BAYES, BRAINS & BABIES:

Electrophysiology and Mathematics of Infant Holistic Processing and Selective Inhibition

Matthew F. Singh
Greg D. Reynolds
Department of Psychology

Introduction

Holistic, (global) processing refers to the integration of local stimulus features into a larger perceptual form. While global precedence is the norm for adults [1], infants' perception depends greatly upon individual differences [2].

The most common measure of these differences is fixation duration to a stimulus which is also negatively correlated to cognitive functions throughout childhood [3]. One explanation is that infants exhibiting shorter looks may be utilizing the more efficient global-local processing sequence, while "long lookers" use the opposite strategy [2]. Previous studies in our lab have supported this [4].

However, we contend that these differences actually stem from dopaminergic (DA) systems such as in dorsolateral prefrontal cortex (dlPFC) and striatum as opposed to basic perceptual intake. A major role of DA in dlPFC is the integration of working memory (WM) and inhibition [9]. While long looking infants demonstrate lower recognition memory [10], this study will be the first to test selective inhibition in accordance with looking-behavior. Moreover, we have developed a model for these effects using Markov Chains and Rescorla-Wagner (RW) learning algorithms which model DA striatal effects [5].

Combining these approaches, we predict that saccade latency can be predicted as a stochastic linear differential equation with parameters related to inhibition and looking behavior. In the current stage we present evidence that infants differ only in learning hierarchical associations, but not temporally proximal cues. As such, our design demonstrates selectively decreased performance by short lookers in a task that actually rewards impairment in holistic processing.

Procedure

Participants consisted of 6-month infants (n=8)

Similar to [5] participants were seated on a parent's lap 55 cm away from a 27" color LCD monitor. A 128 channel EGI EEG system was used for data collection including a 124 geodesic net and 2 EOG electrodes coated with SignaCreme. E-Prime 2.0 was used.

Auditory stimuli consisted of 4 two-tone sequences each lasting a total of 500ms, referred to here on as AX (Target), XA, BY, CZ to indicate shared components.

Participants were presented with a female face until breaking fixation after 20s of accumulated looking. Participants were then presented with a randomly selected auditory sequence followed by a 300 ms inter-stimulus interval (ISI) before the presentation of the reinforcer. The AX sequence signaled appearance on one side while all others signaled the opposite side of the screen. The AX sequence occurred 3 times as often as the others. Data was collected until the participants became bored or fussy (M=70 trials, SD=17).

EOG analyses were performed using NetStation software. "Hits" and "misses" were defined as either correct or incorrect anticipatory looks for the visual reinforcer. Latency was filtered for a 100-500ms prosaccade band pass. Statistical analyses were performed using the JMP 10 interface for SAS.

Behavior

For behavioral measures, the ROC (discrimination) functions for standard stimuli were positively correlated with peak look (r (1, 6) =6,732, p=.039), but not for the target.

Latency

To decompose the factors of peak look, regression analyses found strong relations with the difference in ROC between standard as target as well as the differences in ROC latencies: total (r(2,115)=.795, p<.0001) and (t=13.75, p<.0001; t=-2.61, p=.01) for behavior and latency respectively.

In testing our computational model of a first order, linear stochastic differential equation, multiple regression analyses were performed in ROC latency differences. Model comparisons supported a linear differential equation maximized with ROC difference and looker-type as predictors: total (F (2, 64) =7.0435, p=.0017) and (t=-2.93, p=.0047; t=-3.75, p=.0004) for ROC behavior and looker type respectively. Corrections for multiple comparisons did not affect significance.

Computation

The standard RW equation [6] states that changes in association between a stimulus and reward is proportional to error:

$\Delta V = \alpha\beta(\lambda - \Sigma V)$ [Here $\alpha\beta$ denote cue/reward salience]

Here λ is the true outcome and V is the prediction. We can recreate a continuous time version with stochastic differential equations. Combining this with Bayes' rule we obtain our novel Markov Matrix for weight adjustment:

$$M := \begin{pmatrix} (1-\delta) + \delta\widehat{A}_i & \cdots & \delta(\widehat{A}_m) \\ \delta(\widehat{A}_i) & & \\ \vdots & \ddots & \vdots \\ & & \delta(\widehat{A}_m) \\ \delta(\widehat{A}_i) & \cdots & (1-\delta) + \delta\widehat{A}_m \end{pmatrix}$$

Here δ is a learning parameter and the "A" terms reflect conditional probabilities. With hierarchical neural networks we can further partition variances to configure synaptic weight vectors in neural networks. Using the generated weights after an input we can describe neuron activation with a modified general state equation [7]:

$\widehat{x}_i = \widehat{x}W \quad \frac{d\widehat{x}}{dt} = -\alpha\widehat{x} + Tr(\widehat{x}W)$

Here F denotes a transfer function between neuron input and output. Due to threshold linearity [8] we used the ramp function.

Simulation

Simulations were performed using a custom coded network for MATLAB 2014b for a fully crossed AX-CPT (see computation)

Our first simulation demonstrates the rapid AX learning predicted by RW independent of inhibitory parameters. Total activation of the correct neuron is plotted as a function of δ and block number

In contrast our second simulation demonstrates the vastly greater importance of inhibitory parameters on BX trials (currently under study). Activation is plotted as a function of δ vs. the inhibitory parameter after 100 trials. Note that a lack of general learning (chance) produces a better result than non-inhibitory learning which will only reinforce the prepotent response.

Conclusion

Results strongly support the alternative hypothesis: infant looker groups do not differ in first order learning of stimulus associations. However, conclusions regarding inhibition are

still dependent upon out current study (of which this was a precursor). The selective decreases in short looker's ability to discriminate the standard from target have several explanations. For example, reinforcement frequency may be a lurking variable. Moreover, the suspected AX-XA interaction would require inhibition to be overcome. However, the requirements for that task are challenging even for adults and so lower levels of inhibition linked with WM may have caused the interaction by coding both stimulus components with the reinforcer, instead of just the most proximal one.

while weighted optimizations found our model of stochastic first order-linear DE's to be the best possible model, it left over 4/5 of variance unaccounted for. As such it is likely that high stochasticity will continue to make modeling difficult for the current age-paradigm combination.

REFERENCES

1. Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353-383.

2. Freeseman, L., Colombo, J., Coldren, J. (1993) Individual differences in infant visual attention: Four-month-old's discrimination and generalization of global and local stimulus properties. *Child Development, 64(4),* 1191-1203.

3. Tamis-LeMonda, C. S., & Bornstein, M. H. (1989). Habituation and Maternal Encouragement of Attention in Infancy as Predictors of Toddler Language, Play, and Representational Competence. *Child Development*, 60, 738-751.

4. Guy, M. W., Reynolds, G., D., & Zhang, D. (2013) Visual attention to global and local stimulus properties in 6-month old infants: Individual differences and event-related potentials. *Child Development.*

5. Bayer, H.M & Glimcher, P.W. (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron,* 47, 129-141.

6. Rescorla, R.A. & Wagner, A.R. (1972). A theory of Pavlovian conditioning.: variations in the effectiveness of reinforcement and nonreinforcement. In Black, A.H. & Prokasy, W.F. (Eds), *Classical Conditioning II: Current Research and Theory.* Appleton-Century-Crofts, New York, NY, pp. 64-99.

7. Usher, M. & McClelland, J.L. (2001) The time course of perceptual choice: The leaky, competing, accumulator model. *Psychological Review*, 108(3), 550-592.

8. Ermentrout, B. (1998) Linearization of F-I curves by adaptation. *Neural Computation*, 10, 1721–1729.