Reconciling the Cognitive and Neural Architecture of Multisensory Processing in the Autistic Brain

Michael J. Crosse[1], **Cristiano Cuppini**[2], **John J. Foxe**[1,3] **and Sophie Molholm**[1,3]

[1]Albert Einstein College of Medicine, [2]University of Bologna, [3]University of Rochester

Abstract

Individuals with autism spectrum disorder (ASD) are often impaired in their ability to process multisensory information, which may contribute to some of the social and communicative deficits that are so prevalent in this population. Despite the increasing empirical evidence, our understanding of the cognitive and neural architecture that underpins the development of this ability is lacking. Recent work has shown that response times (RTs) to multisensory stimuli can be modeled using an extension of the classic race model framework [4]. Here, we demonstrate that the basic race model architecture is predictive of multisensory benefits in adult participants with ASD ($R^2_{(23)} = 0.25$, $p = 0.017$), but not in children with ASD ($R^2_{(34)} = 0.012$, $p = 0.531$). We find that by modeling an alternative processing strategy whereby multisensory RTs are determined by the preceding modality is a better predictor of behavior in autistic children than the race model ($R^2_{(34)} = 0.14$, $p = 0.029$). To understand the neural basis of this cognitive framework, we developed a neuro-computational model presenting two levels of multisensory interactions: inhibition and cooperation. This model is based on a previous neural network implementation used to explain acquisition of multisensory integrative abilities at the neuronal level in the superior colliculus [2, 1]. The model suggests that in the absence of substantial experience with multisensory stimuli (i.e., at an early stage of development), the main interaction between sensory modalities is competition, with the preceding modality the stronger competitor. At a later stage of development, experience with crossmodal events appears to promote positive interactions between modalities, thus enhancing behavior. These findings link our cognitive framework to a plausible neural implementation and provide an explanation for the multisensory deficits commonly reported in children with ASD.

Modeling the cognitive architecture

To test whether multisensory RTs in children with ASD comply with a race model architecture or an alternative processing strategy, we tested the ability of several models to predict empirical multisensory benefits. We proposed two alternative strategies to the race model framework: 1) multisensory RTs are biased towards the modality of the previous trial, irrespective of which modality wins the race; 2) multisensory RTs are biased towards a specific modality, irrespective of which modality wins the race. We parametrically varied the contribution of the race model $P_{A\cup V}(n,t)$ and each bias model as follows:

$$\text{Model 1a} = \frac{(1-k)\Sigma_n P_{A\cup V}(n,t) + k(P_A(AV,t) + P_A(A,t) + P_V(V,t))}{3} \tag{1}$$

where k is the contribution of each model ranging from 0 to 1 in increments of 0.25 and n is the preceding modality. Because RTs preceded by AV trials could also be biased towards the V modality, we define separately Model 1b by replacing $P_A(AV,t)$ with $P_V(AV,t)$ in Model 1a. Model 2a and 2b were computed in much the same way as equation 1, except the RT bias was determined by either the A or the V modality respectively.

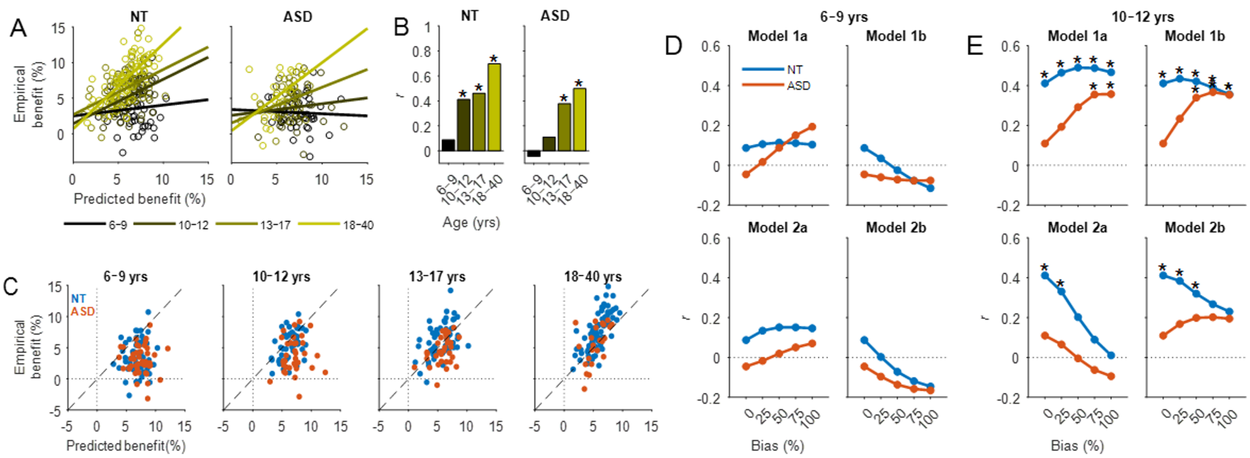

Figure 1: A, Predicted and empirical benefits in neurotypical and autistic participants with linear regression fits. B, Correlation coefficients of regression fits in panel A represent race model performance. C, Predicted and empirical multisensory benefits by age group. D–E, Performance of alternative models as a function of their contribution to the race model for 6–9 year olds and 10–12 year olds.

Neural model implementation

A neural model was developed to focus on the key dynamics believed to underlie the development of multisensory integration. The model contained artificial nodes (units) grouped into topographically-organized regions representing different circuit components. Each of the tectopetal regions were functionally divided into two subregions: visual and auditory. The tectopetal projections from these subregions were excitatory and topographic. Interactions between each of these modality-specific subregions implemented a default computation that approximated a winner-takes-all (WTA) competition [3]. Effectively, only the strongest input signal survived this competition with attrition to influence the target unit(s). For simplicity, competition to control the target unit was implemented via direct inhibitory connections between units in different input regions. In addition, subregions extended a set of non-competitive tectopetal projections that were excitatory, strictly topographic, modifiable by crossmodal experience, and not influenced by the WTA competition. In this way, they were functionally distinct from the competitive projections. Changes in these non-competitive projections, along with changes in the inhibitory balance between inputs, were hypothesized to account for the acquisition of multisensory enhancement capabilities during normal development.

Conclusions

The ability of the race model to predict empirical multisensory benefits shifts from over-predicting it in younger participants, to under-predicting it in adults (Fig. 1C). This fits with the predictions of our neural model that suggest that early in development, interactions between sensory modalities reflects a competition, inhibiting the processing of the other modality and resulting in sub-optimal behavior. It's possible that this competition is the cause of younger participants not exhibiting even basic statistical facilitation. This is further supported by the minimal amount of variance explained by a race model in younger participants compared to adults (Fig. 1A,B). A race architecture represents the most efficient processing strategy given the task demands placed on the participant. These data suggest an alternative, and therefore less efficient, processing strategy is being implemented by children. Our data suggest that the strategy used to couple multisensory decision processes is biased towards the preceding modality, and not the faster modality. In line with our neural model predictions, it is conceivable that if both signals competed for resources, the modality that preceded the current event would be the stronger competitor as more attentional resources would have been allocated therein.

Acknowledgments
We thank J. Bates and N. Russo for clinical and cognitive phenotyping, A. Brandwein for leading the original EEG studies for which these data were collected, D. Horsthuis, E. Tavern and D. DeMaio for technical assistance. This work was supported by NIH grants RO1MH085322 and U54HD090260.

References

[1] C. Cuppini, B. E. Stein, and B. A. Rowland. Development of the mechanisms governing midbrain multisensory integration. *Journal of Neuroscience*, pages 2631–17, 2018.

[2] C. Cuppini, B. E. Stein, B. A. Rowland, E. Magosso, and M. Ursino. A computational study of multisensory maturation in the superior colliculus (sc). *Experimental brain research*, 213(2-3):341–349, 2011.

[3] S. Grossberg. *Contour enhancement, short term memory, and constancies in reverberating neural networks*, pages 332–378. Springer, 1982.

[4] T. U. Otto and P. Mamassian. Noise and correlations in parallel perceptual decision making. *Current Biology*, 22(15):1391–1396, 2012.