Multisensory Integration in Autism-Implications for Treatment

Kaitlyn Clarke, MS, CCC/SLP, Doctoral Candidate
Department of Communication Sciences and Disorders
The Pennsylvania State University

What is multisensory integration and why does it matter?

- Multisensory integration (MSI) is defined as the “convergence of inputs from different sensory modalities onto individual neurons...” [Merideth & Stein, 1986, p.640].
- From birth, newborns are typically exposed to a multitude of sensory experiences, including auditory, visual, and somatosensory inputs [Streri, 2012].
- This ability to perceive various modes of sensory input is crucial in human development, specifically the ability to integrate multiple forms of information that occur at one time [Lewkowicz, 1994].
- Crucial in the regulation of sensory perception during activities such as learning language and communicative interactions [Hillcock, Powers & Wallace, 2011].
- Allows us to make sense of the many streams of available inputs at any one moment [Allen, Stevenson, Wallace & Wenger, 2015].

How and where does MSI occur?

- One of the primary structures responsible for the processing of multisensory information is the superior colliculus (SC).
- Integrates multiple modalities and differentiates between unimodal and multimodal streams of input.
- Located at the juncture of the temporal, occipital, and parietal cortices.
- Secondary structure is anterior ectosylvian sulcus (AES) which contains multisensory neurons.
- We know quite a bit about adult structures and functioning related to unimodal and multimodal sensory input, however, less is known about infant MSI development.
- Currently, we look to the developmental progression in monkeys to investigate this development.

MSI in newborns

- Unsurprisingly, newborns do not have the same integration skills as adults.
- There are two primary theories of how newborns develop their ability to integrate multisensory information: the integration view and the differentiation view.
  - Integration view: unimodal sensory perception is separate and the development of multisensory integration occurs over time through experience of repeated exposure and further neural mapping [Lewkowicz, 1994].
  - Differentiation view: infants are born with the ability to process multisensory information that requires experience to parse single modalities out over time [Lewkowicz, 1994].
What do we know from monkeys?

SC is initially exclusive to unimodal input—similar to the integration view of multisensory input—specifically, somatosensory information.

Next to develop is unimodal auditory input, then multimodal auditory-somatosensory.

Finally, unimodal visual perception, followed by multimodal combinations including vision.

Sours (2017) found a similar trajectory in humans for neural structures being present at birth but being activated in a sequence of unimodal and multimodal responses.

It can be hypothesized that the sequence of sensory responses are likely due to input in the environment.

Temporal Binding

WHAT?

Multiple modes of input are received by individual neurons and converged during what is known as the temporal binding window (TBW).

- The temporal binding window is the maximum amount of time between stimuli that still allows humans to pair information that occurs in close temporal proximity, allowing the perception of integrated auditory and visual stimuli (Hilcock, Powers & Wallace, 2011).

- The TBW in newborns is significantly greater than that of an adult.
  - allows the individual to pair stimuli through repeated exposures and perceived regularities in their environment (Lewkowicz & Flom, 2013).
  - It is through this TBW and pairing of auditory and visual sensory information that infants receive the necessary input and stimuli for the development of speech perception (Lewkowicz & Flom, 2013).

Temporal Binding

WHAT?

- In infancy, the TBW for auditory and visual information ranges from 350 to 400ms, narrowing to approximately 300ms by age four and continuing to decrease to 80-187ms by adulthood (Lewkowicz & Flom, 2013).
  - The narrowing of the TBW is due to repeated exposure and statistical learning of regularities in auditory and visual input patterns, however, the process of the narrowing of the TBW is not rapid (Lewkowicz & Flom, 2013).

- In early childhood development, the TBW is crucial in early language and cognitive development (Lewkowicz & Flom, 2013).
  - pairing of both auditory and visual information allows us to expand our understanding of concepts—such as related noises to actions (e.g. oral motor movements and speech sounds) (Altieri et al., 2015).
So, how about in autism?

- At the brain level, individuals with ASD are found to have similar brain activations to auditory input that is both social and non-social.
- Behaviorally, individuals with ASD are noted to display decreased accuracy for identification of auditory information—especially with competing background noise.
- Varied reports for brain activations (ERPs) during visual processing tasks were noted in individuals with ASD—as well as variable reports of behavioral performance for visual scanning.
- One hypothesis of the variability in visual perception and processing in ASD is related to the saliency of the information presented.
- Ok, but what about MSI?

MSI in autism

- Typically developing two-month-old infants show a preference for the synchrony—or togetherness—of multiple streams of input during infant-directed speech.
- Infants who later are diagnosed with ASD do not display the same preference (Patten, Watson & Baranek, 2014).
- The implications of this lack of preference for synchrony for individuals with ASD may cascade to later skills, such as word learning, general language skills, and joint attention (Patten, Watson & Baranek, 2014).
- Around school age (i.e. seven to eight years old), individuals with ASD still display differences in the integration of multiple streams of sensory information (i.e. auditory and visual).
- Participants with ASD matched their typically developing peers in multimodal perception by the age of 16.4 (Taylor, Isaac & Milne, 2010).
- Why does the synchrony of input matter?

Why does the synchrony of input matter?

- To combine multiple streams of sensory input, individuals rely on the synchrony of input within the temporal binding window (TBW) (Hillcock, Powers & Wallace, 2011).
- In ASD, the TBW is larger than that of typical development—limiting the benefit on synchrony of dual inputs (Stevenson et al., 2015).
- In TD, the benefit of the synchronicity of multiple streams of input is related to temporal processing within the TBW (Martinez-Sanchis, 2014).
- Detection of temporal order and synchronicity of multiple modes of input has been found to be predictive of later language skills (Patten, Watson & Baranek, 2014) and has implications for speech perception in individuals with ASD (Stevenson et al., 2015).
- In a highly social, speech-based task, Righi and colleagues (2018) determined that individuals with ASD spent less time looking at the screen overall, but also less preference for the videos that were more synchronous as compared to their typically developing peers (Righi et al., 2018).
- It was hypothesized that individuals with ASD were less likely to attend to synchronous over asynchronous dual inputs because of the limited ability to detect the temporal regularities of these inputs (Righi et al., 2018).
- Further cascading effects of deviations in the ability to detect synchronicity include differences in attention, localization, and global processing (Bahrick & Todd, 2012).
- It has been hypothesized that temporal processing for individuals with ASD has been disrupted, specifically related to deviations in the frontal cortex and limitations in the motor PFC (Martinez-Sanchis, 2014).
- The deviations in the frontal and PFC are related to difficulties in Complex information processing (i.e. more social and linguistic context), as well as inabilities to pair multiple streams of input within an appropriate time interval (Martinez-Sanchis, 2014).
- Additionally, Righi and colleagues (2018) have proposed that there are underlying differences in the ability to detect temporal (i.e. timing) variations in auditory perception in individuals with ASD that drive difficulties in multimodal perception and integration.
Because of difficulties in the perception of multiple streams of sensory information in individuals with ASD, it has been hypothesized that there is a tendency for “piecemealing” information together, instead of global processing.

For low-level procedures, individuals with ASD are reported to experience decreased effects from the widening gap between sensory inputs (i.e. TBW) and varied ERPs during the processing of these stimuli (Marco et al., 2011).

Stevenson and colleagues (2018) looked at higher-level multimodal input through the effects of varying speech sound (auditory) and speech oral-movement (visual) synchrony to measure accuracy of speech perception.

Findings within this study revealed that the ability to pair multiple forms of sensory input (i.e. multimodal integration) has direct implications on speech perception for individuals with ASD, likely due to the avoidance of globally processing speech through the multimodal integration of both auditory and visual stimuli (Stevenson et al., 2018).

In addition to a reliance on bottom up processing, there is also the difference in TBW length related to the pairing of multimodal input.

Because of this change in TBW, individuals with ASD may continue to pair sensory inputs that are unrelated because of their proximity within this timeframe and “piecemeal” their processing of sensory situations (Bahrick & Todd, 2014).

Losing the forest for the trees?
- Or losing the multimodal input for the unimodal input?

Based on the review of typical development and development in individuals with ASD, we see that there are underlying differences in the benefits of multisensory input.

Thinking about AAC instruction, we know that multi-modal input has been shown to be effective.

Aided AAC input has been shown to improve outcomes for individuals who require AAC regardless of diagnosis.

SLPs are using multimodal input, however, are not necessarily attending to the synchronicity of input provided (Clarke & Williams, 2019-under review).

Of nine participants, only one stated they varied the timing of multimodal input.

Most participants cited “aided language stimulation”, but did not follow the specific guidelines of Goosens’ model.
• My hypothesis is that although multimodal input is effective, is it the most efficient?
  • What if we can make effective therapies more efficient?

Target Pair
Auditory & Visual
Temporal Binding Window
Extraneous auditory input
Non-salient info
Background noise
Visual crowding

Future Research
• How can I determine if the synchronicity of inputs is an agent of change?
  • Single case, adapted alternating treatment design
  • Comparing synchronous vs. asynchronous dual inputs during instruction across matched vocabulary sets
  • Measuring accuracy of response, latency of response, and rate of acquisition

Treatment Implications (cont.)
• What about unimodal input?
  • Vision
    • Many AAC systems rely on visual representations for vocabulary, including graphic symbols
    • Individuals with ASD show greater benefits from and preference for salient visual stimuli
    • What about graphic symbols? Are these salient?
    • Would photographs be more salient for individuals with ASD?
      • Inclusion of salient people/place/term during instruction/use of AAC supports
  • Auditory
    • Many AAC supports provide either digitized or synthesized output
    • What about the synchrony and timing of synthesized output? Digitized?
    • Digitized output may be more salient, but can we better manipulate the timing of synthesized output?
      • Limit background noise with the use of synthesized output?
Discussion

- Thoughts on the presented model
  - Match to current research
  - Applications from lab research (cognitive psychology) to intervention research
  - Application to clinical practice
- Thoughts on use of single case research design
- Recommendations for changes