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 preliteracy skills (letter names & sounds knowledge) in 4- to 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[3]).** Assesses three MASks: duration of attention maintenance, speed of shifting, accuracy of intersensory matching to audiovisual events in the presence and absence of competing stimulation (Fig 1). To see an example video of the MAAP, click here.
- **Intersensory Processing Efficiency Protocol (IPEP[4]).** 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

References

1. 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).
2. Bahrick, L. E., McNew, M. E., Todd, J. T., Martinez, J., Mira, S., Cheatham-Johnson, R., & Hart, K. C. (2017). *Individual differences in intersensory processing predict pre-literacy skills in young children*. Poster presented at the meeting of the Society for Research in Child Development, Austin, TX.
3. Bahrick, L. E., Todd, J. T., & Soska, K. C. (2018). The Multisensory Attention Assessment Protocol (MAAP): Characterizing individual differences in multisensory attention skills in infants and children and relations with language and cognition. *Developmental Psychology*, *54*(12), 2207–2225.
4. Bahrick, L. E., Soska, K. C., & Todd, J. T. (2018). Assessing individual differences in the speed and accuracy of intersensory processing in young children: The intersensory processing efficiency protocol. *Developmental Psychology*, *54*(12), 2226–2239.

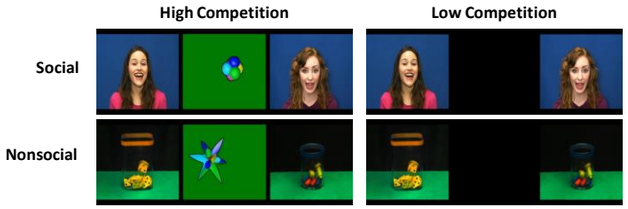

Figure 1. Static images of the dynamic social and nonsocial events from the MAAP. For both social and nonsocial trial blocks, one of the lateral events (social, nonsocial) was synchronous with its appropriate soundtrack. On low competition trials, the central stimulus was turned off during the lateral events, whereas on high competition trials, the central stimulus remained on during the lateral events.

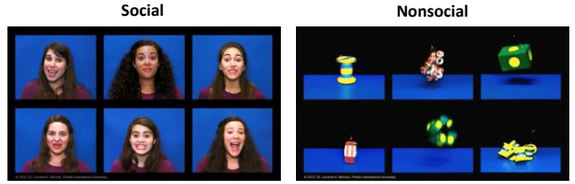

Figure 2. Static images of the dynamic social and nonsocial events from the IPEP. For both social and nonsocial trials blocks, on all trials, all six events (women, objects) were shown moving, but on each trial the movements of a different woman or object was synchronized with the accompanying, natural soundtrack.

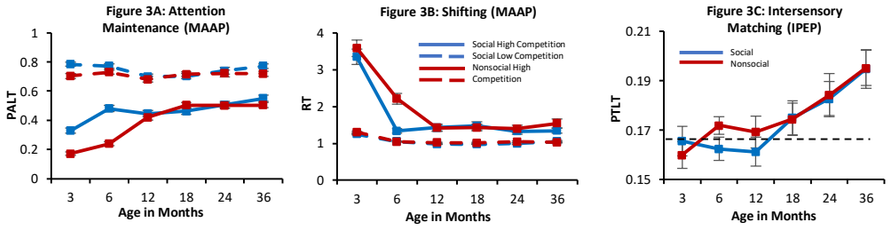

Figure 3. For the MAAP, A) attention maintenance (proportion of available looking time, PALT) and B) shifting (reaction time, RT). For the IPEP, C) intersensory matching (proportion of total looking time to the sound synchronous target event; PTLT).

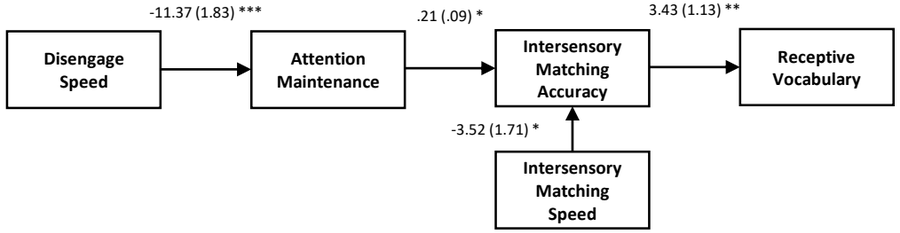

Figure 4. A SEM depicting relations between MASks to social events and language outcomes. Unstandardized regression coefficients appear above each pathway (*SE*s in parentheses). *Note.* * $p < .05$, ** $p < .01$, *** $p < .001$.

Results

- **Improved MASks Across Age:** For both social and nonsocial events, across age, we found longer attention maintenance (*p*s < .001; Fig 3A) and faster speed of shifting (*p*s < .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; *p*s < .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 MASk to directly predict Receptive Vocabulary (CDI words understood) and was a part of a 4-part mediational 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.
  - Also, speed 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 casual pathways to language. Intersensory accuracy of face-voice matching mediates the relation between more basic MASks (disengagement speed, attention maintenance, intersensory speed of face-voice matching) and receptive vocabulary size.

Presented July 6-9, 2020, at the International Congress on Infant Studies Virtual Conference. This research was supported by the NIH Grants 2R01-HD053776-11 and R01-HD094803 awarded to Lorraine E. Bahrick.