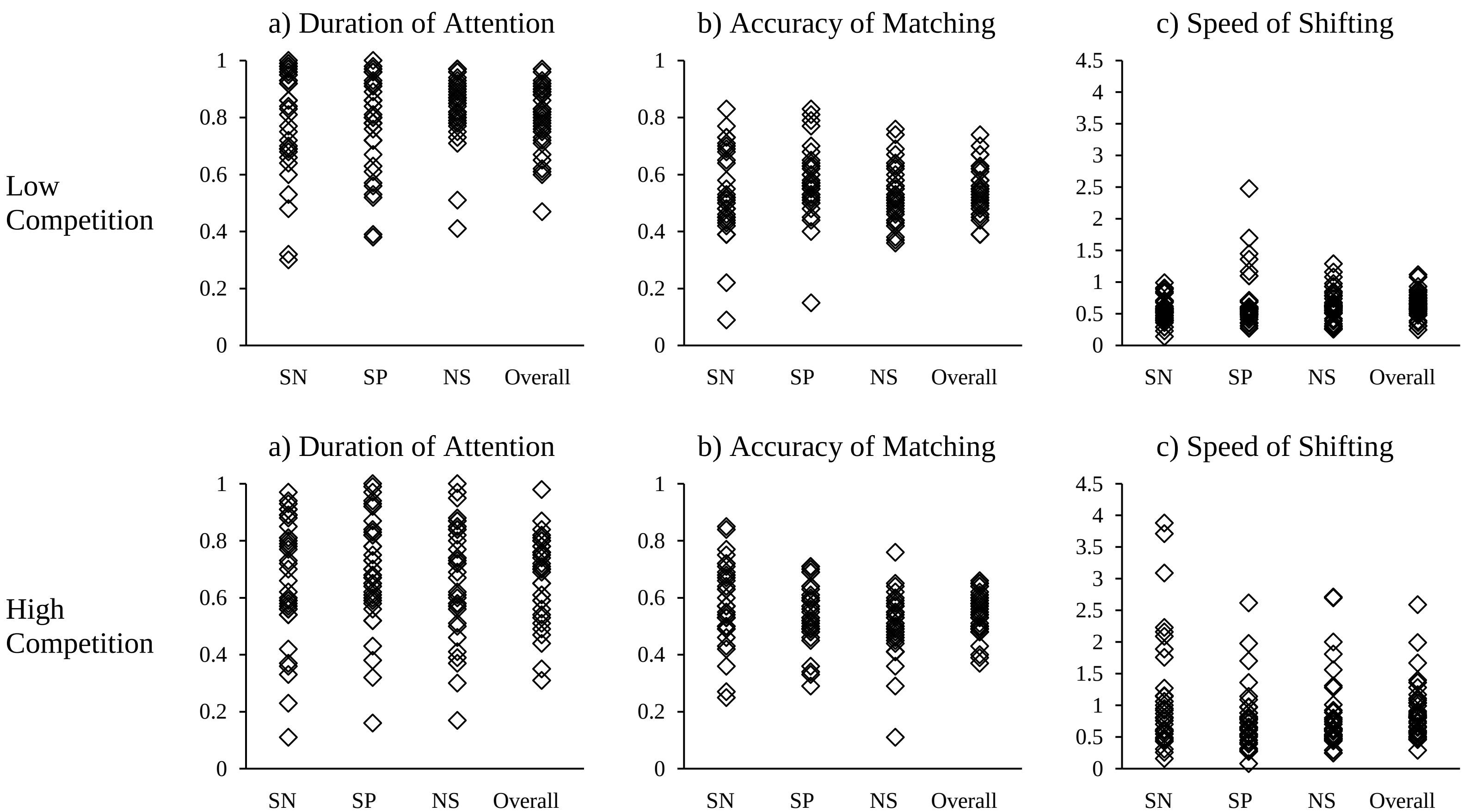

Figure 2. Experiment 1: Scatterplots for the three indices of attention—a) duration of attention maintenance (proportion of available looking time to the lateral events), b) accuracy of intersensory matching (proportion of total looking time to the synchronous event), c) speed of attention shifting (RT to shift to from the central stimulus to a lateral event)—as a function of competition (low, high) and event (social neutral: SN; social positive: SP; nonsocial: NS). Each point denotes the score of a single participant.

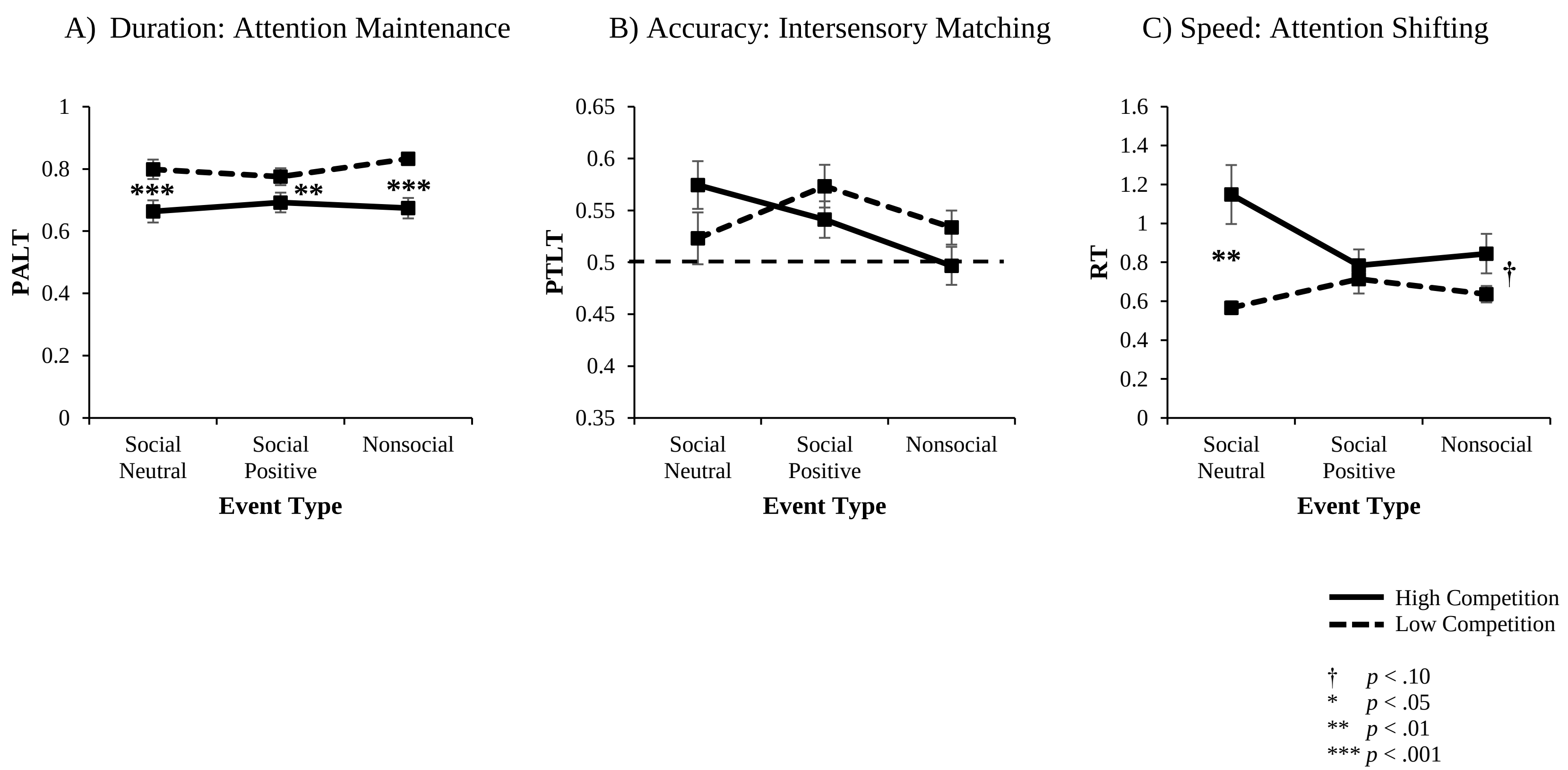

Figure 3. Experiment 1: Performance for each of the three indices of attention—a) duration: attention maintenance (proportion of available looking time, PALT), b) accuracy: intersensory matching (proportion of total looking time, PTLT), and c) speed: attention shifting (reaction time in seconds, RT)—as a function of type of competition (low, high) and event (social neutral, social positive, nonsocial). Error bars depict standard errors of the mean.

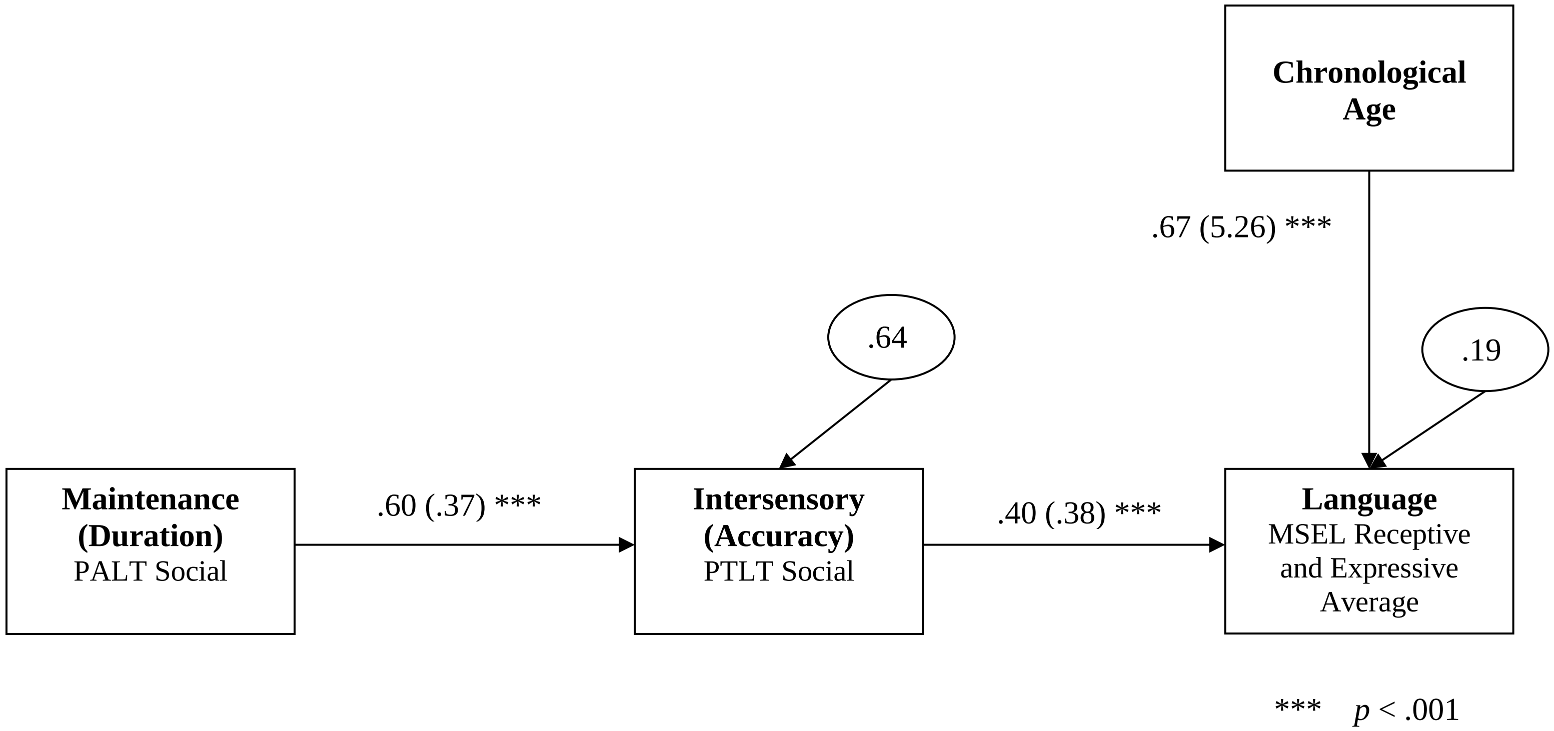

Figure 4. Experiment 1: Structural model depicting relations between attention to social events and language outcomes. Standardized regression coefficients are presented outside the parentheses and unstandardized coefficients are presented inside parentheses. The proportions of variance unaccounted for by predictor variables (error variance) are presented in circles above the outcome variables (Intersensory, Language).

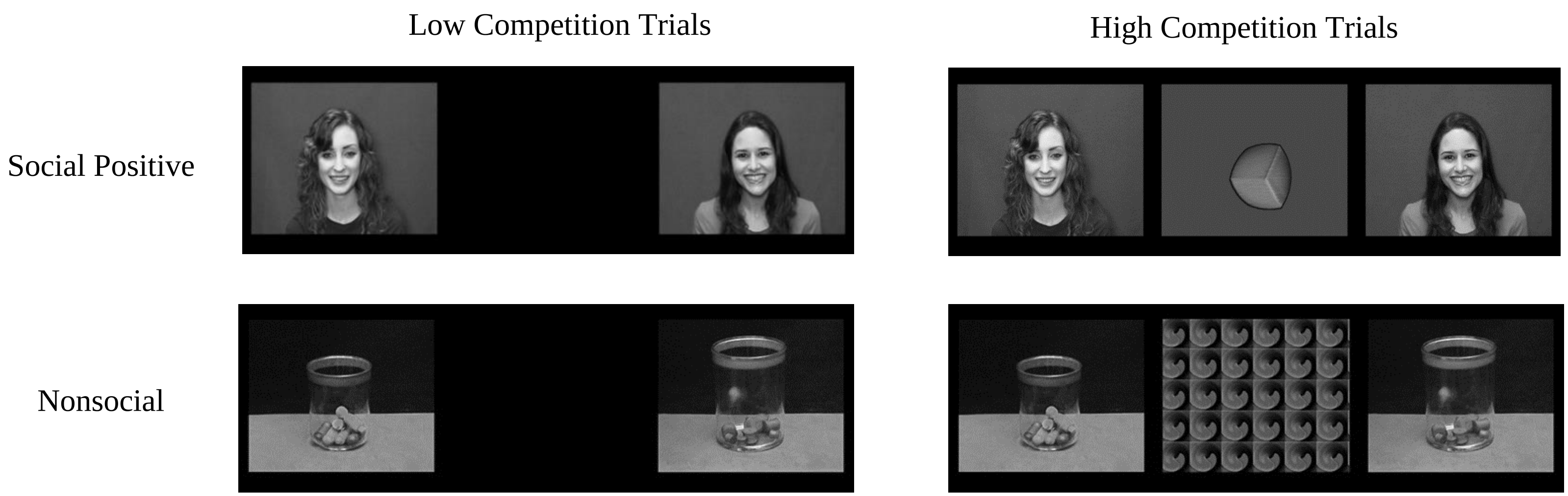

Figure 5. Experiment 2: Static images of the dynamic audiovisual events. Children were presented with a three second central stimulus (computerized geometric shape) followed by two side-by-side lateral events (social, nonsocial), one of which was synchronous with its appropriate soundtrack. On low competition trials, the central stimulus was turned off during the lateral events, whereas on high competition trials, the central stimulus remained on during the lateral events.

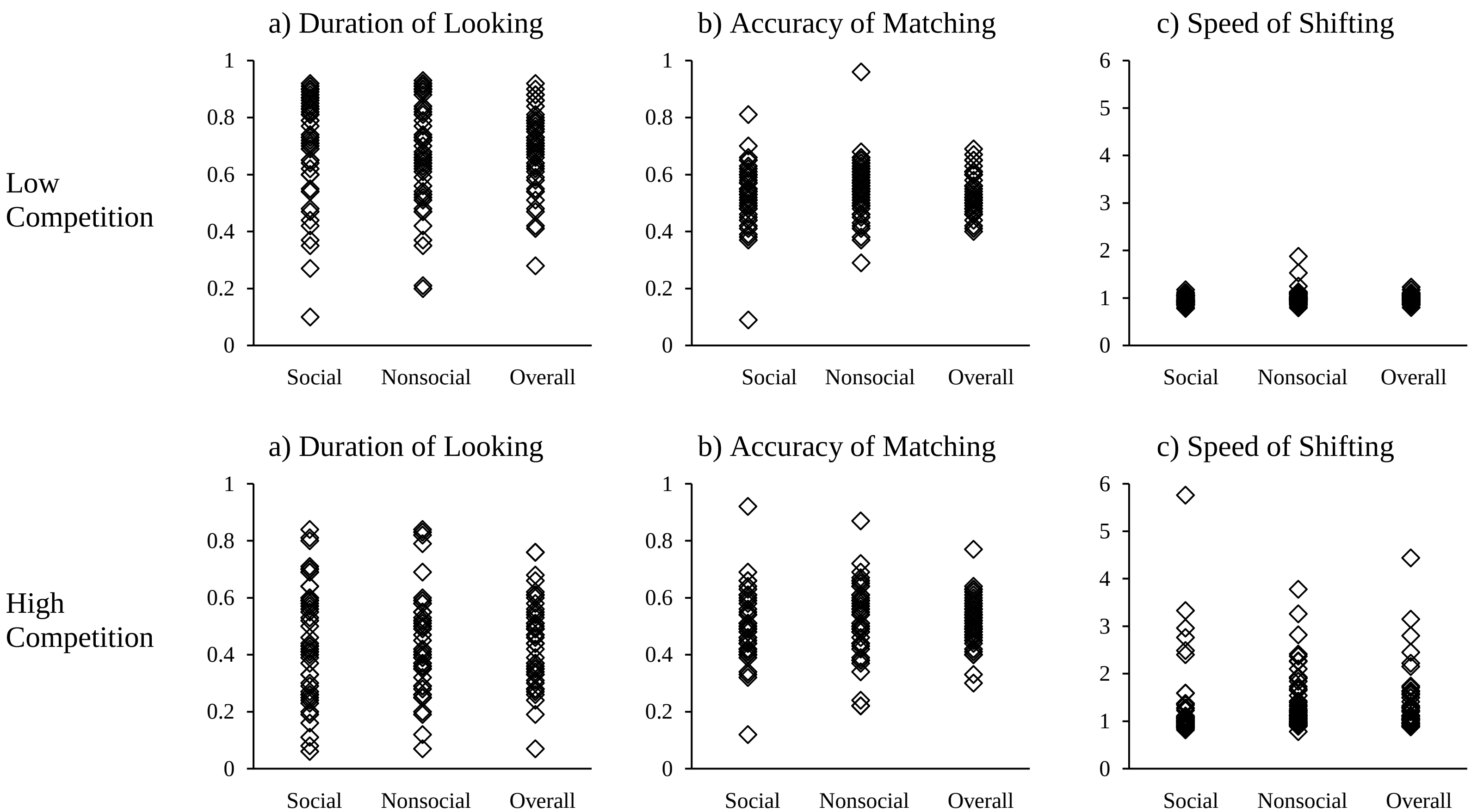

Figure 6. Experiment 2: Scatterplots for the three indices of attention—a) attention maintenance (duration; proportion of available looking time to the lateral events), b) intersensory matching (accuracy; proportion of total looking time to the synchronous event), c) attention shifting (speed, RT to shift to from the central stimulus to a lateral event)—as a function of competition (low, high) and event (social, nonsocial). Each point denotes the score of a single participant.

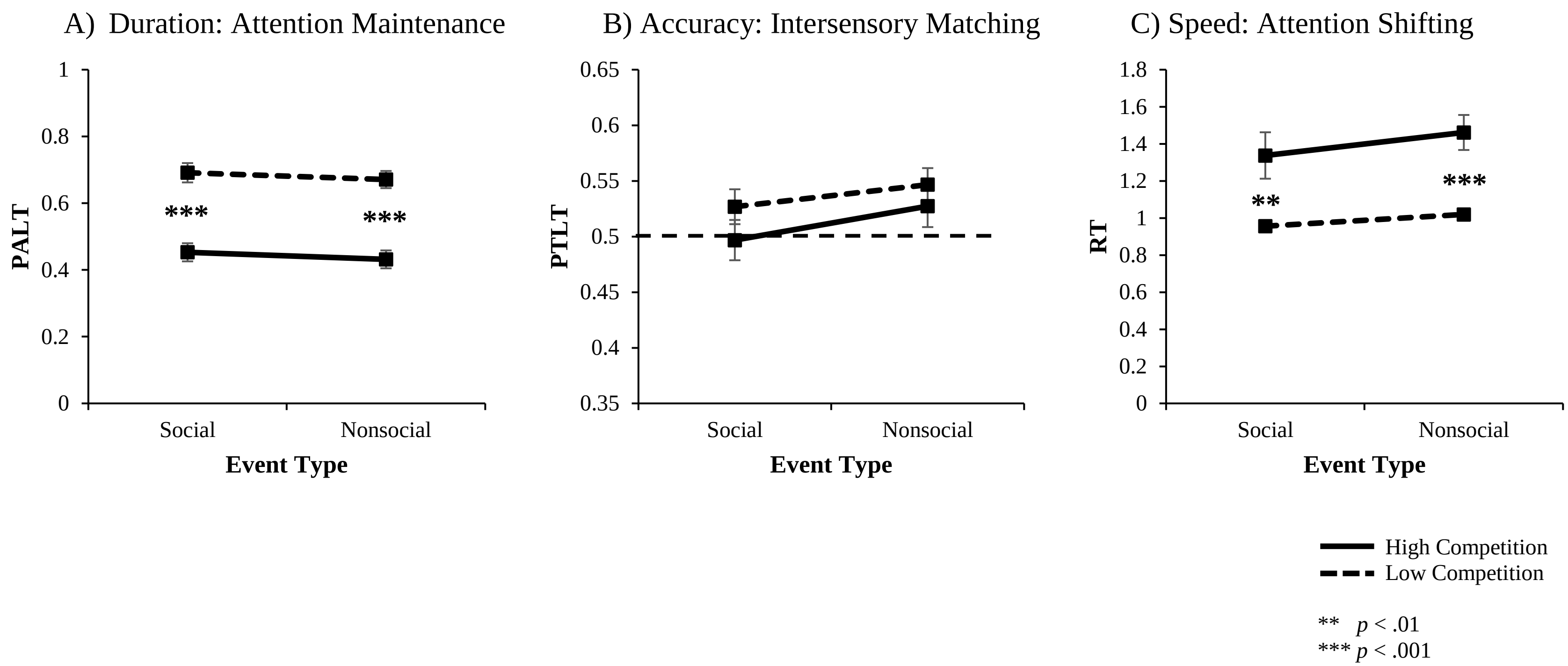

Figure 7. Experiment 2: Performance for each of the three indices of attention— a) duration: attention maintenance (proportion of available looking time, PALT), b) accuracy: intersensory matching (proportion of total looking time to the sound synchronous event, PTLT), and c) speed: attention shifting (reaction time in seconds, RT)—as a function of type of competition (low, high) and event (social, nonsocial). Error bars depict standard errors of the mean.