Characterizing Multisensory Attention in Early Development: Individual Differences, Trajectories, and Relations with Outcomes

James Torrence Todd & Lorraine E. Bahrick

Introduction
Multisensory attention skills (MASks; detecting audiovisual synchrony, shifting and maintaining attention to unitary audiovisual events) are foundations for language development.1

• Until recently there were no measures appropriate for assessing individual differences in MASks in young children. Thus, developmental trajectories of these skills and links with outcomes remain poorly understood.

• We developed the first individual difference measures appropriate for young children for assessing MASks in the context of audiovisual social and nonsocial events.

• Using these measures, we found that intersensory processing predicts prematurity skills (letter names & sounds knowledge) in 4-7 year-olds.2

• Here, we characterize developmental trajectories of MASks, and relations between MASks and language outcomes.

Methods
• Participants. N=104 infants (53 F) tested longitudinally at 3, 6, 12, 18, 24, and 36 months of age.

• Multisensory Attention Assessment Protocol (MAAP). Assesses three MASks: duration of attention maintenance, speed of shifting, accuracy of intersensory matching to audiovisual events in the presence and absence of competing stimuli (Fig 1). To see an example video of the MAAP, click here.

• Intersensory Processing Efficiency Protocol (IPEP). A more fine-grained assessment of speed and accuracy to selectively attend to the sound synchronous audiovisual (target) event amidst five distractors (Fig 2). To see an example video of the IPEP, click here.

• Language Outcome Measures:
  • MB-CDI Words & Gestures: 12 and 18 months
  • PPVT IV: 36 months
  • EVT-2: 36 months

• Results
  • Improved MASks Across Age: For both social and nonsocial events, across age, we found longer attention maintenance (ps < .001; Fig 3A) and faster speed of shifting (ps < .001; Fig 3B) in the presence of competing stimulation on the MAAP, and greater intersensory matching (proportion of total looking time to the sound-synchronous event; ps < .02; Fig 3C) on the IPEP.

  • Intersensory Accuracy for Social Events Predicts Language Outcomes.

  • 12-month intersensory matching (MAAP) predicts 18-month CDI expressive vocabulary, r = .34, p < .03, and 36-month PPVT scores, r = .34, p = .01.

  • 6-month intersensory matching (IPEP) predicts 36-month PPVT, r = .26, p = .04, and EVT scores, r = .26, p = .04.

  • 6-month accuracy in finding the audiovisual target (IPEP) predicts 36-month PPVT, r = .36, p < .003, and EVT scores, r = .26, p = .04.

  • 12-month Intersensory Matching Accuracy of Social Events is a Mediator between Basic MASks and Language Outcomes. SEM analyses using social events in the presence of competing stimulation revealed that, at 12 months, accuracy of matching faces and voices on the MAAP (Intersensory Matching Accuracy) was the only MASks to directly predict Receptive Vocabulary (CDI words understood) and was a part of a 4-part meditational chain (Fig 4).

  • Speed of disengaging from competing stimulation (Disengage Speed) predicts duration of looking to faces (Attention Maintenance), which in turn predicts Intersensory Matching Accuracy, which in turn predicts Receptive Vocabulary.

  • Accuracy of matching faces and voices (Intersensory Matching Speed) predicts Intersensory Matching Accuracy, which in turn predicts Receptive Vocabulary.

  • Together, MASks predict 28% of the variance in Receptive Vocabulary.

Conclusions:
• Findings demonstrate change in MASks across age, with improvements in attention maintenance, shifting, and intersensory matching.

• Greater intersensory accuracy in infancy (6, 12 months) predicts language outcomes in toddlerhood (18, 36 months).

• Finally, our model of 12-month MASks to social events revealed novel causal pathways to language. Intersensory accuracy of face-voice matching mediates the relation between basic MASks (disengagement speed, attention maintenance, intersensory speed of face-voice matching) and receptive vocabulary size.