Prefrontal Cortex Contributions to Learning in Infancy

by

Denise M. Werchan

B.S., The University of Arizona, 2012

Sc.M., Brown University, 2015

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Department of Cognitive, Linguistic, and Psychological Sciences at Brown University

Providence, Rhode Island

May 2020

This dissertation by Denise M. Werchan is accepted in its present form by the Department of Cognitive, Linguistic, and Psychological Sciences as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date ______________________ ____________________________________

Dr. Dima Amso, Advisor

Recommended to the Graduate Council

Date ______________________ ____________________________________

Dr. Michael J. Frank, Reader

Date ______________________ ____________________________________

Dr. Kevin Bath, Reader

Approved by the Graduate Council

Date ______________________ ____________________________________

Dr. Andrew G. Campbell, Dean of the Graduate School

Curriculum Vitae

Denise M. Werchan
Dept. of Cognitive, Linguistic, and Psychological Sciences
Brown University, Providence, RI 02912
Denise_Werchan@Brown.edu

EDUCATION

PhD in Cognitive Science, Brown University 2020
NSF Graduate Research Fellow
Committee: Dr. Dima Amso (chair), Dr. Michael Frank, & Dr. Kevin Bath

ScM in Psychology, Brown University 2013 – 2015
Thesis: Hierarchical Rule Learning and Generalization in 8-Month-Old Infants

BS in Psychology with Honors, The University of Arizona 2009 – 2012
Summa Cum Laude; Excellence in Research Award; Phi Beta Kappa Society

RESEARCH TRAINING

PhD Graduate Student Researcher 2013 – present
Developmental Cognitive Neuroscience Lab, Brown University
Director: Dr. Dima Amso

Functional Near-Infrared Spectroscopy (fNIRS) Training Course 2015
Martinos Institute for Biomedical Engineering at Harvard University

Research Assistant 2009 – 2013
Child Cognition Lab, The University of Arizona
Director: Dr. Rebecca Gómez

Research Assistant 2012
Memory Development and Disorders Lab, The University of Arizona
Director: Dr. Jamie Edgin

SELECTED HONORS, AWARDS & FUNDING

Robin Chemers Neustein Graduate Award in Brain Science	2018 – 2019
Betty R.H. and James M. Pickett Fellowship Fund in Psychology	2017 – 2018
Newton International Fellowship, The University of Cambridge (*declined*)	2017
Travel Award, International Congress of Infant Studies	2016
National Science Foundation Graduate Research Fellowship	2013 – 2016
Travel Award, International Society for Developmental Psychobiology	2014

Graduate Research Fellowship, Brown University	2013 – 2014
Sackler Developmental Psychobiology Fellowship, Cornell University	2013
Travel Award, Sleep Research Society	2013
Harvill Graduate Research Grant, The University of Arizona	2012 – 2013
Excellence in Research Award, The University of Arizona College of Science	2012
William James Award, The University of Arizona Department of Psychology	2012
Galileo Circle Scholarship for Excellence in Research, The University of Arizona	2012
Phi Beta Kappa Society	2012
Undergraduate Research Grant, The University of Arizona	2011
Sigma-Xi Grant-In-Aid of Research	2011
William C. Dement Sleep Research Fellowship, Brown University	2011

PUBLICATIONS

1. **Werchan, D.M.,** & Amso, D. (2020). Top-Down Knowledge Rapidly Acquired Through PFC-Dependent Rule Learning Biases Visual Attention in 9-Month-Old Infants. *Developmental Cognitive Neuroscience*.

2. **Werchan, D.M.,** Lynn, A., Baumgartner, H., Kirkham, N.Z., & Amso, D. (2019). The Emergence of Object-Based Visual Attention in Infancy: A Role for Family Socioeconomic Status and Competing Visual Features. *Infancy*.

3. **Werchan, D.M.**, Baumgartner, H., Lewkowicz, D. & Amso, D. (2018). The Origins of Cortical Multisensory Dynamics: Evidence From Human Infants. *Developmental Cognitive Neuroscience*.

4. **Werchan, D.M.** & Amso, D. (2017). A Novel Ecological Model of Prefrontal Cortex Functional Development. *Psychological Review, 24*(6), 720-739.

5. **Werchan, D.M.**, Collins, A.G.E, Frank, M.J., & Amso, D. (2016). Role of Prefrontal Cortex in Learning and Generalizing Hierarchical Rules in 8-Month-Old Infants. *The Journal of Neuroscience, 36*(40), 10314-10322.

6. Newman-Smith, K., **Werchan, D.M.**, Goldstein, M., Sweeney, L., Bootzin, R.R., Nadel, L., & Gómez, R.L. (2016). Sleep and Statistical Language Learning: Insights into Cortical Development and Retention in 6.5-Month-Olds. *Brain & Language, 167,* 3-12.

7. **Werchan, D.M.**, Collins, A.G.E, Frank, M.J., & Amso, D. (2015). Eight-Month-Old Infants Spontaneously Learn and Generalize Hierarchical Rules. *Psychological Science, 26*(6), 805-15.

8. **Werchan, D.M.** & Gómez, R.L. (2014). Wakefulness (Not Sleep) Promotes Generalization of Word Learning in 2.5-Year-Old Children. *Child Development, 85*(2), 429-436.

9. **Werchan, D.M.** & Gómez, R.L. (2013). Generalizing Memories Over Time: Reinforcement and Sleep Facilitate Transitive Inference. *Neurobiology of Learning and Memory, 100,* 70-76.

REFEREED ABSTRACTS

1. **Werchan, D.M.** & Amso, D. (2018, September). *Rule Learning Supports the Development of Visual Attentional Biases in 9-Month-Old Infants*. Poster presented at the Flux Congress on Integrative Developmental Cognitive Neuroscience, Berlin, Germany.

2. Placido, D., Nketia, J., **Werchan, D.M**. & Amso, D. (2018, September). *Exploring the effects of emotional enrichment on components of executive functions in children of varying socioeconomic status*. Poster presented at the Flux Congress on Integrative Developmental Cognitive Neuroscience, Berlin, Germany.

3. Novick, A. M., **Werchan, D.M**., Amso, D., Parade, S. H., Ackerman, L, Gunther, K., Seifer, R., Philip, N. S., & Tyrka, A. R. (2018, May). *Cognitive and Affective Function in School- Aged Children Experiencing Trauma and Adversity*. Poster presented at the 73rd Annual Meeting of the Society of Biological Psychiatry, New York, NY.

4. **Werchan, D.M.**, Gunther, K., & Amso, D. (2017, September). *9-Month-Old Infants use Higher-Order Contexts to Support Working Memory in the A-Not-B Task*. Poster presented at the Flux Congress on Integrative Developmental Cognitive Neuroscience, Portland, OR.

5. **Werchan, D.M.,** Baumgartner, H., Martino, R., Lewkowicz, D. & Amso, D. (2017, April). *Unisensory and Multisensory Processing in Cortex*. Poster presented at the Society for Research in Child Development, Austin, TX.

6. **Werchan, D.M**., Gunther, K., & Amso, D. (2017, April). *9-Month-Old Infants use Contextual Cues to Support Working Memory in the A-Not-B Task*. Poster presented at the Society for Research in Child Development, Austin, TX.

7. Baumgartner, H, **Werchan, D.M.**, Lewkowicz, D. & Amso, D. (2016, September). *Cortical Plasticity and Specialization in Response to Multimodal Events in Infancy*. Poster presented at the Flux Congress on Integrative Developmental Cognitive Neuroscience, St. Louis, MO.

8. **Werchan, D.M.,** Collins, A. G. E., Frank, M. J., & Amso, D. (2016, May). *8-Month-Old Infants Create Hierarchical Rules to Support Learning and Generalization in Novel Contexts*. Talk presented at the International Congress of Infant Studies, New Orleans, LA.

9. **Werchan, D.M.,** Kim, J., & Gomez, R. L. (2016, May). *Nighttime Sleep Enhances Preschoolers' Long-Term Generalization of Word Learning*. Talk presented at the International Congress of Infant Studies, New Orleans, LA.

10. Newman-Smith, K., **Werchan, D.M.,** Bootzin, R., Nadel, L., & Gomez, R.L. (2015, October). *Retention of Statistical Learning in 6.5-Month-Old Infants*. Talk presented at the annual meeting of the Society for Neuroscience, Chicago, IL.

11. **Werchan, D.M.,** Collins, A. G. E., Frank, M. J., & Amso, D. (2015, September). *Activation of prefrontal cortex during hierarchical rule learning in 8-month-old infants: data from near-infrared spectroscopy*. Poster presented at the Flux Congress on Integrative Developmental Cognitive Neuroscience, Leiden, The Netherlands.

12. Newman-Smith, K., **Werchan, D.M.**, Bootzin, R., Nadel, L., & Gomez, R.L. (2015, March). *Retention of Statistical Learning in 6.5-Month-Old Infants*. Talk presented at the Society for Research in Child Development, Philadelphia, PA.

13. **Werchan, D.M.**, Collins, A. G. E, Frank, M. J., & Amso, D. (2014, November). *Hierarchical rule learning and generalization in infancy*. Poster presented at the annual meeting of the International Society for Developmental Psychobiology, Washington, DC.

14. **Werchan, D.M.**, Collins, A. G. E., Frank, M. J., & Amso, D. (2014, November). *The development of hierarchical rule learning and generalization in infancy*. Talk presented at the Society for Neuroscience, Washington, DC.

15. **Werchan, D.M.**, Collins, A. G. E, Frank, M. J., & Amso, D. (2014, September). *The early emergence of hierarchical rule learning*. Talk presented at the Flux Congress on Integrative Developmental Cognitive Neuroscience, Los Angeles, CA.

16. Newman-Smith, K., **Werchan, D.M.**, Goldstein, M., Bootzin, R., Nadel, L., & Gomez, R.G. (2014, June). *A nap and its relationship to consolidation and retention of new learning in 6.5-month-olds*. Poster presented at the annual meeting of the Society for Sleep Research, Minneapolis, MN.

17. **Werchan, D.M.**, & Gomez, RL. (2013, June). *Wake (not sleep) enhances preschool age children's generalization of word learning*. Poster presented at the annual meeting of the Society for Sleep Research, Baltimore, MD.

18. **Werchan, D.M.**, & Gomez, RL. (2013, June). *Integrating memories over time: reward mediates the formation of inferences during sleep*. Poster presented at the annual meeting of the Society for Sleep Research, Baltimore, MD.

19. **Werchan, D.M.**, & Gomez, RL. (2013, April). *The role of sleep and reinforcement in memory integration*. Poster presented at the annual meeting of the Cognitive Neuroscience Society, San Francisco, CA.

20. **Werchan, D.M.**, & Gomez, RL. (2013, April). *Wake (not sleep) promotes generalization of word learning in toddlers*. Talk presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.

21. **Werchan, D.M.**, & Gomez, RL. (2012, December). *Forgetting promotes generalization of word learning in preschool age children*. Poster presented at the University of Arizona/Arizona State University Cognitive Science Conclave, Phoenix, AZ.

22. **Werchan, D.M.**, & Gomez, RL. (2012, October). *The effects of sleep and reinforcement on transitive inference*. Poster presented at the annual meeting of the Society for Neuroscience, New Orleans, LA.

TEACHING EXPERIENCE

Teaching Certificate 1: Reflective Teaching 2016
Brown University Sheridan Center for Teaching and Learning

Teaching Assistant, Statistics for the Behavioral Sciences 2016
Brown University, Cognitive, Linguistics, and Psychological Sciences Dept.

Guest Lecturer, "Neurodevelopmental Disorders", Elementary Psychology 2015
Brown University, Cognitive, Linguistics, and Psychological Sciences Dept.

Teaching Assistant/Section Leader, Elementary Psychology 2015
Brown University, Cognitive, Linguistics, and Psychological Sciences Dept.

Grader, Introduction to Cognitive Science 2014
Brown University, Cognitive, Linguistics, and Psychological Sciences Dept.

Teaching Assistant/Section Leader, Scientific Computing in Python 2011 – 2012
The University of Arizona, Computer Science Department

SERVICE & OUTREACH

Organizer, Communicating Science to Society Symposium 2016
Brown University Friedman Family Fund Lecture Series

Session Chair, Symposium on Word Learning, Sleep, and Contexts 2016
Annual Meeting for the International Congress for Infant Studies

STEM Workshop Leader, Introduction to Cognitive Neuroscience 2015 – 2016
Girl Scouts of America Senior Leadership Conference

Graduate Student Representative 2015 – 2016
Brown University Cognitive, Linguistics, and Psychological Sciences Dept.

Colloquium Committee Member 2012 – 2013
The University of Arizona Psychology Department

Preface and Acknowledgements

I would like to thank my family, friends, colleagues, and mentors for all of the support and encouragement that they have provided throughout my graduate education. This journey would not have been possible without having such a supportive and caring group of individuals by my side. In particular, I would like to thank my advisor, Dima Amso, for her continual guidance and support over the years. Dima has given me the freedom and flexibility to pursue my own research from day one, and has provided encouragement and feedback at every step along the way. I am grateful to have worked with such a dedicated and inspiring graduate advisor.

I would also like to thank my committee members, Michael Frank and Kevin Bath. I am truly grateful for all of the feedback and the time they have devoted to my project. Additionally, I would like to acknowledge Anne Collins, who provided much needed assistance and support when I began this research. I would also like to thank all of the members of Developmental Cognitive Neuroscience Lab that I have had the privilege of working with during my graduate career. I especially want to acknowledge Andrew Lynn for his continual support and camaraderie, and Laura Ackerman, Kelley Gunther, and Diego Placido, whose efforts and dedication as lab manager have made this process considerably easier.

Finally, I would like to thank the many families and infants who have participated in my studies over the years. I am truly indebted to these wonderful families who have volunteered their time for science, and this research would not have been possible without their generosity.

Table of Contents

3. Neural Underpinnings of Hierarchical Rule Learning in Infancy 72

4. A Rule Learning Account of Infants' Errors on Canonical Measures of Executive Functions 98

Figures and Tables

Chapter 5

Chapter 1

1. Introduction

The human prefrontal cortex (PFC) is perhaps the most elaborated and highly interconnected neocortical region, and is necessary for complex thought and action that is characteristic of higher-level human cognition (Badre, 2008; Badre & D'Esposito, 2009; Badre & Wagner, 2004; Koechlin, 2016; E. K. Miller & Cohen, 2001; O'Reilly, 2006; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005). This region also purportedly has one of the most extended developmental trajectories, with changes in its functional and anatomical properties occurring at least through late adolescence (Diamond, 2002; Giedd et al., 1999; Gogtay et al., 2004) and possibly throughout the lifespan (Anguera et al., 2013; Lee, Ratnarajah, Tuan, Chen, & Qiu, 2015; Li et al., 2014; Lövdén et al., 2010). In this dissertation, I offer a novel approach to understanding PFC functional development: I argue that far from having a protracted developmental course, the PFC continuously adapts its computations to accommodate the demands present in the changing ecological niche of the growing child. In this view, PFC functional development is emergent from change in the feedforward inputs from the rest of the brain, the physical structure of the growing body, the content and nature of existing knowledge, and the challenges and opportunities present in the external environment.

The novel model of PFC functional development proposed in this dissertation is inspired by ecological explanations for developmental change in cognition and behavior (Gibson, E. J., Pick, 2000; Rovee-Collier & Cuevas, 2009; Schneirla, 1957; Spear, 1984; Turkewitz & Kenny, 1982) and by a recent application of these ideas to brain development and risk for developmental psychopathology (M. H. Johnson, Jones, & Gliga, 2015).

Ecological accounts consider infants, children, and adults to be different organisms who occupy different ecological niches, each of which carries its own unique demands and challenges. Ecological approaches emphasize that organisms from all species have evolved to be adapted to their unique niches at each point in development, since optimal development of phenotype depends on adaptation to *all* environments, rather than adaptation only to the *final* environment (Lehrman, 1953). In this ecological view, infants and children have different sets of problems to solve for learning and behavior (Rovee-Collier & Cuevas, 2009). Thus, we may be limiting our understanding of ontogenetic brain development if we measure developmental change as relative only to the adult state.

In this first chapter, I will begin by reviewing the existing literature on the structural and functional development of the PFC. I will argue that these accounts are highly apt descriptions, but that they offer little mechanistic insight into how the system is developing, its catalysts, and its influences. I will then consider recent evidence that points to the hypothesis that *adaptation,* and not *maturation,* best describes the process of PFC developmental change. Throughout, I will highlight novel predictions raised by this account of PFC development, and will examine implications of this ecological model for redefining executive functions and for informing typical and atypical developmental trajectories. In the subsequent chapters, I will then present a series of empirical studies that test a key model prediction: that the PFC is adapted in infancy for learning flexible and generalizable rule structures.

1.1 PFC: The State of the Art

The human prefrontal cortex is a collection of interconnected neocortical regions that send and receive projections from nearly all primary sensory and motor systems, as well as many subcortical regions in the brain (Gilbert & Li, 2013; E. K. Miller & Cohen, 2001). The PFC has a number of unique characteristics. It is domain-general; through direct and indirect connections, it integrates and processes signals from almost every other neural region in the brain (Duncan & Owen, 2000; Fedorenko, Duncan, & Kanwisher, 2013). In addition, the PFC develops in the absence of direct input from sensory registers (Cahalane, Charvet, & Finlay, 2012). This is in contrast to more domain-specific neural regions that receive direct, stable sensory input, such as primary visual cortex. The PFC remains plastic at least through late adolescence (Diamond, 2002; Giedd et al., 1999; Gogtay et al., 2004) and possibly throughout the lifespan (Anguera et al., 2013; Lee et al., 2015; Li et al., 2014; Lövdén et al., 2010), providing increased opportunities for the changes in the internal and external environment to shape PFC development.

The PFC is anatomically defined as the projection zone of the mediodorsal nucleus of the thalamus in both primates and non-primates (Fuster, 2008). Within the PFC, there are a number of subregions that are delineated based on anatomical connections and granular structure (Barbas & García-Cabezas, 2016). These include the orbitofrontal PFC, ventrolateral PFC, dorsolateral PFC, rostrolateral, and medial PFC (Badre & D'Esposito, 2007, 2009; Bunge & Zelazo, 2006; Crone & Steinbeis, 2017; Koechlin, 2016).

PFC subregions are separately critical for supporting flexible and goal-directed control over cognition and action, emotion, and social behaviors across time and contexts (Badre, 2008; Cohen, Braver, & Brown, 2002; Kolb et al., 2012; E. K. Miller & Cohen, 2001; O'Reilly, 2006; Rougier et al., 2005). Collectively, these are referred to as *executive functions*. Different regions of the PFC appear to be specialized for different executive functions. For example, lateral portions of the PFC (including dlPFC) are important for goal-directed thought or action (Badre, 2008; Badre, Kayser, & D'Esposito, 2010; Barch et al., 1997; Bunge, 2003; Curtis & D'Esposito, 2003), whereas dorsomedial portions of the PFC are involved in decision-making under risk and uncertainty (Hadland, 2003; Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006; Rushworth, Walton, Kennerley, & Bannerman, 2004). Ventromedial PFC (vmPFC, including the orbitofrontal cortex) is implicated in decision making relevant to maximizing benefit and minimizing cost based on stimulus-reward relationships (Fellows & Farah, 2003; Hornak et al., 2004; Tsuchida, Doll, & Fellows, 2010). The functional specialization of the various subregions of PFC is due in part to differences in the input to each subregion from connections with other neural regions (Duncan & Owen, 2000; Fedorenko et al., 2013; E. K. Miller & Cohen, 2001).

1.1.1 Prenatal PFC Structural Development

During prenatal brain development, the primary structural features of PFC, and its anatomical connections with other regions, are formed by two major patterns of maturation, which in turn have significant implications for the subsequent function. First, during prenatal brain development, the cortex is populated with neurons in an anterior-to-posterior or "front-to-back" gradient, such that cell migration ceases in the anterior rostrolateral cortex first, and then progresses caudomedially towards more posterior regions (Cahalane et al., 2012; Finlay & Uchiyama, 2015, 2017). This gradient of neurogenesis results in frontal neurons becoming differentiated *earlier* than neurons in more posterior regions. It also occurs opposite of the gradient of thalamic innervation, where thalamic input innervates sensory regions prior to anterior association regions (Cahalane et al., 2012; Finlay & Uchiyama, 2017). These opposing patterns, of thalamic innervation and relative to cortical neurogenesis, are hypothesized to bias the PFC for its association role. That is, this anterior-to-posterior neurogenesis in PFC occurs in the absence of sensory input from thalamic registers and is therefore primarily shaped by the PFC's intrinsic cortical activity (Cahalane et al., 2012; M. H. Johnson et al., 2015).

Prenatal brain development is simultaneously characterized by an "inside-out" gradient with respect to the formation of cortical laminar structure. The laminar structure of the subregions of PFC varies. There are agranular and dysgranular regions that have fewer than six layers, such as the orbitofrontal (pOFC), anterior cingulate cortex (ACC), and vmPFC. There are also eulaminate regions that have six well-defined and delineated layers, such as the dlPFC (Barbas & García-Cabezas, 2016). This layered structure is formed during prenatal development in an inside-out pattern, where deeper layers are formed prior to upper layers (Sidman & Rakic, 1973).

The variance in laminar structure across the subregions of PFC is thought to be driven by differences in prenatal developmental timing of neural migration (Barbas & García-Cabezas, 2016; Dombrowski, Hilgetag, & Barbas, 2001). Specifically, agranular and dysgranular regions with fewer than six layers have a shorter developmental period relative to eulaminate regions. These differences in developmental timing are thought to arise from decreased neuronal density in the upper layers of agranular and dysgranular regions relative to eulaminate regions (Barbas & Garcia-Cabezas, 2016).

The systematic variations in laminar structure across the subregions of PFC bias their subsequent functions and connectivity with other regions (Barbas & Garcia-Cabezas, 2016). For example, regions with fewer or less complex layers, such as the anterior cingulate cortex (ACC) and vmPFC, primarily have feedforward and feedback connections with subcortical structures including the amygdala, hippocampus, and hypothalamus. Consistent with their patterns of connectivity, these subregions of PFC primarily process "internal" environmental information, such as emotions, rewards, motives, and drives (Barbas & Garcia-Cabezas, 2016). In contrast, eulaminate regions with six well-defined layers, such as the dlPFC, have feedforward and feedback connections from other eulaminate sensory regions, such as the primary sensory cortices, in addition to the agranular and dysgranular subregions of PFC. These patterns of connectivity thus make dlPFC well-placed to process information from sensory cortices as well as internal emotional and motivational information.

1.1.2 Postnatal PFC Structural Development

The existing understanding of PFC functional development has been primarily influenced by findings on PFC's *postnatal* structural development. The existing consensus is two-fold: (1) PFC postnatal structural maturation has a protracted course (e.g., Diamond, 2002; Fuster, 2002; Giedd et al., 1999; Gogtay et al., 2004; Shaw et al., 2008); (2) This structural development then supports associated executive functions development as well as individual differences in executive functions efficiency (e.g., Diamond, 2002; Fuster, 2002; Kharitonova, Martin, Gabrieli, & Sheridan, 2013; Paus, 2005; Sheridan, Kharitonova, Martin, Chatterjee, & Gabrieli, 2014). This view is further supported by evidence from twin studies that executive functions are 99% heritable (Friedman et al., 2008). However, these estimates of genetic heritability are likely inflated due to shared environments among twins. For instance, twins in low socioeconomic environments have much lower estimates of cognitive heritability (Harden, Turkheimer, & Loehlin, 2007; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003).

The existing framework is shaped by findings that PFC postnatal structural maturation is greatly protracted in comparison to other cortical and subcortical neural regions. During the first few years of life, the brain rapidly quadruples in size, reaching 90% of its adult volume by 6 years of age (Courchesne et al., 2000; Knickmeyer et al., 2008). During this period of rapid early development, PFC expands over twice as much as other cortical regions (Hill et al., 2010). This period is characterized by large increases in synaptogenesis and neurogenesis, followed by periods of neuronal pruning and synaptic death to accommodate the early increases in neural and synaptic formation (Stiles & Jernigan, 2010). Some findings suggest that these rates of neural and synaptic formation and elimination occur heterochronously across different cortical regions (Rakic, 2002). During early gestational development, synaptogenesis and neurogenesis are fairly evenly distributed across cortical regions, and neuronal density within PFC is similar to other cortical regions (Shankle, Rafii, Landing, & Fallon, 1999). However, some studies suggest that soon thereafter, rates of neural and synaptic formation across different cortical regions quickly change, with PFC being one of the latest regions to reach its peak. For example, work analyzing post-mortem human brains by Huttenlocher & Dabholkar (1997) found that synaptogenesis peaked at approximately 3 months of age in sensorimotor cortex, whereas synaptogenesis occurred much more slowly in PFC, not reaching its peak until around 3.5 years of age (Huttenlocher & Dabholkar, 1997). That the number of synapses in PFC appears to peak early in childhood does not necessarily indicate that new synapses are no longer forming. Rather, it may reflect that the rate of synaptic elimination exceeds the rate of synaptogenesis (Petanjek et al., 2011; Selemon, 2013). However, note that several other studies of non-human primate brains found no differences in synaptogenesis across cortical regions (Bourgeois, Goldman-Rakic, & Rakic, 1994; Rakic, Bourgeois, Eckenhoff, Zecevic, & Goldman-Rakic, 1986; Zecevic, 1998; Zecevic, Bourgeois, & Rakic, 1989). These discrepancies have been suggested to reflect methodological confounds between primate and human data (Goldman-Rakic, Bourgeois, & Rakic, 1997).

Similarly, PET studies with human infants have shown that PFC has a lag of up to 8 months in reaching peak levels of glucose metabolism – a measure of neural activity (Chugani & Phelps, 1986). This is in contrast to temporal, parietal, and occipital cortices, where peak levels of glucose metabolism are observed soon after birth (Chugani & Phelps, 1986). Another useful index of brain development is cortical thickness. Measuring cortical thickness provides a composite index of overall maturity that includes neurons, synapses, axons, dendrites, and glia. During prenatal development, cortical thickness increases linearly across the entire brain as a function of time (Rabinowicz, de Courten-Myers, Petetot, Xi, & de los Reyes, 1996). During postnatal development, however, overall cortical thickness of PFC follows an inverted U-shaped trajectory, increasing throughout childhood, peaking in adolescence, and then slowly decreasing and stabilizing in early adulthood (Shaw et al., 2006). Protracted changes in white and grey matter volume within PFC also occur well into adolescence (Giedd et al., 1999; Gogtay et al., 2004; Reiss, Abrams, Singer, Ross, & Denckla, 1996). These increases are more protracted in PFC compared to other regions of the cortex, where increases are primarily observed up to around 6-9 years of age (Courchesne et al., 2000). We note that these agranular and dysgranular PFC subregions seem to also structurally develop, based on cortical thickness, postnatally earlier than eulaminate PFC subregions (Shaw et al., 2008).

Recent work also shows protracted changes in myelination of pathways between PFC and other neural regions. Maturation of white matter tracks is found to broadly follow a posterior-to-anterior gradient, such that sensory and motor regions myelinate first and cortical association regions including PFC myelinate last (Deoni et al., 2011; Deoni, Dean, O'Muircheartaigh, Dirks, & Jerskey, 2012). While changes in myelination are important as it is assumed that they result in an increase the efficiency of neural communication, undermyelinated pathways are also capable of transmitting signals, albeit less efficiently (M. H. Johnson & De Haan, 2015).

1.1.3 PFC Functional Development

The existing framework of PFC structural and functional development has primarily linked PFC postnatal structural development to age-related changes in executive functions (Diamond, 2002; M. H. Johnson, 1990; M. H. Johnson, Posner, & Rothbart, 1994; Luna & Sweeney, 2004; Stuss, 1992). The argument that structural maturation drives cognitive development is particularly compelling since the functional developmental course of executive functions closely follows the developmental course of PFC maturation (Amso, Haas, McShane, & Badre, 2014; M. C. Davidson, Amso, Anderson, & Diamond, 2006; Snyder & Munakata, 2010; Wendelken, Munakata, Baym, Souza, & Bunge, 2012; Zelazo et al., 2003). The close parallels between PFC maturation and improvements in executive functions have therefore led many to propose that structural maturation of PFC allows new or more advanced abilities to "come online".

In human infants, one of the earliest and most comprehensive lines of work relating PFC maturation to changes in executive functions comes from work using Piaget's A-not-B task (Piaget, 1952). In standard versions of this task, infants watch as an experimenter hides a desirable object in one of two possible locations, and infants are allowed to reach to search for the object after a brief delay. During the task, the object is typically hidden in the same location (A) for multiple trials before it is reversed and hidden in the alternate location (B). The A-not-B error occurs when infants reach to the previously correct location (A) rather than the new location (B) on these reversal trials. Piaget was the first to observe that infants younger than 7 months tend to make this perseverative A-not-B error (Piaget, 1952). That is, infants younger than 7 months of age typically reach to the location where the object was hidden on the immediately preceding trial, rather than to the new location.

Between 7.5-9-months, infants begin to reach to the new location on reversal trials at successively longer delays ranging from 1-5 seconds (Diamond, 1985, 2001). However, as the delay between hiding and searching is incrementally increased, infants continue to make the A-not-B error until around 12 months of age (Diamond, 1985). Infant monkeys show similar developmental progressions on this task as human infants (Diamond & Goldman-Rakic, 1989). Anatomical work with monkeys provides direct evidence that behavioral improvements parallels maturation of dorsolateral PFC (dlPFC), suggesting that structural maturation of PFC supports cognitive changes (Diamond & Goldman-Rakic, 1989). This idea is further supported by evidence that adult monkeys with dlPFC lesions also show the A-not-B error when there is a delay between hiding and searching (Diamond & Goldman-Rakic, 1989). In human infants, evidence that improvements on the A-not-B task relate to PFC development is found from work using electroencephalography (EEG) and near-infrared spectroscopy (NIRS) to record infants' brain activity during cognitive testing. For example, frontal EEG responses in 7-12-month-old infants correlate with behavioral performance in the A-not-B task (Bell & Fox, 1992; Fox & Bell, 1990). A longitudinal NIRS study with 4-12-month-old infants also indicates that PFC blood oxygenation levels correlate with behavioral performance in this task (Baird et al., 2002).

Data support the linkage between PFC maturation and executive functions development into late childhood and adolescence. For example, a large cross-sectional study of 3-25-year-olds who were tested on an extensive battery of cognitive tasks showed that tasks that are dependent on more posterior brain regions, such as recognition memory, tend to stabilize around 8 years of age (Luciana, 2003; Luciana & Nelson, 1998, 2002). However, tasks that recruit prefrontal cortical regions show more protracted development, not nearing adult levels of performance until around 12 years of age. Maturation of white matter tracts within frontoparietal networks also correlates with improvements in executive functions. For example, in children ages 8-18 years, maturation of white matter tracts, as measured by diffusion tensor imaging (DTI), is associated with increased working memory capacity (Olesen, Nagy, Westerberg, & Klingberg, 2003), as well as with increased task-dependent functional activation within frontal and parietal brain regions (Klingberg, Forssberg, & Westerberg, 2002).

Taken together, this very brief review of an extended body of evidence provides support for a structural maturation account of PFC functional development. However, by necessity, these studies reflect inferences drawn based on correlations between structure and function, most often across different studies and populations, and no evidence of directional causation. Further, the existing framework has primarily examined developmental changes in PFC functional development using a model of function that was developed based on adult data. Thus, the existing account, of the causes of PFC developmental change, offers little in the way of mechanistic predictions other than the assumption that functions "come online" at certain ages. Moreover, these viewpoints consider PFC functional development largely with a spotlight on frontostriatal and frontoparietal change, using a construct of executive functions that is defined relative to an adult state (Dosenbach et al., 2007). However, PFC is a widely interconnected region, sending and receiving inputs from auditory, emotion, memory, and motor areas as well. The narrow spotlight constrains a full ontogenetic understanding of PFC functional development by limiting explanatory power to descriptions of how PFC functions become more "adult like", rather than considering the *mechanisms* driving developmental change.

1.2 PFC: Revising Assumptions

Recent behavioral, neuroanatomical, and electrophysiological research provides novel insights into the functional development of the PFC. Specifically, the PFC is functionally active and involved in organizing complex behaviors from the first months of life, that its rate of development is not deterministic but can be impacted and even accelerated by extreme experience, and that executive functions reflect changes in whole brain connectivity above and beyond simple PFC structural maturation. We review each of these findings in turn.

1.2.1 PFC is Active and Implicated in Behavior as Early as Birth

Research examining resting-state brain activation in infants using fMRI shows that infants have PFC activity in frontoparietal networks from birth (Doria et al., 2010). This indicates that PFC is active even before the emergence of classic executive functions later in infancy. Evidence from near-infrared spectroscopy (NIRS) and positron emission tomography (PET) studies also shows that infants frequently have PFC activation during cognitive tasks, even in cases where adults *do not* have PFC activation. For example, one study used NIRS to examine newborn infants' PFC activation when listening to adult-directed compared to infant-directed speech (Saito et al., 2007). The authors found that infants showed strong cortical activation over anterior PFC when listening to infant-directed speech, but not when listening to adult-directed speech. Other work showed that 3-month-olds have activation in the dorsolateral PFC when listening to forward speech, but not backward speech (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002). This finding was evident only when infants were awake and attentive and not when they were asleep. These studies suggest that PFC is involved in processing linguistic input that is relevant to infants during early life.

PFC might also help process social information in young infants. For example, work by Grossmann and colleagues (2008) used NIRS to test whether PFC is implicated in early social cognition in young infants. During this study, 4-month-olds were shown videos where an actor either established mutual gaze with the infants, or averted their gaze away from infants. Infants had increased right fronto-polar cortical activation only when viewing videos where the actor established mutual gaze, and not when the actor averted their gaze. Similar work shows that 5-month-olds also have strong dorsal PFC activation when attempting to coordinate their attention with others (Grossmann & Johnson, 2010; Grossmann, Lloyd-Fox, & Johnson, 2013). As infants get older, the specificity of PFC activation to social information may increase. For example, 9-12-month-olds have increased anterior orbritofrontal activation only when they view their mother smiling and not when they view strangers smiling (Minagawa-Kawai et al., 2009). These studies suggest that PFC may play an important role in processing socially relevant information in infants' environments.

PFC also appears to be implicated when infants plan motor or oculomotor actions, or when they view actions of others. For example, work by Csibra, Tucker, and Johnson, (2001) examined 4-month-old infants' frontal ERP responses while viewing sequences of predictable stimuli presentations. When making proactive (anticipatory) eye movements, infants had positive eye movement related potentials over frontal regions. Importantly, however, they did not have these frontal potentials when making reactive eye movements, which suggests that PFC may be involved in planning oculomotor actions from very early in life (Richards, 2000). Other work using NIRS shows that 5-month-olds also have increased inferior frontal cortical activation when viewing videos of human motor actions, but not when viewing other forms of motor actions, such as mechanical, non-human actions (Lloyd-Fox, Blasi, Everdell, Elwell, & Johnson, 2011).

A body of research also points to a role for PFC in regulating and processing emotional responses (Ahmed, Bittencourt-Hewitt, & Sebastian, 2015; Gilmartin, Balderston, & Helmstetter, 2014; Ochsner & Gross, 2008) as early as infancy. For instance, work using EEG has found that left frontal asymmetry, as measured by increased activation in left relative to right frontal regions, relates to emotional regulation and temperament in infancy (R. J. Davidson & Fox, 1982, 1989; Dawson, Panagiotides, Klinger, & Hill, 1992; Diaz & Bell, 2012; Fox, 1991, 1994; Fox & Davidson, 1987). Frontal asymmetry is also modulated by environmental factors. For example, infants of depressed mothers typically have opposite patterns of frontal asymmetry relative to typical controls (e.g., right frontal asymmetry, rather than left frontal asymmetry), which parallels findings from inhibited infants and from chronically depressed adults (Field, Fox, Pickens, & Nawrocki, 1995; Jones, Field, & Almeida, 2009; Jones, Field, Fox, Lundy, & Davalos, 1997; Lusby, Goodman, Bell, & Newport, 2014).

Given evidence for architectural constraints on the prenatal organization of the PFC, as well as constraints on the computational properties of neurons in the human brain (e.g., Elman et al., 1996), it would seem unlikely that the PFC and its subregions are performing fundamentally different tasks in infants than in children or adults. We propose that each PFC subregion is likely performing similar computations across the lifespan. The domain generality of PFC regions affords opportunities for those computations to be applied to information that is relevant and available to the organism in their current developmental state. Taken together, this body of evidence suggests the possibility that PFC may be involved in creating rule structures that help organize linguistic, social, emotional, and oculomotor learning and action as early as in infancy.

1.2.2 Connectivity, not PFC structural maturation per se, may be key to executive functions development

Methodological and statistical advances, resting state fMRI combined with functional connectivity analyses, have furthered our understanding of the structural and functional changes that occur during human brain development. Recent data from these methods show that changes in PFC *connectivity*, rather than solely structural maturation of PFC, may support changes in executive functions.

Connectomics is a recent addition to the field of functional magnetic resonance imaging (fMRI) research that examines the functional coupling of brain regions into networks, as measured by correlated activation of brain regions (Sporns, 2013). One benefit of connectomics research for understanding PFC development is that it examines how PFC changes and develops *within the context of the entire brain*, rather than as a singular region. Examining changes within the context of a large interconnected system is especially important for understanding PFC development, since PFC has extensive connections with other neural regions.

Research in this area demonstrates the existence of highly connected “hub” regions that play an important role in global information integration between different regions of the brain (van den Heuvel & Sporns, 2013). Interestingly, hub regions in the human connectome change across development. In early postnatal life, the first functional hubs are found in unimodal cortical areas, including auditory, visual, and sensorimotor cortices (Fransson, Åden, Blennow, & Lagercrantz, 2011). The location of these hubs shifts throughout development, eventually settling to hubs in the posterior cingulate, insula, and other heteromodal cortices by adulthood (Fransson et al., 2011).

These shifts are thought to partially reflect changes from segregation to integration with respect to regional functional connectivity. During early infancy a “small-world” architecture dominates functional connectivity across the entire brain, with increased short-range connections within cortical regions and decreased long-range connections between cortical regions. This pattern reflects segregation. The overall number of connections remains constant across development; however, the number and strength of long-range connections between different cortical regions increases while the quantity of short-range connections within regions decreases (Fair et al., 2007; Gao et al., 2011; Supekar, Musen, & Menon, 2009). This pattern reflects regional integration and relates to executive functions in tasks assessing reasoning ability. For example, shifts in functional connectivity from short-range intracortical connections within PFC to long-range frontoparietal connections correlates with age-related improvements in relational reasoning abilities during the transition from childhood to adolescence (Wendelken, Ferrer, Whitaker, & Bunge, 2016). Similarly, developmental improvements in inhibitory control during an antisaccade task correlate with increases in long-range functional connections between multiple regions of the PFC (including the ACC, medial frontal gyrus, inferior frontal gyrus, and frontal eye fields) and subcortical and parietal regions, which is thought to reflect increased top-down modulatory control over behavior (Hwang, Velanova, & Luna, 2010). However, note that the magnitude of shifts from short- to long-range functional connections may have been exaggerated due to increased motion artifacts in younger children relative to older children and adults (see Power, Barnes, Snyder, Schlaggar, & Petersen, 2012).

It is notable that changes also occur at different rates for anatomical compared to functional connections. Nonhuman primate work has shown that structural constraints, in the form of feedforward and feedback cortical connections, are in place prenatally (Goldman-Rakic, 1987). Anatomical connections within the frontoparietal network are adult-like by around 9 months of postnatal life (Conel, 1939); however, functional connectivity between these regions only *begins* to be evident between 6-9 months of age (Fransson et al., 2007) and develops into young adulthood, which coincides with increases in myelination (Giedd et al., 1999; Klingberg et al., 2002).

Simultaneous increases in functional connectivity and myelination are thought to occur, at least in part, in response to experience. A principle idea in the postnatal neural development of cortex is that there is a minimum of genetic instruction necessary for organized developmental change, and that developmental change instead occurs largely in response to experience (Finlay & Uchiyama, 2017; Kolb et al., 2012). There is, however, strong evidence that cortical areas are hierarchically organized along a rostro-caudal gradient (e.g., Van Essen & Maunsell, 1983). In this context and given its location, PFC necessarily computes over inputs from sensory, motor, and basal forebrain regions (Amso & Scerif, 2015; Finlay & Uchiyama, 2015; Gilbert & Li, 2013).

Experience-based changes in the mammalian cortex are thought to relate to changes in connectivity (Caroni, Donato, & Muller, 2012; Holtmaat & Svoboda, 2009), which can be improved via changes in axonal myelination, as well as synaptic plasticity via dendritic remodeling and synaptic pruning. For instance, in recent work, C. M. Johnson and colleagues (2016) examined changes in PFC dendritic spine formation and elimination in response to experience using two-photon imaging and a rule learning task in a rodent model. They found that animals in the rule training group had greater OFC to dmPFC bouton turnover. At the same time, the total bouton density did not differ between the groups and boutons gained were not lost after rule reversal training. These data represent a change in synaptic structure in PFC in response to learning a novel rule structure. They argue that there is not a net pruning of dendritic spines in response to rule reversal learning, but that experience instead results in persistent changes to PFC synaptic structures (Hofer, Mrsic-Flogel, Bonhoeffer, & Hübener, 2009; Holtmaat, Wilbrecht, Knott, Welker, & Svoboda, 2006; Muñoz-Cuevas, Athilingam, Piscopo, & Wilbrecht, 2013; Yang, Pan, & Gan, 2009).

As such, changes in functional connectivity that occur in response to environmental experience may be relevant to functional PFC development. Connectivity thus becomes an important piece in the ecological account, which considers PFC development as a process of experience-dependent adaptation.

1.2.3 Early life adversity modulates PFC development

A third line of work shows that, rather than PFC having a strict maturational course, early childhood adversity and deprivation can modulate the development of PFC and its functional connections with other neural regions thereby impacting PFC function (Brito & Noble, 2014; Farah et al., 2006, 2008; Hackman, Farah, & Meaney, 2010; Hackman, Gallop, Evans, & Farah, 2015; McLaughlin et al., 2014; Noble et al., 2015). For instance, a number of studies have found that low socioeconomic status results in overall slower development of cortex, as measured by delayed gray matter maturation and decreased cortical thickness within PFC (Hair, Hanson, Wolfe, & Pollak, 2015; Mackey et al., 2015; McLaughlin et al., 2014; Noble et al., 2015), as well as significant reductions in white matter tracts between PFC and other cortical and subcortical regions (Eluvathingal, 2006; Hanson et al., 2013; Hanson, Knodt, Brigidi, & Hariri, 2015; Sheridan, Fox, Zeanah, McLaughlin, & Nelson, 2012). These structural changes are paralleled by worse performance on classic measures of executive functions (see Brito & Noble, 2014;

Hackman & Farah, 2009; S. B. Johnson, Riis, & Noble, 2016, for review), potentially reflecting adaptation to environments with little enrichment or challenges for learning and behavior. This idea is also supported by evidence showing that the most effective interventions for children at risk for poor executive functions development tend to be those that increase overall environmental enrichment and that progressively increase the demand placed on PFC processing (Diamond & Lee, 2011). Thus, these studies highlight that PFC structural maturation is not deterministic, but is greatly impacted by the environment and experience, potentially as an adaptive response to the characteristics of an individual's environmental constraints.

Additional evidence for the role of adverse early environments in modulating PFC development comes from work examining early life stress in the form of maternal deprivation. This work has found that early life stress impacts the functional development of circuitry linking ventromedial PFC (vmPFC) and amygdala (for a review, see Callaghan & Tottenham, 2016). In typical development, maturation of vmPFC-amygdala circuitry occurs during the transition from childhood to adolescence, and it is also associated with improvements in emotional regulation (Callaghan, Sullivan, Howell, & Tottenham, 2014). For example, typically developing children have immature patterns of vmPFC-amygdala connectivity (e.g., positive coupling between vmPFC and amygdala activity) during emotional regulation tasks (Gee, Humphreys, et al., 2013). This positive coupling mirrors behavioral patterns of fear responses, such as increased (but developmentally normative) levels of separation anxiety. During the transition to adolescence, vmPFC-amygdala connectivity begins to mature to a more adult-like pattern, such that adolescents have negative coupling between vmPFC and amygdala (Gee, Humphreys, et al., 2013). These changes also parallel decreases in reported levels of separation anxiety.

Recent work examining the impact of early life stress on the timing of vmPFC-amygdala circuitry maturation provides support for the idea that maturation may reflect ecological adaptation. For example, rodent models indicate that early life adversity in the form of maternal deprivation *accelerates* the development of vmPFC-amygdala circuitry, such that adult-like states are reached at an earlier age than in typical development (for a review, see Callaghan & Tottenham, 2016). This accelerated development is modulated by elevated cortisol levels (Callaghan & Richardson, 2012, 2014), and it also relates to the early emergence of adult-like fear learning that is supported by this circuitry (Callaghan & Richardson, 2011). Similar findings have also been observed in humans. For example, work by Gee, Gabard-Durnam, and colleagues (2013) found that previously institutionalized children who were maternally deprived during early childhood had accelerated maturation of vmPFC-amygdala circuitry. Specifically, maternally-deprived children exhibited positive coupling between vmPFC and amygdala activity when viewing emotional faces earlier in development than typical comparisons. This accelerated development was also paralleled by decreases in developmentally normative levels of separation anxiety, which suggests that accelerated maturation of vmPFC-amygdala circuitry may be an adaptive response to environments where extreme maternal separation is experienced. Similar work also shows that history of maternal separation influences exploration versus exploitation strategy use in children (Humphreys et al., 2015), providing further support that atypical developmental trajectories associated with early life stress might reflect adaptations to adverse early environments.

Note that whether adaptation results in protracted or accelerated development may depend on whether the driver is "external" (e.g., sensory input) or "internal" (e.g., emotions, motives, and drives). As discussed in the prenatal development section above, vmPFC has fewer cortical layers, matures earlier, and primarily sends and receives connections from subcortical structures, whereas dlPFC has six well-defined layers, matures more slowly, and sends and receives projections from sensory regions as well as from other subregions of PFC (Barbas & Garcia-Cabezas, 2016). Thus, internal drivers, such as heightened amygdala reactivity and stress resulting from maternal deprivation, may lead to accelerated development of vmPFC through direct feedforward pathways from amygdala and other subcortical structures to vmPFC (Barbas & Garcia-Cabezas, 2016). In contrast, delays in dlPFC development, as often seen in low socioeconomic environments, may be mediated through either connectivity with sensory regions or indirectly with subcortical limbic regions (Barbas & Garcia-Cabezas, 2016), which may result in delayed maturation (Hair et al., 2015; Mackey et al., 2015; McLaughlin et al., 2014; Noble et al., 2015). This is supported by recent neuroimaging findings indicating that low socioeconomic environments appear to influence PFC's cortical thickness and white matter microstructure through limbic regions (Lawson, Duda, Avants, Wu, & Farah, 2013; Ursache & Noble, 2016).

1.3 An Ecological Account of PFC and Executive Functions Development

Taken together, recent findings require a reconceptualization of PFC functional development. The PFC is active and implicated in organizing learning and action, particularly in social, linguistic, emotional, and oculomotor domains relevant to infants. In childhood, the PFC adapts to be involved in regulating thought and action control processes. Rather than being on a fixed maturational time-course, PFC development can be delayed or accelerated in response to environmental demands. Finally, executive functions development is linked to PFC connectivity with the rest of the brain, and not to volumetric changes within PFC *per se*.

This literature supports the hypothesis that PFC development may be an *adaptive response* to changes in the organism's internal or external environment. The term *environment* here reflects the integrity of inputs to PFC from the rest of the brain (which are additionally constrained by the physical structure of the changing body), the nature and content of the knowledge gained, and the opportunities and challenges present in a child's external environment. As noted earlier, PFC serves as a domain general processing system that performs computations across a range of different inputs. While neuroconstructivist viewpoints that incorporate interactive specialization may be better suited for understanding cortical specialization in regions that process specific, stable input, such as visual cortex (M. H. Johnson, 2000, 2011), an ecological approach may be more apt for understanding how domain general systems develop as a function of adaptation to changing multi-modal input that necessitates organization. This account rests on the tenet the PFC performs similar computations across the lifespan, and that these computations are adapted to the unique ecological niche occupied by individuals at each point in development. This is a process-oriented account of PFC and executive functions, which dissociates the cognitive processes supporting executive functions from the specific representations that these processes operate over (e.g., (Duncan, 2001; Duncan & Miller, 2002).

The key premises of the ecological account of PFC functional development are:

(1) The PFC performs the same basic computations across the lifespan, but these computations are continually adapted for novel learning demands relevant to an individual in their unique ecological niche.

(2) PFC functional development will be reciprocally influenced by adaptation to changes in the input to the PFC via feedforward connections, as well as through niche construction via PFC's feedback connections to other neural regions.

(3) Both adaptation and niche construction will be constrained by anatomical proximity of regions in relation to the PFC.

We will expand on each of these key ideas and highlight existing empirical support for them, as well as discuss novel testable predictions arising from them, in the subsequent sections.

1.4 PFC Computations and Redefining Executive Functions

A neural computation is defined here as a mathematical means of describing a neural pattern of response relevant to function. Computational models provide support for the idea that many of the broad thoughts and behaviors that PFC supports can be captured by a set of neural computations. For instance, a recent broad view implicates PFC in mounting adaptive responses in uncertain or variable environmental contexts (Koechlin, 2016). In this view, computations in vmPFC are used to learn the expected reward values of stimulus-action pairings, whereas those in lateral portions of the PFC implement more complex and hierarchically nested state-action mappings. A derivative process within lateral PFC, for example, is that it supports the active maintenance of information, which is represented by distributed patterns of neural activity (Badre, 2008; Cohen et al., 2002;

D’Esposito & Postle, 2015; E. K. Miller & Cohen, 2001; O’Reilly, 2006; Rougier et al., 2005). This process allows information to be maintained in working memory over time, protecting it from interference from distracting or irrelevant inputs. Recurrent excitatory connections within PFC, as well as the intrinsic bistability of PFC neurons support this computation, which is also modulated by dopamine activity within PFC (Cohen et al., 2002; Constantinidis & Klingberg, 2016; Durstewitz, Seamans, & Sejnowski, 2000; Fallon, Williams-Gray, Barker, Owen, & Hampshire, 2013; Lew & Tseng, 2014; O’Reilly, 2006; Rougier et al., 2005).

A second computation supported by PFC is the adaptive updating of patterns of neural activity by dynamically switching between active maintenance and rapid updating of new representations (Braver & Cohen, 2000; Chatham & Badre, 2015; Chatham, Frank, & Badre, 2014; Chiew & Braver, 2017; Cohen et al., 2002; M. J. Frank, Loughry, & O’Reilly, 2001; O’Reilly, 2006; Rougier et al., 2005). This computational process is thought to be supported by a selective, dopamine-modulated gating mechanism that controls the flow of information into PFC (Braver & Cohen, 2000; Chatham & Badre, 2015; Chatham et al., 2014; Chiew & Braver, 2017; Rougier et al., 2005). VMPFC is involved in switching from exploration to exploitation states, relevant to uncertainty in choice value outcomes (Domenech & Koechlin, 2015; Koechlin, 2016). In particular, data suggest that, together with the vmPFC, the dACC is involved in making the switch from exploration to exploitation policy (Donoso, Collins, & Koechlin, 2014; Kolling, Behrens, Mars, & Rushworth, 2012).

A third PFC computation involves modulating neural processing in other cortical areas that are required for successful task execution (Braver, Paxton, Locke, & Barch, 2009; Buschman & Miller, 2007; Duncan, 2001; Fuster, 2008; Gilbert & Li, 2013; B. T. Miller & D'Esposito, 2005; E. K. Miller & Cohen, 2001; O'Reilly, 2006; Rougier et al., 2005). This process is supported by PFC's extensive interconnectivity with other subcortical and cortical neural regions. As mentioned previously, PFC integrates inputs from other neural regions; however, rather than acting simply as an "integrator" and "transmitter" of these inputs, PFC modulates the flow of activity in subcortical and posterior regions via top-down feedback pathways to align with the goals and values of currently relevant information that is maintained within PFC (Miller & Cohen, 2001; Gilbert & Li, 2013).

Framing PFC's role in learning and behavior as a set of computations is useful in that allows for understanding of how a common underlying process or computation can support a diverse range of functions that are appropriate for humans in each stage of their development. For example, an ecological account would predict that during early infancy, maintenance and gating may underlie planning of saccadic eye movements to sample information required for learning environmental contingencies. This aligns with findings showing that 4-month-old infants have frontal ERPs during anticipatory but not reactive eye movements during a simple stimulus-response learning task (e.g., Csibra et al., 2001). As infants begin learning to reach and grab objects, these same computations may be used to plan goal-directed reaching for objects, as supported by work indicating that 5-month-old infants have increased frontal activation when viewing human motor actions, but not non-human actions (e.g., Lloyd-Fox et al., 2011). During early childhood, these computations may then be adapted to support learning of novel complex social and moral action rules, mathematical concepts etc., in the environment. In adults, the same PFC region may implement its computations of gating and maintenance of task-relevant information to select complex rules for actions (e.g., planning the necessary actions required for cooking a complex meal, balancing a budget, planning for the future, following a diet).

The ecological account, with its focus on adaptation, also offers a revised definition of executive functions. Executive functions can be defined in this framework as those computations performed by PFC at any point in the lifespan. This redefinition offers the prediction that developmental continuity should be observed in PFC's computations when ecologically appropriate tasks are used at different developmental stages. For example, the PFC is involved in hierarchically organizing and processing language in infancy, recalling the data on infant-directed speech (Saito et al., 2007). Early language development has been shown to be a strong predictor of later, more classic frontostriatal executive functions tasks in childhood (e.g., (Gooch, Thompson, Nash, Snowling, & Hulme, 2016; Kuhn, Willoughby, Vernon-Feagans, & Blair, 2016; Noble, McCandliss, & Farah, 2007; Noble, Norman, & Farah, 2005). *If executive functions reflect the computations performed by PFC at any point in the lifespan, then these developmental relationships might reflect the efficacy with which PFC's computations are adapted to different demands across the lifespan.* This framework thus provides a novel mechanistic explanation for the relationship between delayed language acquisition and executive functions delays (e.g., (Figueras, Edwards, & Langdon, 2008; Henry, Messer, & Nash, 2012).

It is also worth emphasizing how this ecological account differs from other accounts that argue that protracted maturation of PFC may be an adaptation during early childhood. According to these alternative accounts, delayed maturation of PFC may be adaptive during early childhood in that it helps individuals in the species more efficiently learn the natural statistics of the environment, without top-down PFC control imposing constraints on this process (Chrysikou, Novick, Trueswell, & Thompson-Schill, 2011; Chrysikou, Weber, & Thompson-Schill, 2014). This account assumes a PFC protracted developmental course is accurate and assigns evolutionary value to this developmental course (Thompson-Schill, Ramscar, & Chrysikou, 2009). In contrast, we argue against protracted maturation and an evolutionary explanation for this being adaptive. We suggest instead that the PFC is instead *directly adapted* to support learning demands that are relevant to individuals in their unique ecological niche, whether it be learning and generalizing relevant environmental contingencies in infancy and early childhood or organizing and exerting control over complex action and thought in adolescence and adulthood.

Thus, rather than having a protracted developmental course, executive functions may be continually adapting to processing information relevant to developing children. Once the physical structure of the body and environmental demands on learning and behavior begin to stabilize from adolescence to early adulthood, our ecological account makes the novel prediction that PFC may then stabilize in its development. However, as PFC's connectivity and the structure and capabilities of the physical body begin to destabilize again in ageing, PFC may again show changes and deficits in its functions in comparison to the young adult state (Anguera et al., 2013; Lee et al., 2015; Li et al., 2014; Lövdén et al., 2010).

1.5 Adaptation and Niche Construction

In this section, we expand on how the concepts of adaptation and niche construction are relevant to PFC functional development. Niche construction is an approach in evolutionary biology that stresses how organisms adapt to the environment by actively modifying it to best suit their current needs and abilities (Laland, Odling-Smee, & Feldman, 2000). Niche construction and adaptation require balancing a trade-off between sampling the environment – to estimate the current state of the environment – and specialization – to efficiently adapt to the demands of the sampled environment (Frankenhuis & Del Giudice, 2012; Frankenhuis & Panchanathan, 2011; Nepomnaschy & Flinn, 2009). This is not unlike Piaget's ideas of assimilation and accommodation driving developmental change (Piaget, 1952) or the exploration/exploitation computations discussed earlier (Koechlin, 2016).

In the context of brain development, niche construction involves adapting and changing neural pathways such that an organism selects information in line with what they most need and with what they can best process given the abilities and neural architecture available (M. H. Johnson et al., 2015). PFC may orchestrate this process of niche construction through computations that facilitate the construction of rules and norms that guide learning and behavior, as well as by exerting top-down control over activity in other neural regions to align with currently relevant rules, goals, or norms. This aligns with PFC's computation of exerting top-down modulatory control over posterior neural regions (Braver et al., 2009; Buschman & Miller, 2007; B. T. Miller & D'Esposito, 2005; E. K. Miller & Cohen, 2001; O'Reilly, 2006; Rougier et al., 2005).

In turn, niche construction itself is afforded by changes in the integrity of feedforward inputs into PFC, which reflect the development of physical, perceptual, and cognitive skills and abilities that support novel methods for information sampling and interacting with the environment. Thus, PFC developmental change may be driven by adaptation to changes in information sampling, possibly via feedforward connections, and specialization of functions via its feedback connections.

The idea of adaptation involving a trade-off between sampling and specialization is captured in "dynamic optimization" models of development (Frankenhuis & Panchanathan, 2011). These models suggest that individual differences in timing and plasticity of development occur by balancing sampling and specialization. When sampling indicates that the current state of the environment has changed, PFC must then adapt to specialize for the demands of the new environmental state. Models of switching behavior in PFC have offered a similar conclusion with respect to its online role in stability versus flexibility of function. Successful adaptation to changing environments involves balancing cognitive flexibility, or the ability to adjust behavior to align with changing demands and environments, and cognitive stability, or the ability to maintain behavior in the face of distractors (Armbruster, Ueltzhöffer, Basten, & Fiebach, 2012). Switching between cognitive flexibility and cognitive stability requires organisms to monitor changes in the environment to form and maintain attentional biases towards relevant information, which is supported by dopamine-modulated updating and active maintenance in PFC (Armbruster et al., 2012; Fallon et al., 2013; Rosa, Dickinson, Apud, Weinberger, & Elvevåg, 2010).

1.5.1 Neural Adaptation in Response to Sampling Changes via Feedforward Connectivity

There are a variety of causes of neural uncertainty that can drive adaptive responses. The developmental state of neural regions with feedforward connections into PFC may constrain the information available for sampling. Specialization in other regions of the brain as well as the development of new connections to PFC could modify the input that PFC processes, requiring PFC to adapt to different quantity and quality inputs (Amso & Scerif, 2015). One hallmark of early cortical development is the differentiation of various cortical areas into specialized regions (M. H. Johnson & Vecera, 1996). This process results in previously combined information processing streams specializing for particular types of information, leading to less overlap and interference between them. This may lead to apparent 'improvements' in PFC processing, since it is now computing over inputs that have a higher signal-to-noise ratio. As such, the ecological account predicts that some apparent developmental improvements in PFC processing may originate from adaptation to developmental changes in neural regions that provide feedforward input to PFC, rather than structural maturation of the PFC *per se*.

Apparent development improvements in PFC processing may additionally reflect myelination of connections between PFC and other posterior neural regions, which may influence the efficiency or strength of the input to PFC. Myelination of neuronal axons plays an important role in establishing and maintaining rapid and efficient neural communication across development (Deoni et al., 2012). During the first year of life, myelination proceeds rapidly in a posterior to anterior gradient with early sensory regions being some of the first to become myelinated (Deoni et al., 2011, 2012), likely reflecting increased experience with sensory input (Barres & Raff, 1993; Demerens et al., 1996; Stevens, Porta, Haak, Gallo, & Fields, 2002; Wake, Lee, & Fields, 2011). Myelination may affect PFC processing in two ways. First, myelination of feedforward connections to PFC may influence PFC functional development by impacting the quality or signal-to-noise ratio of sensory input to PFC. Second, increased efficiency of interneural communication between PFC and other regions as a result of myelination may increase the quality and integrity of input to PFC, leading to apparent improvements in PFC processing due to changes in the quality of the input, rather than due to PFC maturation *per se*.

Thus far, the discussion has focused on neural-level analysis of PFC adaptation. Here we offer complementary data that suggest that changes in cognitive, behavioral, and motor abilities also alter the availability of information for maintenance, updating, and planning at the level of PFC and executive functions. This point is easily illustrated with data from developmental transitions in locomotor abilities. This body of work indicates that interactions with the environment changes for infants as they develop new or more refined locomotor skills (Karasik, Tamis-Lemonda, & Adolph, 2011). For example, when infants first learn to sit stably on their own, their hands become free to manipulate and use objects. The novel visual information that is generated by the infant's object manipulation supports the development of more advanced abilities such as object memory (Ruff, 1981), object discrimination (Soska, Adolph, & Johnson, 2010), and view-invariant object recognition (James, Jones, Smith, & Swain, 2014). Manipulating objects while sitting also allows infants to bring objects close to their eyes such that the objects dominate their visual field. This creates opportunities for further development, such as learning names of objects that were not as readily available before infants could sit stably (Pereira, Smith, & Yu, 2014; Yu & Smith, 2012). Thus, developing new motor abilities – such as learning to sit stably – creates opportunities for development by changing the information that infants can sample and process from the world. In relation to PFC functional development, these changes require adaptation at the level of executive functions, as PFC adapts to organize and process the influx of information that was previously unavailable. In other words, the ecological account predicts that as new quality or quantity input becomes available, PFC must adapt its computations of active maintenance (Badre, 2008; D'Esposito & Postle, 2015; E. K. Miller & Cohen, 2001; O'Reilly, 2006; Rougier et al., 2005) and rapid updating (Braver & Cohen, 2000; Chatham & Badre, 2015; Chatham et al., 2014; Chiew & Braver, 2017; Cohen et al., 2002; M. J. Frank et al., 2001; O'Reilly, 2006; Rougier et al., 2005) to operate over the new input.

The ecological approach also makes the prediction that some apparent PFC and executive functions 'deficits', that are observed in a number of neurodevelopmental and learning disorders, may develop as *an adaptive response* to early disruptions in physical, perceptual, and motor abilities (M. H. Johnson et al., 2015). In other words, it predicts that early disruptions in perceptual and motor abilities will change the feedforward input to PFC and ultimately the top-down influence PFC exerts through feedback connections in response. We will return to this idea in the niche construction section below.

This focus on sampling in the ecological account has some similarity with a graded-representations account of PFC function, which suggests that PFC's efficiency is dependent on the strength of the representation it is acting on (Morton & Munakata, 2002). A graded-representations account suggests that maturation of PFC increases the strength of representations that can be actively maintained within PFC, thereby leading to improvements in executive functions. The ecological account makes a similar prediction regarding how the strength of active representations influences PFC function; however, instead of suggesting that PFC maturation *per se* increases infants' ability to maintain active representations, we suggest that representation strength is influenced by adaptation to the internal and external environment through changes in feedforward connectivity.

1.5.2 Specialization and Niche Construction Through Feedback Connectivity

Simultaneously, PFC is involved in the feedback orchestration of widespread neural reorganization and specialization of posterior cortical regions. This ontogenetic mechanism aligns with a recent phylogenetic description of brain evolution, which suggests that cognitive control may have emerged from the increasing convergence of hierarchically organized input to PFC across evolutionary time (Finlay & Uchiyama, 2015). A primary feature of vertebrate brains is conservation of a hierarchical organization across the whole brain, with more rostral regions typically exerting modulatory control over more caudal regions. Across developmental time, the human brain also becomes increasingly hierarchical to accommodate the increased levels of control that become necessary as other neural regions, skills, and abilities develop (Supekar et al., 2009).

A prevalent hypothesis that has been articulated previously is that PFC may be a key player in this process of hierarchical reorganization (M. H. Johnson, 2000, 2011; M. H. Johnson et al., 2015; Thatcher, 1992). Even though functional connectivity innervates from back to front, cortical neurogenesis occurs in the opposite direction along a rostral to caudal progression across the entire brain, with frontal neurons becoming differentiated before more posterior neurons (Cahalane et al., 2012); it has been suggested that this byproduct of developmental timing may place PFC in an optimal position to facilitate the organization and development of other cortical regions (M. H. Johnson et al., 2015). In other words, the early differentiation of neurons within PFC, combined with the early formation of feedback pathways from PFC to other cortical regions (Conel, 1939; Rakic, 2002), may allow PFC to exert modulatory control over the development and specialization of posterior neural regions from very early in development. This possibility bears similarity to work indicating that PFC exerts modulatory control over neural activity within posterior regions in adults (Gilbert & Li, 2013; E. K. Miller & Cohen, 2001).

Computational models shed light on how this bears on PFC development. In particular, a knowledge-based cascade correlation model has been used to explain how PFC might orchestrate other neural regions during cognitive tasks (Shultz, Rivest, Egri, Thvierge, & Dandurand, 2007, cited in M. H. Johnson et al., 2015). This model is built using an architecture where one central control network (such as the PFC) orchestrates and recruits previously trained networks when they are required for more complex tasks. This unique architecture helps the model to learn tasks faster and more efficiently than models without a central control network. This is because the central control network in this model is able to recruit other previously trained networks as units when the “skills” or “knowledge” that are represented by those networks are required for a learning task. In the context of PFC functional development, as tasks become more complex and new computational units are needed, functional changes may be observed as PFC adapts to recruit new networks to support the increased task demands (M. H. Johnson et al., 2015; Shultz et al., 2007). In other words, changes may be seen as PFC learns to select and organize appropriate neural regions for new learning problems.

Importantly, we suggest that adaptation to feedforward changes (sampling) and niche construction through feedback changes (specialization) are tightly and intrinsically coupled (Figure 1). Thus, specialization of feedback pathways may influence adaptation by biasing information sampling to align with current goals and demands for learning and behavior. This may reciprocally modify the input to PFC necessitating adaptation at the level of executive functions. In this view, PFC plays an active role in shaping its own development by biasing sampling of inputs to align with relevant demands for learning and behavior, which is similar to prior ecological theories of development (e.g., Gibson, E. J., Pick, 2000; M. H. Johnson, 2000, 2011; M. H. Johnson et al., 2015). We suggest that PFC's functions may gradually stabilize as feedforward connections and the requirements for learning and behavior in the external environment stabilize with development.

1.6 Constraints on Adaptation and Niche Construction via Anatomical Proximity to PFC

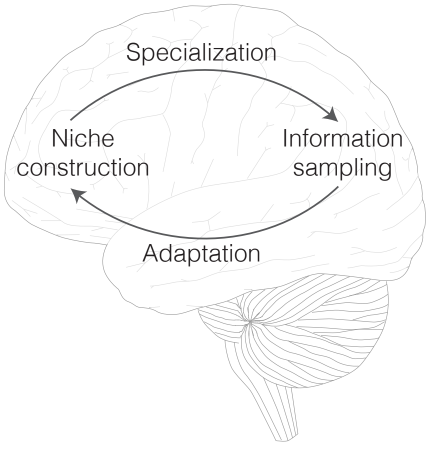

Figure 1. Changes in information sampling necessitates adaptation to new input. Niche construction through specialization of feedback pathways reciprocally biases information sampling to align with current goals.

Timing of adaptation and niche construction across ontogenetic development may be constrained by the anatomical proximity of other neural regions to the PFC. The existing maturationist explanation of ontogenetic brain development suggests that functional development is driven by structural changes in the form of grey matter volume reduction and cortical thinning (e.g., Gogtay et al., 2004; Shaw et al., 2008). Yet, these developmental progressions do not directly map to functional changes. For example, heuristics based on these developmental progressions in grey matter reduction suggested that visual cortex matures first (Gogtay et al., 2004); yet many aspects of visual processing continue to improve throughout childhood (Konrad et al., 2005; Rueda et al., 2004), which is difficult to reconcile with structural maturation accounts. Further, due to constraints in early imaging methodologies, these heuristics were based upon studies of children and adolescents, with little data from ages 4 and under when a majority of fundamental brain development and neural organization occurs.

The ecological account offers a more theoretically structured alternative for control over thoughts and actions based on the ideas discussed above, data on the timing of PFC functional connectivity, and structural proximity of domain specific regions to the PFC. The testable prediction offered by this account is that regions with the closest anatomical proximity and shortest direct projections to the PFC are the first to become specialized for control over action and behavior (Figure 2). In other words, what we sample (and hence specialize) first is based on which regions have the shortest direct connections to the PFC. Thus, regions such as the frontal eye field (FEF) within the frontal lobe may be one of the first regions to specialize, which coincides with the relatively early emergence of oculomotor control (Canfield & Kirkham, 2001), and also relates to PFC activity during saccade planning in 4-month-old infants (Csibra et al., 2001). Similarly, short direct projections between anterior temporal lobe and PFC that support auditory processing and speech perception (Romanski & Goldman-Rakic, 2002; Scott & Johnsrude, 2003) may also specialize relatively early to support language development. This prediction aligns with evidence showing increased PFC activation in infants when hearing infant-directed but not adult-directed speech (Saito et al., 2007), and forward but not backward speech (Dehaene-Lambertz et al., 2002). Other areas within the frontal lobe, such as the premotor area, may begin to specialize for control over more complex motor actions soon thereafter. We predict that regions with more distant connections to the PFC, such as the parietal and occipital lobes, may have the most protracted trajectories for specialization. This prediction coincides with findings showing that functions that depend on frontoparietal and fronto-occipital connectivity, such as visual or executive attention, continue to develop into adolescence (Konrad et al., 2005; Rueda et al., 2004). Thus, when PFC functional development is examined relevant only to frontoparietal change, using a construct of executive functions that is defined relative to an adult state (Dosenbach et al., 2007), it may seem erroneously protracted.

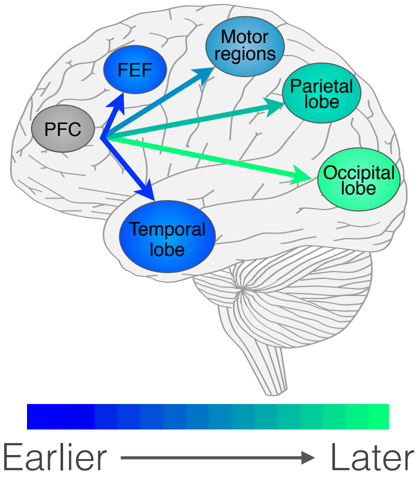

Figure 2. Schematic of the hypothesized developmental trajectory of PFC functional development. Regions with the shortest connections to PFC may specialize for control over action and behavior earlier than more distal regions.

This new pattern of functional developmental change is supported by research showing that PFC's functional connectivity broadly shifts from predominately short-range local connections within the frontal lobe to more long-range connections between the PFC and more distal regions across development (Fair et al., 2007; Gao et al., 2011; Supekar et al., 2009). These shifts are thought to partially reflect changes from segregation to integration with respect to regional functional connectivity, and relate to changes in executive functions. For example, shifts in functional connectivity from short-range intracortical connections within the PFC to long-range frontoparietal connections correlate with improvements in relational reasoning abilities from childhood to adolescence (Wendelken et al., 2016). Similarly, developmental improvements in inhibitory control during an antisaccade task correlate with increases in long-range functional connections between the PFC and subcortical and parietal regions (Hwang et al., 2010). Additionally, other studies show that developmental improvements in cognitive control of eye movements in antisaccade tasks coincide with shifts in functional connectivity from short-range local connections between the dlPFC and FEF in children, to more long-range connections between the dlPFC and visual association cortex in adolescents and adults (Simmonds, 2015). Further, this framework aligns with work indicating anterior to posterior shifts in functional activity for many brain networks (M. H. Johnson & De Haan, 2015). However, note that the magnitude of shifts in short- to long-range functional connectivity in prior reports may have been exaggerated due to age-related differences in motion artifacts (see Power et al., 2012).

1.7 Neural Adaptation and Niche Construction Occur in Response to Shifts in Allostatic Load

Finally, it is necessary to elucidate the hypothesized physiological mechanisms that drive the described PFC changes. In terms of PFC development, the ecological model predicts that increased quantity or quality of multimodal input to the PFC may push the system out of allostasis, which is the maintenance of stable functioning through physiological adaptations (McEwen & Wingfield, 2003). Allostatic overload occurs when energy demand exceeds current supply, which activates adaptive physiological responses to restore a net positive energy balance (McEwen & Wingfield, 2003). Short term fluctuations in function could arise from changes in allostatic load that occur in response to mismatches between the current input to the PFC and the input that the PFC was previously specialized to process at an earlier state. We suggest that this may activate physiological adaptations to increase processing efficiency, such as by promoting experience-dependent myelination in PFC's connections with other neural regions (e.g., Markham & Greenough, 2004), by promoting accelerated development of functional connectivity through mediators such as cortisol (e.g., Callaghan & Richardson, 2012, 2014) or excitatory and inhibitory neurotransmitters such as glutamate and GABA (e.g., Ghisleni et al., 2015; Hensch et al., 1998), or by promoting synaptic pruning within PFC (e.g., Selemon, 2013). These physiological responses may help PFC adapt to process increased or more complex input, helping restore allostasis or a net positive energy balance with time.

Evidence for this idea comes from the data discussed above showing that early life adversity can modulate PFC development as an adaptive response to adverse early environments. For instance, early life stress in the form of maternal deprivation may create an earlier need for PFC to modulate amygdala reactivity due to the lack of external modulation by a maternal caregiver, which may be signaled by increased cortisol levels. This may then push PFC out of allostasis at an earlier stage in development than typical. Accelerated maturation of vmPFC-amygdala circuitry may then occur as an adaptive response to maternal deprivation, via increased levels of cortisol (Callaghan & Richardson, 2012, 2014), thereby helping vmPFC modulate amygdala reactivity in place of an external caregiver (Callaghan & Richardson, 2011; Callaghan & Tottenham, 2016; Gee, Gabard-Durnam, et al., 2013). Similarly, increases in environmental enrichment in low relative to high socioeconomic environments may also result in physiological adaptations due to an abundance of positive acute stressors.

Based on these ideas, the ecological account of PFC development thus makes the counterintuitive prediction that rapid change in one or more developing domains, e.g. vision, language, motor etc., as happens in infancy through childhood, may 'stress' the system sufficiently to result in apparent transient deficits in PFC's functions. These may appear over moments or days, until PFC has time to adapt to managing the rapid influx of information. This transient period of disorganized behavior would theoretically be followed by rapid improvements in executive functions and the newly developed skill. These apparent deficits may take the form of perseverative errors, poor emotion regulation when a child is overwhelmed with information, or they may be a seeming loss of the ability to perform tasks that children had previously mastered.

This pattern is often evident in task performance and on short time-scales. For example, infants make the A-not-B error will eventually adapt to search in the correct location provided a certain number of B trials (Diamond, 1985). Similar findings are seen in deductive reasoning tasks that require participants to relate or integrate a set of variables to verify an abstract rule. When the complexity of information (e.g., number of variables that must be integrated) is parametrically increased, short-term decrements in performance are observed, as well as increased activity within cognitive control networks as PFC adapts to process the more complex information (Cocchi et al., 2014; Shokri-Kojori, Motes, Rypma, & Krawczyk, 2012). Thus, what have been traditionally interpreted as PFC immaturity may reflect processes of adaptation to novel demands on the organism.

1.8 Predictions of an Ecological Model Of PFC Development

Some of the key principles and testable predictions of the ecological account of PFC functional development are briefly summarized here. First, the ecological account posits that PFC performs the same computations across the lifespan, but that these computations are adapted for the learning demands and challenges relevant to an individual in their unique ecological niche. This leads to the first prediction that similar patterns and levels of PFC organization and involvement should be seen in infants, children and adults, but in different tasks, contexts, and activities. Recent advances in imaging technologies, such as NIRS, combined with advancements in machine learning classification techniques for neural data, may allow for direct tests of this prediction by recording and classifying PFC activation in naturalistic contexts across different age groups and in different tasks. For instance, the ecological account predicts that similar patterns of PFC activation may be seen during object exploration or while learning to grasp in infants, during pretend play in children, or while driving a car in adults. Moreover, if executive functions reflect the computations performed by PFC at any point in the lifespan, then a second prediction is that developmental continuity might reflect the efficacy with which PFC's computations are adapted to novel demands across the lifespan. This explanation may apply to the observed relationship between delayed language acquisition and executive functions delays (e.g., Figueras et al., 2008; Henry et al., 2012), as well as early language proficiency in executive functions development in children from different SES homes (Noble et al., 2007, 2005).

Second, the ecological account also posits and predicts that PFC functional development will be reciprocally influenced by changes in feedforward (for adaptation) and feedback connectivity (niche construction). A direct prediction is that early disruptions in perceptual and motor abilities will change the feedforward input to PFC and ultimately the top-down influence PFC exerts through feedback connections in response. While PFC's profuse connectivity, protracted plasticity, and dependence on other neural systems makes it highly adaptable, it also makes PFC highly vulnerable to atypical developmental trajectories due to deviations in species expected environments. Small differences in the experienced environments of typically compared to atypically developing infants and children could create cascading effects that impact PFC development and adaptation of the brain to the environment more generally (Karmiloff-Smith, 2009). This can take the form of altered PFC function from subtle differences in early visual or auditory processing (Baruth, Casanova, Sears, & Sokhadze, 2010), in congenital blindness or deafness (Figueras et al., 2008; Tadic et al., 2009), or from disrupted thalamic organization, such as in individuals with schizophrenia (Cheng et al., 2015).

Finally, the ecological model, in that it relies on the neurobiology of stress-based adaptation, also makes the prediction that changes in the quality or quantity of input to the

PFC will push the system out of allostasis. This leads to the prediction that increased input – whether due to the acquisition of new skills and knowledge over ontogenetic time or due to increased input in a specific task over short time scales – would lead to temporary 'deficits' in PFC function. We noted earlier that these 'deficits' are often observable on short time-scales when the task requires adaptation to a novel demand (e.g., Cocchi et al., 2014; Diamond, 1985; Shokri-Kojori et al., 2012). Future work can test whether deficits may be accompanied by short-term increases in cortisol, reflecting the neurobiological stress response, which may temporarily interfere with the efficiency of neural firing (Arnsten, 2009).

Thus, an ecological model that considers PFC functional development as a process of adaptation to the environment is particularly relevant for understanding mechanisms of typical development, as well as the multiple pathways that might lead to deviations from typical developmental trajectories. A wide and diverse array of neurodevelopmental disorders, including ADHD, autism spectrum disorder, schizophrenia, depression, obsessive compulsive disorder, and anxiety disorders, share a commonality in that they are all frequently associated with atypical PFC functioning. Reconsidering PFC functional development as a process of adaptation and niche construction, and not maturation, provides novel mechanistic insights into the multiple pathways that may lead to these deviations from typical development.

Chapter 2

2. Examining PFC Contributions to Rule Learning in Infancy

Far from having a protracted developmental course, the ecological model of PFC functional development proposed in Chapter 1 argues that the PFC continuously adapts its computations to accommodate the demands present in the changing ecological niche of the growing child. In this view, PFC development is emergent from change in the feedforward inputs from the rest of the brain, the physical structure of the growing body, the content and nature of existing knowledge, and the challenges and opportunities present in the external environment. A central tenet of this ecological model is that the PFC performs the same computations (updating, active maintenance, and modulatory control) across the lifespan. Importantly, this model argues that these computations are adapted for the types of learning and behaviors that are germane to the demands present in an individual's unique ecological niche at each point in development. Thus, the following chapters (2-5) will present illustrative studies designed to test a key prediction of this model: that the PFC is adapted in infancy for learning flexible rule structures that support generalization of past learning in new contexts.

Rules specify relationships between stimuli, actions, and anticipated outcomes (Botvinick, 2008; Botvinick et al., 2009; Badre et al., 2008; Frank & Badre, 2012). For example, a child may learn that when they are indoors, they should use a soft voice, but when they are outside, they can shout. In this example, a stimulus (inside or outside) determines an appropriate action (how the child should speak) and an associated outcome (whether the child is reprimanded). Importantly, rules can be hierarchical, such that they are cued by a higher-order context. Extending this example, a child may learn that the “inside/outside voice” rule only applies if their caregiver is present. Convergent evidence supports a role for the PFC in representing these types of hierarchically-organized rules, with increasingly rostral portions of the PFC representing rules at progressively higher levels of hierarchical organization (Badre, 2008; Badre & D’Esposito, 2007, 2009; Koechlin et al., 2003).

This type of hierarchical organization has two important benefits for learners. First, separating information by higher-order contexts ensures that new information does not interfere with prior learning in other contexts. Second, it allows existing rules to be generalized to novel contexts with similarities to prior contexts. Together, these two benefits have potential to explain the combinatorial explosion in thought and action that is characteristic of early infant development.

As such, the aim of my first study was to examine whether the PFC is adapted in infancy to support learning of flexible rule structures, as predicted by the ecological model proposed in Chapter 1. To test this key prediction, I adapted a canonical hierarchical rule learning task that requires individuals to learn and generalize abstract rules. This work is based off of prior research showing that adults spontaneously extract hierarchical rule structures during incidental learning, which supports generalization in novel contexts (Collins & Frank, 2013; Collins et al., 2014; Collins & Frank, 2016). Importantly, computational models and neuroimaging data in adults suggest that this learning mechanism is supported by the PFC (Collins & Frank, 2014; Frank & Badre, 2012). However, whether infants are capable of using similar learning mechanisms to structure information into hierarchically-organized rules is unknown. As such, I tested whether infants also use this hallmark skill to structure simple audiovisual inputs into generalizable rule structures relevant for learning in infants' environment.

2.1 Eight-Month-Old Infants Spontaneously Learn and Generalize Hierarchical Rules

2.1.1 Abstract

The ability to extract hierarchically organized rule structures from noisy environments is critical to human cognitive, social, and emotional intelligence. Adults spontaneously create hierarchical rule structures of this sort. The developmental origins of this hallmark skill are unknown. In Experiment 1, we exploited a visual paradigm previously shown to elicit incidental hierarchical rule learning in adults. In Experiment 2, we used the same learning structure to examine whether these hierarchical rule learning mechanisms are domain general and can help infants learn spoken object/label mappings across different speaker contexts. In both experiments, we showed that 8-month-olds created and generalized hierarchical rules during learning. Eye blink rate, an exploratory indicator of striatal dopamine activity, mirrored behavioral learning patterns. Our results provide direct evidence that the human brain is predisposed to extract knowledge from noisy environments, adding a fundamental learning mechanism to what is currently known about the neurocognitive toolbox available to infants.

2.1.2 Introduction

Seminal work has shown that infants exploit the statistical properties of the environment to learn visual and auditory information (Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996). Infants can also learn simple algebraic rule structures in patterned sequences of speech sounds, for example, where the abstract rule maps onto an arbitrary rather than concrete class of items (Frank, Slemmer, Marcus, & Johnson, 2009; Marcus, Vijayan, Rao, & Vishton, 1999; Marcus, Fernandes, & Johnson, 2007). However, especially in early postnatal life, infants are faced with complex changing and noisy environments that require learning and action. What mechanisms are available to help young infants transform these signals into organized behavior in the absence of repetitive patterns, cues, or incentives? Here we provide evidence that infants exploit latent hierarchical rule learning mechanisms that to date have been considered characteristic of more mature learning.

This work is drawn from a long line of theoretical and experimental research examining hierarchical action and reinforcement learning. In reinforcement learning, an agent selects among multiple actions in response to stimuli to learn stimulus-action-outcome (S-A-O) contingencies. In a hierarchical framework, these contingencies depend on a higher-order rule-set, which can be cued by multiple contexts. Thus, a hierarchical agent can select the valid S-A-O contingencies in a context-appropriate fashion, and can also transfer those contingencies to novel contexts without having to learn them anew (Collins & Frank, 2013; Collins, Cavanagh, & Frank, 2014; Donoso, Collins, & Koechlin, 2014; Frank & Badre, 2012; Monsell, 2003). This hierarchical framework is domain general and can apply to rule learning across multiple classes of stimulus inputs. For example, a child growing up in a bilingual environment may learn that when she is with her mother (context), she should expect to hear and respond in English (the rule-set), but when she is with her father, expect to hear and respond in Spanish. Thus, the child may use a higher-order context (mother or father) to determine the appropriate lower rule-set to use (language which then specifies object/label mappings, Figure 1A). This mechanism could then help infants learn and separate multiple languages without having to experience every word in each speaker context. Importantly, in this framework the rule-sets are distinct from the contexts that cue them. Therefore, the infant may learn that the context "grandmother" is also associated with the "Spanish" rule-set. Then, when she hears her grandmother use "botella" for bottle, she can immediately infer that her father, but not her mother, will also respond to "botella", even if she had never encountered a bottle with her father. Note that this example describes a different type of hierarchy than those found in linguistic structures, such as embedded clauses in syntax (Chomsky, 1988). Instead, it describes a domain general rule learning approach based on higher-order contexts governing lower-level rule structures.

Previous work in adults shows that hierarchical organization has a dual learning benefit (Collins & Frank, 2013; Collins et al., 2014). First, using higher-order contextual information to specify lower-order rule-sets helps adults structure learning and behavior in such a way that learning new information does not interfere with behaviors learned in *other* contexts. Extending the example of a child in a bilingual home, receiving the label "cat" and the label "gato" should not interfere as long as the labels are governed by unique higher-order rules (as in Figure 1A). Simultaneously, the rule-sets are *latent:* they are not tied to a specific higher-order context and can thus be transferred to novel contexts when useful. Further, novel stimulus-action-outcome contingencies can be appended to latent rule-sets (e.g., one can always learn a new object/word label in an existing language). Hierarchical organization of this sort (Collins & Frank, 2013; Collins et al., 2014) is incidental and automatic during learning, raising the possibility that it may be functional early in life. However, computational models and EEG data suggest that this type of learning depends on hierarchical nesting of dopamine innervated frontostriatal loops (Badre & Frank, 2012; Collins & Frank, 2013; Collins et al. 2014), the anterior components of which are involved in motor action selection and thought to be underdeveloped in infancy (van Hofsten, 2004). We thus assessed whether infants exhibit a predisposition for organizing behavior into latent rule-sets using an oculomotor task environment that does not require motor action selection. Since, infants are capable of attention-guided oculomotor control by roughly 6-8 months of age (Amso & Johnson, 2006, 2008), we predicted that in environments that involve oculomotor responses, infants might also automatically apply hierarchical structure to facilitate learning and generalization across contexts.

We adapted a canonical adult paradigm for assessing incidental hierarchical rule learning (Collins & Frank, 2013; Collins et al., 2014). In Experiment 1, we used this task to ask whether 8-month-olds spontaneously apply hierarchical structure to organize visual information (Figure 1B). In Experiment 2, we used an identical hierarchical learning structure (Figure 1B & C) to test whether this mechanism is useful for word learning, a relevant domain for young infants (e.g. Xu, Cote, & Baker, 2005). We tested the idea that if hierarchical rule-sets are latent, then one should be able to append a novel object label to an existing rule-set (i.e. language) and then transfer it back to other speakers associated with that language, thereby helping infants learn multiple languages without having to concretely experience every word in each speaker context. The juxtaposition of the two experiments asks whether these hierarchical rule learning mechanisms operate across inputs from multiple domains including visual, auditory, and multisensory information (Figure 1). As noted, hierarchical structure learning is thought to depend on dopamine innervated frontostriatal loops (Collins & Frank, 2013; Collins et al., 2014). Therefore, we also measured infants' eye blink rate (EBR) as an exploratory measure of dopamine activity. EBR is thought to be an indirect marker of striatal dopamine activity in infants (Bacher & Smotherman, 2004) and adults (Karson, 1983), and has been implicated in similar cognitive control and rule learning tasks in adults (Dreisback et al., 2005; Muller et al., 2007).

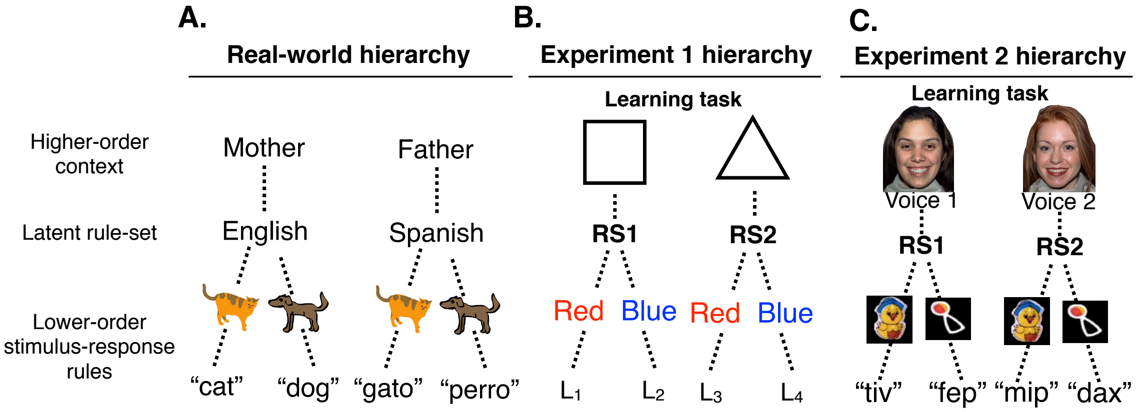

Figure 1. Examples of hierarchical structures in (a) a real-word context and in the learning tasks from (b) Experiment 1 and (c) Experiment 2. During development, children may learn that specific higher-order contexts are associated with distinct rule sets that determine lower-order stimulus- response rules. For example, a child raised in a bilingual environment may come to expect that each parent will speak in a different language and, therefore, different words will be used to label the same objects. This mechanism was manipulated in two experiments. Experiment 1 used a visual hierarchical structure, in which two higher-order shapes each cued a separate rule set that dictated which quadrant (Q) of the screen the shape would appear in, given its color. Experiment 2 used a word-learning hierarchical structure, in which two higher-order face-voice combinations each cued a separate rule set that dictated which artificial words a pair of animated toys were associated with.

2.1.3 Experiment 1

2.1.3.1 Method

2.1.3.1.1 Participants

Twenty healthy 8 month-old infants (8 females; M = 8.5, SD = 1.00) were recruited via advertisements and birth records. Sample size was determined based on similar studies in our lab that used the same age group. We continued collecting data until we reached this target sample size. An additional nine infants were tested, but data were discarded due to fussiness or crying (n = 5), technical/experimenter error (n = 3), or parental interference (n = 1). All families were compensated for participating in the study.

2.1.3.1.2 Materials

We used eye tracking to streamline calculation of speed of eye movements or reaction times, to target locations. Infants sat on their parents' laps approximately 75 cm from a 22" monitor in a dark room. Infants' eye movements were recorded using remote eye tracking software (SMI RED system) and the task was presented using E-Prime software. Before the study began, infants' point of gaze (POG) was calibrated by presenting two target stimuli in the upper left and lower right corners of the monitor. The POG was validated by presenting four stimuli in each corner of the monitor. The reaction times were calculated from trial onset (presentation of the center cue) and ended when the infant's point of gaze arrived at the target location. Target locations were defined in SMI native analysis software package BeGaze and encompassed the target location stimulus (Figure 2A, red box surrounding toys). Eye blink rates per trial were recorded and computed using SMI RED and native software.

2.1.3.1.3 Procedure

2.1.3.1.3.1 Task overview. Eight month-old infants participated in a Learning Task and a Generalization task, during which they saw cue/target location pairings (Figure 2A). The central cues varied by both shape (e.g. square or triangle) and color (red or blue), and the target location consisted of an animated toy presented in one of four quadrants on the screen (Figure 2B). These pairings could simply be learned as individual associations between the central cues and the target locations. Alternatively, infants could apply a hierarchical structure to learn the pairings (as depicted in Figure 2C) and as previously observed in adults (Collins & Frank, 2013; Collins et al., 2014). In this case, one dimension

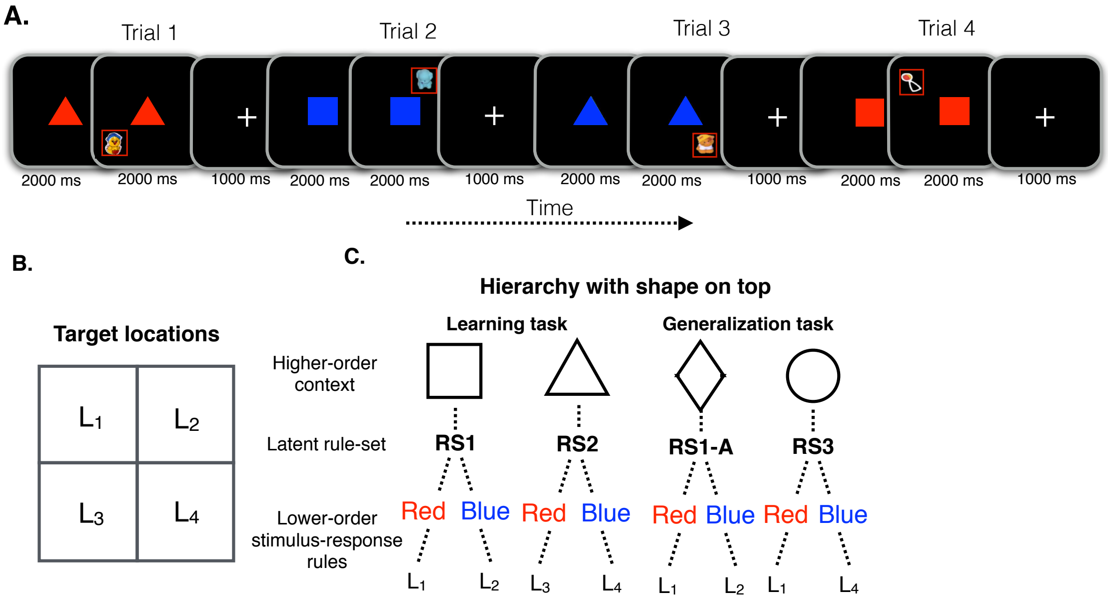

Figure 2. Sample trial sequence and paradigm from Experiment 1. Each trial in the learning task (a) began with a centrally presented cue that varied in color (red or blue) and shape (square or triangle). Then an animated toy (the target) appeared in one of four quadrants of the computer screen (b). Eye movements were measured to determine how quickly infants looked toward the quadrant containing the target stimulus (highlighted here by the red box). Infants could use shape as a higher-order context to cluster the pairings into latent rule sets specifying lower-order color/target- location rules (c). The generalization task was similar to the learning task, except that the shapes were a diamond and a circle. The color pairings for one shape were the same as in the learning task, but the color pairings for the other shape required a new rule set.

(shape) is used as a higher-order context that cues a latent rule-set, which then groups together simpler rules between the lower-order feature (color) and the target location.

After the Learning task infants saw two novel shape contexts during the subsequent Generalization task. The task is designed such that if infants learned latent rule-sets, they could subsequently transfer these rule-sets to novel contexts (e.g. novel shapes) during the Generalization task, as evidenced by faster learning of a rule-set that analogously groups together the same color-location associations that comprised an existing set, as compared to a control rule-set which also involves previously experienced color-location pairings but not in a coherent set.

The mappings between rule-sets, shapes, and target locations were counterbalanced. The dependent variables were 1) reaction time (RT) to the location of the target location (animated toy) and 2) eye blink rate (EBR) per trial. RT was calculated as the time it takes to arrive at the target location. The calculation began at trial onset (presentation of the center cue) and ended at the time the point of gaze arrived at the target location. EBR was calculated as the average number of eye blinks per trial and was computed per trial in BeGaze. We predicted that if infants were learning the pairings, they would show a reduction in RTs with trial exposure, indicating that they are learning to correctly predict or anticipate the target location after the onset of the cue (Canfield & Haith, 1991).

2.1.3.1.3.2 Learning task. During the Learning task, we presented infants with four cue-target location pairings, where the central cues varied by shape (e.g. square or triangle) and color (red or blue). The target location consisted of an animated toy presented in one of four screen quadrants (Figure 2B). In principle, the cue/target location pairings could be learned efficiently as four separate shape-color-target location rules with no latent hierarchical structure. In this case, the two dimensions (color and shape) of each cue are used in conjunction as a single state, with no privilege given to either shape or color. Thus infants might learn the following rules: a red circle means the toy will appear in quadrant 1, a blue circle means quadrant 2, a red square means quadrant 3, and a blue square means quadrant 4.

Alternatively, infants could apply a hierarchical structure to learn the pairings (as depicted in Figure 2C) and as previously observed in adults (Collins & Frank, 2013; Collins et al., 2014). In this case, one dimension is used as a higher-order context that cues a latent rule-set, which then groups together simpler rules between the lower-order feature (e.g. color) and the target location. Thus, infants might learn sets of rules during the Learning task as follows: if the higher-order context "shape" is a triangle, then the "color" red means that the toy will appear in quadrant 1 *and* blue means the toy will appear in quadrant 2 (rule-set 1, RS1). If the higher-order context "shape" is a square, then different rules apply: in this case, red predicts the toy appearing in quadrant 3 and blue in quadrant 4 (RS2). If infants learned in this hierarchical format, we predicted that they would more likely use shape rather than color as a higher-order context, based on pilot data and the known shape bias in infants and children (e.g. Graham & Diesendruck, 2010; Landau, Smith, & Jones, 1988). The Learning task was designed such that there were no clues or incentives offered to structure the input in a hierarchical format. While there is no immediate benefit to creating this sort of hierarchical structure, computational models and adult work have shown that it affords future generalization opportunities (Collins & Frank, 2013; Collins et al., 2014).

Infants received a total of 10 trials per rule-set during Learning (where each trial per rule-set consisted of presentation of two stimuli from the rule-set). The presentation order of the stimuli was intermixed and pseudo-randomized, with the constraint that the randomization resulted in an equal number of trials where the shape changed and trials where the color changed from one trial to the next. During each stimulus presentation, the central cue was shown for 2000 ms, followed by presentation of the animated toy associated with the central cue for 2000 ms (Figure 2A). The central cue remained on the screen while the animated toy was presented. There was a 1000 ms inter-trial interval between each stimulus presentation. We binned every two consecutive trials per rule-set to create five learning blocks for each rule-set. We measured RTs (beginning at cue onset) to the target stimulus and define learning as a speeding of RTs with trial exposure.

2.1.3.1.3.3 Generalization task. Immediately after the Learning task, infants saw four new cue/target location pairings. These new pairings were associated with the same colors (red or blue), but they were composed of new shapes (e.g. diamond or circle). These novel pairings could again be grouped by shape to form rule-sets (Figure 2C). One such rule-set (RS1-A) was the same as a previously learned rule-set (RS1: red- quadrant 1, blue-quadrant 2). The other rule-set (RS3) consisted of two color-target location rules that had both been experienced individually before, but across different rule-sets (RS1 and RS2). If infants learned latent rule-sets that were not tied to the particular shape contexts that they were learned in, then we predicted that this should facilitate positive transfer for one novel context (RS1-A) and negative transfer for another (RS3). That RS3 preserved the same color-target location rules from Learning was a control that lower level stimulus-response learning did not drive generalization performance. That is, it ensured that any difference in learning of RS1-A and RS3 could only be due to transfer of the *set* of color-target location rules, rather than *individual* color-target location rules. Therefore any benefit to learning RS1-A over RS3 can only be attributed to participants having created latent rule-sets during learning that can then be generalized across shape contexts.

It is also critical to note that generalization could only occur if infants used one dimension (shape) as a higher-order context that cues a latent rule-set, which then groups together a set of lower-order (color) target location pairings. If infants only used shape, then there would be no generalization at test since both shapes were entirely novel. If infants only used color, then generalization would occur in both conditions at test, since both conditions preserved the color-target location associations from the Learning task. Thus, if we only observed generalization in an analogous rule-set, we can confidently adopt a model in which infants created a latent hierarchical structure during learning, and then reused this structure to support learning in a novel context.

Infants again received 10 pseudo-randomized trials per rule-set, analogous to the Learning task, and RTs to the target locations from cue onset were measured. We again binned every two consecutive trials per rule-set to create five learning blocks. We measured learning as change in infants' RTs with trial exposure. If infants learned a hierarchical structure, we predicted to find faster RTs (positive transfer) for the consistent rule-set (RS1-A) and slower RTs (negative transfer) for the novel rule-set (RS3). If infants did not learn a hierarchical structure, we expected to find no differences between learning of these rule-sets during the Generalization task.

2.1.3.2 Results

2.1.3.2.1 Learning Task Behavioral Performance

We measured learning as change in infants' eye movement reaction times (RTs) from cue onset to arrival at the toy location. RTs significantly decreased over trial exposure, for both ostensibly formed RS1 and RS2, indicating that infants were anticipating or predicting the correct quadrant after cue presentation more quickly with exposure to both rule-sets, $F(4,76) = 6.221$, $p < .001$, $\eta_p^2 = .247$ (Figure 3A).

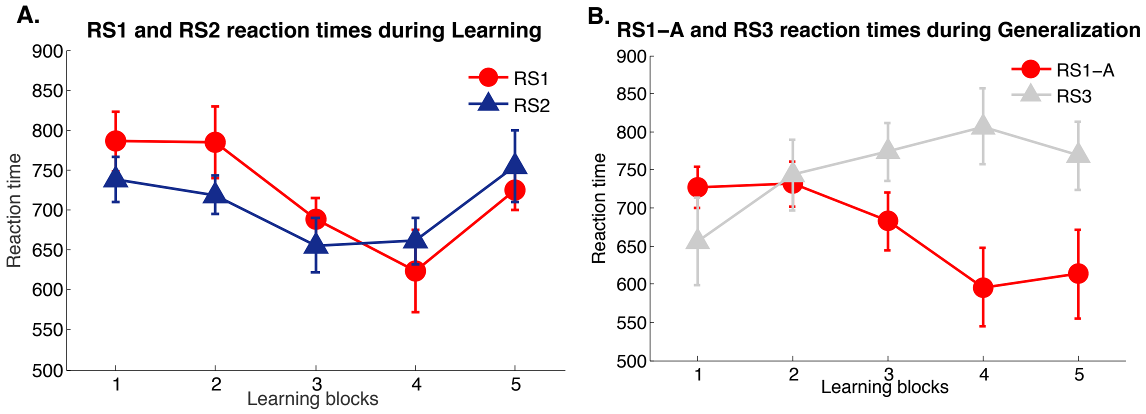

Figure 3. During the Learning task (A) exposure to both rule-sets resulted in learning, as measured by increasingly rapid anticipation of target appearance. During the Generalization task (B), infants generalized learning of rule-set RS1 to RS1-A, as indicated by faster learning of RS1-A than RS3. Error bars indicate SEM.

Notably, the majority of infants provided evidence of learning a hierarchical rule structure as in Figure 2C, rather than individual shape-color-target location rules. In analogous tasks in adults, RT costs are commonly observed when the higher-order rule switches on a trial basis and thus has to be updated into working memory (Collins & Frank, 2013; Collins et al., 2014; Monsell, 2003). We reasoned that if infants learned rule-sets based on shape as in Figure 2C, then we would expect slower RTs (a switch cost) when the shape rule switched on consecutive trials (indicating a switch to a different rule-set, or group of color-target location pairings) relative to when the shape rule repeated (indicating the same rule-set, or lower-level color-target location pairings as the previous trial), regardless of color. We calculated two switch cost values, one assuming a higher-order shape structure (RT shape switch - RT shape repeat) and, as a fidelity check, a second assuming a higher-order color structure (RT color switch – RT color repeat). Fifteen (of 20) infants had a greater (more positive) cost to shape rule switches than color rule switches (Sign Test, $p = .041$). Additionally, these shape rule switch costs were significantly greater than zero ($t(14)=2.657$, $p = .019$; Mean cost = 27.04 ms). These data provide the first evidence that infants are establishing a hierarchical rule structure from ambiguous input, as indicated by a selective RT cost related to updating of higher-order rules into working memory.

2.1.3.2.2 Learning Task Eye Blink Rate

Neuroscience and computational modeling research provides mechanistic evidence that the formation of hierarchical rule structures is supported by interactions between the prefrontal cortex (PFC), striatum, and their dopaminergic innervation in adults (Collins & Frank, 2013; Collins, et al., 2014; Frank & Badre, 2012). These models posit that frontostriatal loops are hierarchically nested, such that a context cues a higher-order level that selects the appropriate rule structure, which in turn constrains a lower stimulus-response selection level (Collins & Frank, 2013; Frank & Badre, 2012). Learning of which rule structures apply is thought to rely on dopaminergic signals in frontostriatal pathways. We used this information to generate a prediction about a physiological indicator of striatal dopamine function, namely eye blink rate (EBR) (Blin, Masson, Azulay, Fondarai, & Serratrice, 1990; Karson, 1983; Kleven & Koek, 1996; Taylor et al., 1999). In adults, increased EBR is correlated with better performance in cognitive control tasks that require updating of rule representations into working memory (Dreisback et al., 2005; Muller et al., 2007), where this same updating function is related to striatal activity (Collins & Frank, 2013; Frank & Badre, 2012; Frank, Loughry, & O'Reilly, 2001; McNab & Kingberg, 2007). Infants also show increased EBR during feeding and presentation of novel stimuli (Bacher & Smotherman, 2004), both of which are modulated by dopamine agonists (Pitts & Horvitz, 2000). These data hint at a link between EBR and dopamine activity as early as infancy. Therefore, we used this information to generate the prediction that this EBR measure would only be engaged on precise trial types relevant to switching the higher-order rule. We compared infants' EBR on trials where the shape switched (and color stayed the same) to infants' EBR on trials where the shape repeated (and color again stayed the same). We controlled for color switches in this way to ensure that any difference in EBR was only due to changes in the higher-order shape rule, and not to factors related to color switches, such as a change in luminance. We found that trials where the shape switched—indicating a switch to a new rule-set—elicited more eye blinks than trials where the shape rule repeated, specifically during the second half of the Learning task, $F(1,19) = 11.262$, $p = .003$. EBR for shape-switch vs. repeat trials was not different during the first half of Learning, $t(19) = 1.259$, $p = .223$. However, by the second half of Learning, when the rule-sets are learned (see Figure 2A), EBR for shape-switch trials was significantly higher than EBR for shape-repeat trials, $t(19) = 3.951$, $p = .001$ (Figure 3). As a control, we ran the same analysis assuming a higher-order context of color, and controlling for changes in shape, and found no differences in EBR for color-switch vs. repeat trials, $F(1,19) = .531$,

$p = .475$. We take this exploratory measure as suggestive that the neural system supporting this learning in infants may engage dopamine-dependent mechanisms.

2.1.3.2.3 Generalization Task Behavioral Performance

Finally, and most importantly, we found that infants treated these hierarchical rule-sets as *latent*, meaning that they were not tied to the particular shape contexts and could be generalized to novel contexts. In the Generalization task, we presented the same infants with two novel shape contexts. One of these (RS1-A) had the same set of rules governing color-target location pairings as a rule-set from the first task (RS1), except with a novel shape. The other control shape signified two lower-order rules (RS3) that had each been experienced individually before, but across different rule-sets, thereby controlling for simple low level stimulus-response learning (Collins and Frank, 2013). Hence any difference in performance between RS1-A and RS3 can only be related to the recognition of the latent rule-sets across contexts. We found that infants indeed reliably transferred the rule structure from RS1, as indicated by faster learning of the analogous rule-set (RS1-A) relative to the novel rule-set (RS3), $F(4,76) = 4.102$, $p = .005$, $\eta_p^2 = .178$ (Figure 3B). This positive transfer is consistent with the prediction that infants built rule-sets during Learning and reused one of these rule-sets to support learning in a novel context. In contrast, the relative slowing of RTs for RS3 may be indicative of negative transfer (Collins & Frank, 2013): RS3 pairings involved individual rules that reminded them of either RS1 or RS2; hence an incidental tendency to apply hierarchical structure would lead to incorrect predictions and slower RTs.

2.1.3.3 Experiment 1 Discussion

These results provide the first evidence that infants create hierarchical rule structures during incidental learning. Although it is possible that infants could have learned the pairings using alternate mechanisms, such as statistical learning, this is an unlikely explanation, as our input did not contain a statistical or patterned structure that infants could exploit to facilitate generalization in novel contexts. Infants also could not have used simple associative mechanisms to facilitate learning in a novel context, since learning of the analogous and novel rule-sets should be equivalent during the Generalization task if this were the case. That infants learned the analogous rule-set faster than the novel rule-set, along with a significant reaction time cost to higher-order rule switches, is strong evidence that infants created and reused a hierarchical rule structure.

2.1.4 Experiment 2

Experiment 2 replicated this result using the same basic hierarchical structure as Experiment 1, but different learning and response requirements (Figure 1). We examined whether this mechanism is useful for word learning – a domain relevant to 8-month-olds, and whether such a mechanism would support the ability to append novel lower-order contingencies (object-label pairings) to existing latent rule-sets (languages). Infants use several mechanisms to facilitate language acquisition and word learning, including statistical learning to segment words from strings of syllables (Saffran et al., 1996; Kirkham et al., 2002), and abstract rule-based mechanisms to form simple rules from syllable sequences (Frank et al., 2009; Marcus et al., 1999; Marcus et al., 2007). However, infants have difficulty extracting statistical regularities when more than one artificial statistical grammar is presented in the same auditory sequence without explicit cues to indicate a change to novel statistical structure (Gebhart, Aslin, & Newport, 2009). Monolingual 12-month-old infants are also unable to simultaneously learn two separate abstract rule structures (e.g. “AAB” and “ABA” patterns) from syllable sequences using simple first-order rule learning mechanisms (Kovacs & Mehler, 2009). Yet, bilingual infants are capable of learning multiple languages (e.g. Genesee & Nicoladis, 2007; Person, Fernandez, & Oller, 2003) and appear to reach language acquisition milestones at similar ages as their monolingual counterparts (e.g. Petitto et al., 2001). This suggests that infants might have access to additional learning mechanisms that help them learn and separate multiple languages across contexts. We examined here whether hierarchical rule learning mechanisms serve this goal. In Experiment 2, we tested 8-month-olds using an identical hierarchical structure as in Experiment 1 (Figure 1C). We designed Experiment 2 to be similar with respect to the hierarchical learning structure that could be formed, but unique with respect to the response requirements as well as the information to be learned. We sought to verify the domain generality of this mechanism and especially that it was not constrained by the visuospatial dimensions and oculomotor response requirements of Experiment 1.

2.1.4.1 Method

2.1.4.1.1 Participants

The final sample consisted of twenty-two healthy 8-month-old infants (9 females; $M = 8.5$, $SD = 1.03$). An additional five infants were tested, but data were discarded due to fussiness or crying (n = 4) or parental interference (n = 1). All families were compensated for time and travel.

2.1.4.1.2 Procedure

2.1.4.1.2.1 Task overview. Infants were familiarized with several trials that consisted of a face followed by audiovisual toy/word pairings during a Learning task and a Generalization task (Figure 4A and B). Infants experienced a face on the left half of the screen followed rapidly by a toy on the right of the screen being labeled by a female voice. The faces were two discriminable female faces and the toys were two different animated toys. Four monosyllabic words (jic, mip, dax, and tiv) were used, which were uniquely spoken by two female speakers (e.g. Xu et al., 2005).

The Learning task was constructed such that infants could use the faces/voices as higher-order contexts to create two rule-sets (Figure 4A). Critically, the same two toys were used (e.g. cartoon duck and rattle) in both rule-sets. However, each face/voice higher-order context labeled the toys using different words, thereby creating RS1 and RS2, akin to learning in a bilingual environment (Figure 1). As in Experiment 1, infants could simply learn four associations, but this would not allow them to transfer rule-sets or pass the upcoming inference test. Specifically, as in Experiment 1, the Generalization task was designed so that infants could reuse a rule-set structure that was identical to one shown during the Learning task (e.g. RS1-A). This rule-set could either now be transferred to a novel face/voice higher-order context or it could be relearned as a novel set of simple associations (Figure 4). We also added a novel toy/word to the rule-set that was not previously experienced as part of the analogous RS1 during Learning. The critical test in this experiment is whether they would now transfer the novel word to the appropriate face/voice higher-order context originally experienced during the Learning phase of the task, that is whether they appended a novel association to an existing latent rule-set. Thus, we included a final Inference test trial where we paired the higher-order face/voice contexts from Learning with the novel toy/word presented as part of RS1-A during Generalization, and examined looking time when the new toy/word was paired with the consistent (RS1) versus inconsistent (RS2) face/voice context from Learning (Figure 4C). We predicted that if infants formed hierarchical rule-sets using the face/voice as a higher-order context, then they would look longer at the inconsistent trials that violated the learned rule-set structure. If infants did not adopt a hierarchical rule-set structure, then we expected to find no difference in looking time between the consistent and inconsistent trials.

The mappings between faces, voices, toys, and words were counterbalanced. The dependent measures were the average looking time to the consistent vs. the inconsistent

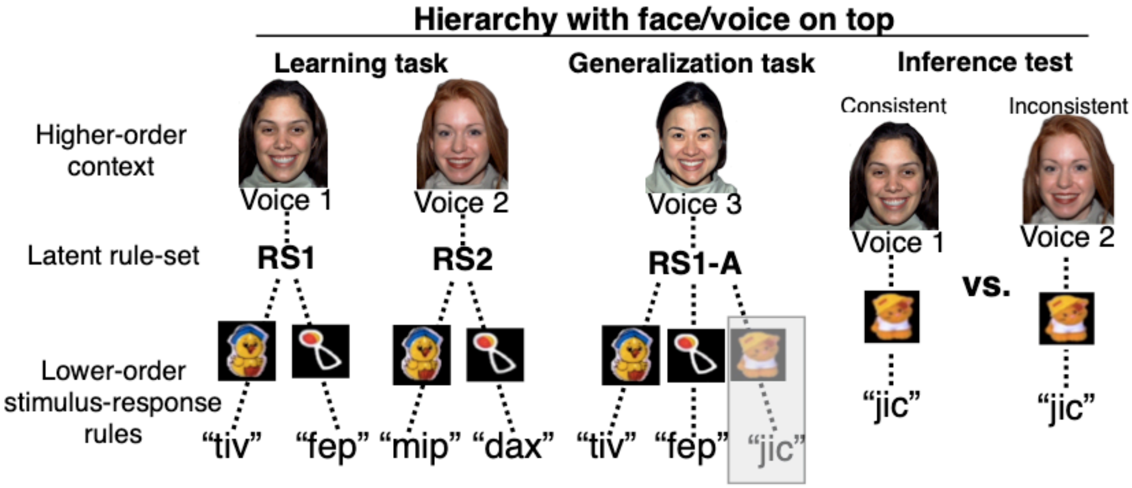

Figure 4. In the Learning task, infants could use face/voice as a higher-order context to cluster the toy/word pairings into rule-sets. In the Generalization task, infants were shown a learned rule-set now associated with a novel context (face/voice); an additional toy/word was also added to the set. During the Inference test, infants were shown pairings that were consistent and inconsistent with the rule-set structure. All mappings between faces, voices, toys, and words were counterbalanced.

face/voice context during the Inference test trial and the average eye blink rate (EBR) during the Learning task as in Experiment 1.

2.1.4.1.2.2 Learning task. During the Learning task, infants saw four different face-voice/toy/word pairings. Two female faces and voices, two toys, and four words were used in these pairings. All mappings between faces, voices, toys, and words were counterbalanced. The pairings were constructed such that each face was associated with a unique voice (e.g. face 1 was always associated with words spoken by voice 1 and face 2 was associated with words spoken by voice 2). Both faces/voices were associated with the same two toys (e.g. both face 1 and face 2 were paired with a cartoon duck and rattle); however, each face/voice used different words to label the toys, similar to a bilingual environment. Infants received a total of 32 trials during the Learning task (8 trials per pairing). During each trial, infants would first see the face presented on the left side of the screen. After 500 ms, a toy appeared on the right side of the screen for an additional 1500 ms, while a recorded female voice said the artificial word associated with the pairing. There was a 1000 ms interval between trials. The pairings could be learned simply as individual face-voice/toy/word associations, using simple associative learning mechanisms. Alternatively, infants could use the faces/voices as higher-order contexts to learn the pairings as rule-sets (RS1 and RS2) grouping together simpler toy/word rules or associations (Figure 4).

2.1.4.1.2.3 Generalization task. Immediately following the Learning task, we presented infants with three new face-voice/toy/word pairings. These pairings could again be grouped by face/voice to form a rule-set identical to one experienced during the Learning task (e.g. RS1-A); however, this rule-set was now associated with a novel higher-order face/voice context. Additionally, one novel toy/word association was added to the rule-set (Figure 4). Infants again received 8 trials per pairing, analogous to the Learning task, for a total of 24 trials.

2.1.4.1.2.4 Inference test. After the Learning and Generalization tasks, infants saw the faces/voices from the Learning task paired with the novel toy/word association from the Generalization task. One of these face-voice/toy/word pairings was consistent with the rule-set structure formed during the tasks, whereas the other pairing was inconsistent with this rule-set structure (Figure 5A). Infants received two consistent trials and two inconsistent trials during the Inference test. The order of the consistent and inconsistent test trials was intermixed and counterbalanced across subjects. During each test trial, infants saw the face and toy, while a recorded voice said the word associated with the toy once every 3 seconds. Each trial continued until the infant looked away for more than 2 seconds, for a maximum of 60 seconds. The dependent measure was the average looking time to the consistent trials compared to the inconsistent trials.

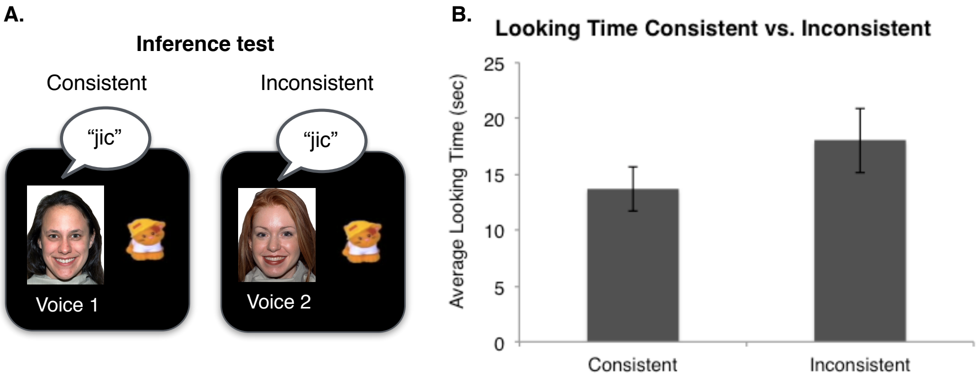

Figure 5. During the Inference test, infants saw pairings that were consistent and inconsistent with the hierarchical structure (A). Infants looked longer at the inconsistent pairing that violated the hierarchical structure (B).

2.1.4.2 Results

To determine whether there were any differences in looking time between the consistent and inconsistent trials, we conducted a two-tailed paired samples t-test, which indicated that infants looked significantly longer at the inconsistent pairing compared to the consistent pairing, $t(21) = 2.461$, $p = .023$ (Figure 5B).

2.1.4.2.1 Learning Task Eye Blink Rate

We had the opportunity to ask whether Experiment 1 EBR finding replicates here. We examined differences in EBR for trials where the higher-order face/voice rule switched from the previous trial—indicating the need to update the current rule-set in working memory—compared to trials where the higher-order face/voice rule repeated during the Learning task. We conducted a 2 (Rule Switch vs. Repeat) x 2 (First vs. Second half of Learning) repeated measures ANOVA. Replicating the findings from Experiment 1, we identified a time by trial type interaction $F(1,21) = 7.47$, $p = .013$. EBR for face-switch vs. repeat trials was not different during the first half of the Learning task, $t(21) = 1.16$, $p = .26$. However, by the second half of Learning, EBR for face-switch trials was significantly higher than EBR for face-repeat trials, $t(21) = 2.96$, $p = .008$.

2.1.5 General Discussion

Across two experiments, we showed that infants spontaneously adopt hierarchical rule-set structures during incidental learning. Importantly, our findings from both experiments suggest that these rule-sets were not tied to a particular context but were instead *latent*, as evidenced by the finding that infants could generalize the sets to support learning in novel contexts – over and above the generalization that might be expected from low level stimulus-response learning. Eye blink rate, an exploratory physiological indicator of dopamine activity in infants (Bacher & Smotherman, 2004) and adults (Karson, 1983), mirrored the behavioral findings.

Prior research shows that infants exploit several mechanisms to drive learning in environments that have a statistical or patterned structure (Kirkham et al., 2002; Saffran et al., 1996). Infants can also learn and generalize simple first-order rules from patterned auditory sequences (Frank et al., 2009; Marcus et al., 1999; Marcus et al., 2007). However, these mechanisms are unlikely to account for our findings, as neither of our experiments contained a statistical or patterned structure that infants could exploit to facilitate learning. Infants also could have learned the cue-target location pairings in Experiment 1 and the face-voice/toy/word pairings in Experiment 2 using simple associative learning mechanisms. However, if this were the case, then performance during the Learning task and Generalization task should be equivalent in Experiment 1, and looking time to the consistent and inconsistent trials should be equivalent in Experiment 2. That infants exhibited faster learning of an analogous rule-set during the Generalization task in Experiment 1 is clear evidence that infants spontaneously constructed rule-sets during initial learning and flexibly reused one of these sets to facilitate learning in a novel context. Experiment 2 replicates and extends these findings by showing that hierarchical rule-sets are latent. Specifically, 8-month-old infants were able to append a novel object label to an existing rule-set during Generalization and then transfer the novel item back to other higher-order contexts associated with that rule-set. This type of mechanism may thus help infants learn multiple languages without having to experience every word in each speaker context.

Finally, we offer here speculation about the neural systems involved in this learning. Evidence from computational modeling and neuroscience research suggests that hierarchical rule learning is supported by dopamine-innervated pathways between the PFC and striatum (Collins & Frank, 2013; Collins, et al., 2014; Donoso et al., 2014; Frank & Badre, 2012). Clearly, it may well be that some other neural system supports hierarchical rule learning observed in our infant sample. However, our data showing that infants have a greater switch cost for the higher-level dimension, paired with higher EBR specifically on rule switch trials, is remarkably consistent with behavioral patterns traditionally associated with frontostriatal working memory updating mechanisms. While the PFC does not reach maturity until adolescence (e.g., Huttenlocher, 1979), the basal ganglia show relatively high functionality, as measured by glucose metabolism, compared with most of the cerebral cortex in the newborn period (Chugani, 1996). One hypothesis then is that frontostriatally-mediated hierarchical rule learning may be dependent on frontostriatal loops in infancy in such a way that weights striatal involvement more heavily than prefrontal involvement. Thus, these frontostriatal loops perform similar computations across the lifespan, but on inputs that are appropriate to learning in infants' unique ecological niche (Rovee-Collier & Cuevas, 2009). As the individual's ecological niche changes and adapts across development, this mechanism may then be coopted to support increasingly complex tasks, such as cognitive control of complex thought and action. Another possibility is that this form of hierarchical learning may require less anterior frontal involvement than adult versions of the task. In adult work, participants must learn correct responses through reinforcement. In contrast, infants are simply shown the toys in the target locations, which directly indicate the correct actions (e.g. screen quadrants to direct gaze to). This form of hierarchical rule learning may require less extensive anterior PFC processing than adult reinforcement learning versions, since infants do not have to learn to select motor actions using reinforcement learning. As such, the PFC may not need to be fully developed to support hierarchical rule learning as tested here using an oculomotor task environment. Future work using computational and neuroimaging tools appropriate to infants will bear directly on these questions.

Chapter 3

3. Neural Underpinnings of Hierarchical Rule Learning in Infancy

In Chapter 2, I examined whether the PFC is adapted for learning in infancy by behaviorally testing whether infants use PFC-dependent learning mechanisms to organize arbitrary visual and auditory inputs into hierarchical rule structures. In a first experiment, I adapted a simple visual paradigm known to elicit spontaneous hierarchical rule learning in adults, and used eye tracking to show that infants used the shape of an object as a higher-order context to organize these visuospatial inputs into simpler color-location rules. Importantly, I found that infants were able to use this hierarchical organization to generalize these color-location rules to novel shape contexts. In a second experiment, I used this same hierarchical task structure to show that this learning mechanism is domain general and helps infants learn and generalize object-label mappings across different speaker contexts.

Prior computational and neuroscience work in adults indicates that hierarchical rule learning is governed by the PFC and its dopaminergic connections with the striatum (Collins & Frank, 2013; Collins et al., 2014). However, it is unclear whether similar frontostriatal circuitry is also implicated in this learning mechanism in infants, or whether it is supported by alternate neural circuitry in infants until the PFC is sufficiently mature to contribute to this form of learning later in life. Therefore, the aim of Chapter 3 was to examine the neural underpinnings of hierarchical rule learning in infancy. To examine this, I used a combination of functional near-infrared spectroscopy (fNIRS) recording over the PFC and behavioral eye blink rates, a purported measure of dopamine activity, to understand whether similar neural mechanisms support hierarchical rule learning in infancy.

3.1 Role of PFC in Learning and Generalizing Hierarchical Rules in 8-Month-Old Infants

3.1.1 Abstract

Recent research indicates that adults and infants spontaneously create and generalize hierarchical rule sets during incidental learning. Computational models and empirical data suggest that, in adults, this process is supported by circuits linking prefrontal cortex (PFC) with striatum and their modulation by dopamine, but the neural circuits supporting this form of learning in infants are largely unknown. We used near-infrared spectroscopy (NIRS) to record PFC activity in 8-month-old human infants during a simple audiovisual hierarchical-rule-learning task. Behavioral results confirmed that infants adopted hierarchical rule sets to learn and generalize spoken object-label mappings across different speaker contexts. Notably, infants had increased activity over right dorsal lateral PFC when rule sets switched from one trial to the next – a neural marker related to updating rule sets into working memory in the adult literature. Infants' eye blink rate, a possible physiological correlate of striatal dopamine activity, also increased when rule sets switched from one trial to the next. Moreover, the increase in right dlPFC activity in conjunction with eye blink rate also predicted infants' generalization ability, providing exploratory evidence for frontostriatal involvement during learning. These findings provide evidence that PFC is involved in rudimentary hierarchical rule learning in 8-month-old infants, an ability that was previously thought to emerge later in life in concert with PFC maturation.

3.1.2 Introduction

A defining feature of flexible human cognition is the ability to derive hierarchical rules from experience. Hierarchical rules group together sets of lower-order rules that can be cued by higher-order contexts (Collins et al., 2014; Collins and Frank, 2013; Collins and Frank, 2016; Donoso et al., 2014; Frank and Badre, 2012). For instance, individuals in bilingual environments might use hierarchical rules to help learn and separate multiple languages specifying object-label mappings. In this framework, individuals may use a higher-order context, such as the identity of a speaker, to determine an appropriate rule set to use (e.g., language grouping together object-label mappings; Figure 1).

Prior behavioral work suggests that adults spontaneously and incidentally create hierarchical rules when learning simple stimulus-action rules through reinforcement, which supports learning in new contexts (Collins and Frank, 2013; Collins et al., 2014; 2016). Computational models suggest that the formation of hierarchical rules is supported by dopamine-innervated circuits between PFC and striatum (Collins and Frank, 2013; Frank and Badre, 2012). Specifically, these models posit that frontostriatal circuits are hierarchically nested, such that a higher-order context cues a valid rule set to update and maintain in PFC in an anterior circuit, which then constrains the stimulus-response rules that can be selected in a posterior circuit. Dopaminergic reinforcement learning signals allow the network to learn which rule sets are valid for a given context, as well as the valid stimulus-response rules for a given rule set. They also allow a learner to discover when a previously valid rule set can be transferred to a novel context.

Recent research provides behavioral evidence that 8-month-old infants also spontaneously construct hierarchical rules when learning simple visual stimulus-response associations, which supports generalization in novel contexts (Werchan et al., 2015). Infants' eye blink rate, which is an exploratory measure of striatal dopamine activity (e.g., Karson, 1983), also increased when higher-order rules switched during learning. However, PFC has a protracted developmental course – not reaching maturity until adolescence (e.g., Huttenlocher, 1979). As such, one prediction is that behavioral evidence of hierarchical rule learning in infants is not supported by the same frontostriatal circuitry as seen in adults. For example, since the striatum shows relatively high functionality in comparison to PFC in early infancy (e.g., Chugani, 1996), hierarchical rule learning may weight striatal involvement more heavily than frontal involvement until PFC is sufficiently mature to contribute to this form of learning later in life.

Alternatively, hierarchical rule learning might be supported by similar frontostriatal circuitry in infants as in adults. In this view, protracted development might reflect *adaptation* to increasingly complex niches over development, rather than changes in the nature of the computations performed by PFC. In infancy, PFC may support hierarchical rule learning relevant to structuring language or social cognition, for example. When the environment requires learning of more complex rules for action during childhood and adolescence, such as those required to play a musical instrument or drive a car, this circuitry (and hierarchical extensions thereof in anterior PFC) may be adapted to support these increasingly complex demands.

Therefore, in the current study we used near-infrared spectroscopy (NIRS) to examine PFC activity during a simple hierarchical rule learning task that required 8-month-old infants to learn sets of audiovisual toy-word pairings across different higher-order face-voice contexts (Figure 1). We predicted that if PFC is involved in hierarchical rule learning in infants, then increased PFC activity should be observed when infants switch between two higher-order rules (Switch conditions) compared to when the higher-order rule stays the same (Stay conditions). In this context, “rule switching” refers not to response switching but rather to switching between observing two different hierarchical rule structures (see Figure 1). Importantly, we expected these differences would emerge to a greater extent over the course of learning. Specifically, computational models and prior work with adults suggest that switching-related PFC activity should be greater *after* rule structures are learned (Collins and Frank, 2013; Collins et al., 2014).

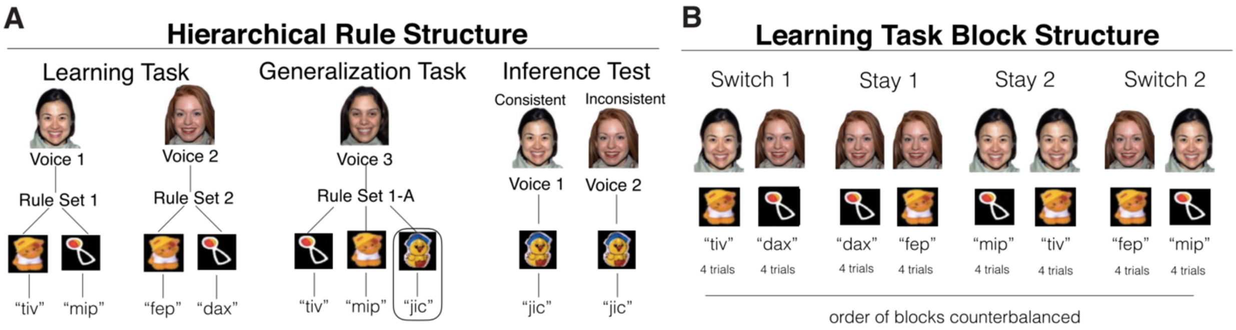

Figure 1. (*A*) Hierarchical rule structure used during the task, which was modeled after Werchan et al. (2015). During the learning task, infants saw face–voice/toy–word mappings that could be grouped into hierarchical rule sets using the faces and voices as higher-order contexts. During the generalization task, infants saw a previously learned rule set now paired with a novel face and voice (RS1-A) and one new toy–word pairing was added to the rule set. During the inference test, infants saw the faces and voices from the learning task paired with the novel toy–word mapping from the generalization task. Infants’ looking time to pairings that were consistent versus inconsistent with the hierarchical structure was measured. (*B*) The learning task was split into two 24 s blocks in which the higher-order rule switched from one trial to the next (Switch 1 and Switch 2) and two 24 s blocks in which the higher-order rule stayed the same from one trial to the next (Stay 1 and Stay 2). The order of blocks was counterbalanced.

3.1.3 Method

3.1.3.1 Participants

The final sample consisted of 37 healthy 8-month-old human infants (20 females, 17 males; mean age = 8.5 months, *SD* = 0.43). All infants were full-term with no current or past history of severe health problems, developmental delays, or birth complications. An additional 3 infants were tested, but their data were discarded due to fussiness or crying (*n* = 2) or NIRS recording interference from hair (*n* = 1). The study was approved by the local Institutional Review Board, and all parents or legal guardians provided written, informed consent prior to participation. All families were compensated for time and travel to our lab.

3.1.3.2 Procedure

3.1.3.2.1 Behavioral procedures

3.1.3.2.1.1 Task overview. Behavioral procedures replicated the methods used by Werchan et al., (2015). During the study, infants participated in a learning task, a generalization task, and a violation-of-expectation inference test. During the learning and generalization task, infants were familiarized with trials that consisted of a face followed by an audiovisual toy-word pairing during a learning task and a generalization task (Figure 1). The faces were discriminable female faces (taken from the NimStim Face Stimulus Set; Tottenham et al., 2009), and the toys were different animated toys. Four artificial monosyllabic words ("jic," "mip," "dax," and "tiv") were used, with a different word assigned to the same object in each of the two rule sets. Each unique word was spoken by one of two female speakers. The mappings between faces, voices, toys, and words were counterbalanced.

The learning task was designed so that infants could learn the pairings in multiple ways: infants could either learn the face-voice/toy-word pairings as four simple individual associations, or infants could use the face-voice mappings as higher-order contexts to create two rule sets (RS1 and RS2) grouping together simpler toy-word associations or rules (i.e., a rudimentary language, which could then be transferred to other contexts cueing that same set). This is similar to learning that mom is associated with an "English" rule set and that dad is associated with a "Spanish" rule set in our earlier example. The initial learning task was split into two blocks where the higher-order stimulus (e.g., face-voice speaker contexts) switched from one trial to the next, requiring infants to update the currently relevant rule set into working memory (Switch 1 and Switch 2 conditions), and two blocks where the higher-order rule stayed the same on each trial (Stay 1 and Stay 2 conditions; Figure 1b). We recorded infants' frontal cortical activity across left and right dorsal lateral PFC (dlPFC) and medial PFC (mPFC) during the learning task using NIRS (Figure 2a and 2b).

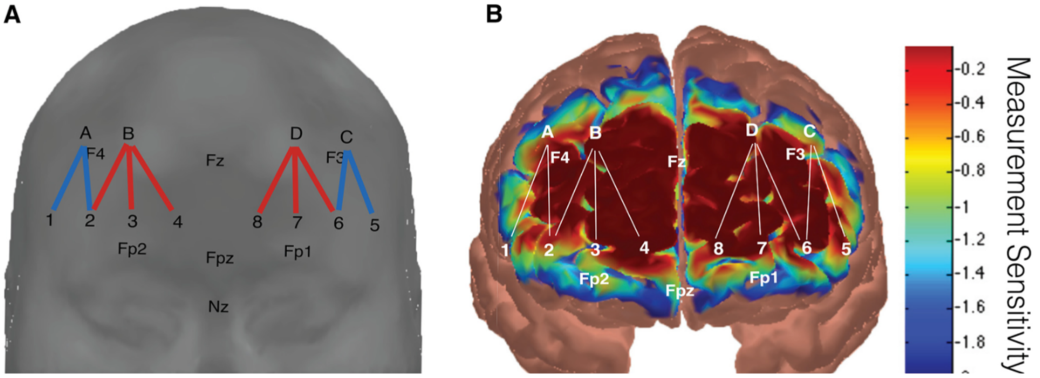

Figure 2. (*A*) Sources (letters) and detectors (numbers) were arranged in a lattice pattern and placed inside of a neoprene headband with the lower edge of the headband positioned in line with the Fp1-Fpz-Fp2 line in the international 10 – 20 system. Red lines represent channels over mPFC and blue lines represent channels over dlPFC. (*B*) Measurement sensitivity to frontal cortex was estimated using AtlasViewer (Aasted et al., 2015), which indicated that the source-detector channels likely targeted a broad area over frontal cortex, including mPFC and dlPFC.

After the learning task we presented infants with a novel face and voice that was associated with a previously learned rule set during a generalization task (RS1-A; similar to observing that grandma also uses object-label mappings consistent with a "Spanish" rule set). Importantly, we also introduced a novel toy-word pairing that could be appended to this rule set (similar to observing grandma speaking a Spanish word for a new object that had not been previously experienced with dad; Figure 1). During a final inference test, we presented infants with trials in which faces and voices from the initial learning task were presented with the novel toy-word pairing from the generalization task. The purpose of this inference test was to examine whether infants had appended the novel toy-word pairing to the existing latent rule set (rather than only to the specific face with which it was presented) and transferred it back to other speakers associated with that rule set (similar to inferring that dad but not mom should use the novel object-label mapping used by grandma; Figure 1). If infants learned the rule structures and used these structures to make inferences about novel face-voice/toy-word pairings, then we predicted that infants should look longer to the inconsistent pairings that violated the learned rule structures. If infants did not adopt a hierarchical rule set structure, then we expected to find no difference in looking time between the consistent and inconsistent trials.

3.1.3.2.1.2 Learning task. During the learning task, infants saw four different pairings of faces and voices with toys and words. Two discriminable female faces and voices, two toys, and four monosyllabic pseudowords were used in these pairings. All mappings between faces, voices, toys, and words were counterbalanced. Each face was associated with a unique voice (e.g., Face 1 was always associated with Voice 1, and Face 2 was always associated with Voice 2). Both of the face-voice mappings were associated with the same two toys (e.g., both Face 1 and Face 2 were paired with a cartoon duck and a rattle); however, each face-voice mapping used different words to label the same toys, akin to learning in a bilingual environment.

Infants received four blocks of 8 trials during learning, for a total of 32 trials (8 trials per pairing). During each trial, infants would first see the face on the left side of the screen. After 500 ms, a toy appeared on the right side of the screen for an additional 1,500 ms, while a recorded female voice said the artificial word associated with the pairing. There was a 1,000-ms interval between trials. The four blocks consisted of two rule switch conditions (Switch 1 and Switch 2)– where the higher-order face-voice context changed each trial – and two rule stay conditions (Stay 1 and Stay 2) – where the higher-order face-voice context stayed the same each trial. Each block was preceded by a 10 second baseline interval to allow blood volume to return to baseline levels during which time a black screen with a white fixation cross was presented. Previous studies indicate that 10 seconds is sufficient for blood volume to return to baseline levels (Wilcox et al., 2008, 2009). The order of blocks was counterbalanced between participants.

3.1.3.2.1.3 Generalization task. Immediately following the learning task, we presented infants with three new pairings of faces, voices, toys, and words. These pairings could again be grouped by the face and voice to form a rule set (RS1-A) identical to one experienced during the learning task (RS1); however, this rule set was now associated with a novel higher-order face-voice context. Additionally, one novel toy-word association was added to the rule set (Figure 1a). Infants again received 8 trials per pairing, as in the learning task, for a total of 24 trials.

3.1.3.2.1.4 Inference test. After the generalization task, infants saw the faces and voices from the original learning task paired with the novel toy-word association from the generalization task. One of these pairings of faces, voices, toys, and words was consistent with the rule set structure formed during the tasks, whereas the other pairing was inconsistent with this rule-set structure (Figure 1a). Infants received two consistent trials and two inconsistent trials during the inference test. The order of test trials was intermixed and counterbalanced across subjects. During each test trial, infants saw the face and toy while a recorded voice said the word associated with the toy once every 3 s. Each trial continued until the infant looked away for more than 2 s, up to a maximum of 60 s. The dependent behavioral measure was the average looking time during the consistent trials compared to the average looking time during the inconsistent trials. Looking time was measured by condition-blind manual coding of the video recordings. Reliability was verified by a second rater coding a subset of the videos (n=10). Inter-rater reliability was high ($r = .97$).

3.1.3.2.2 NIRS recording procedures

Infants' frontal cortical activity was recorded during the learning task using a TechEn CW6 NIRS system (Milford, MA) with wavelengths set at 695 nm and 830 nm. Raw signals were continuously sampled at 50 Hz. An array consisting of 12 optodes (4 sources and 8 detectors, resulting in 10 source-detector channels) with an inter-optode separation of 3 cm was placed over infants' frontal brain regions on each hemisphere. The array was arranged in a lattice pattern (see Figure 2a), which was fixed on sturdy, flexible plastic to ensure that the distance between the sources and detectors remained constant at 3 cm. The optode array was then attached inside of an adjustable neoprene headband to secure the optodes to the scalp. The array was placed over infants' scalps using standardized coordinates corresponding to frontal cortical regions, with the lower edge of the optode array positioned in line with the Fp1-Fpz-Fp2 line in the international 10–20 system (Figure 2a; Jasper 1958). The vertical midline of the optode array was centered across the nasion–inion line.

After recording, the data was pre-processed in HomER 2.0 using the default preprocessing pipeline (see Huppert et al., 2009). First, the raw signals (acquired at 60 Hz) were digitally band-pass filtered at 0.01-0.1 Hz to remove systematic physiological and movement artifacts (see Homae et al., 2010; White et al., 2009). Second, the change in optical density was calculated for each wavelength relative to the 10 second baseline prior to block onset, during which a black screen with a white fixation cross was presented. Third, changes in the concentration of oxygenated and deoxygenated hemoglobin were calculated from the changes in optical density using the modified Beer-Lambert law. Next, motion artifacts were detected by identifying signal fluctuations greater than ±5 μM over a .5 second range in each channel (see Emberson et al., 2015; Lloyd-Fox et al., 2009). This indicated that no motion artifacts occurred during the 4 blocks of interest (Stay 1, Stay 2, Switch 1, Switch 2), so all blocks were retained for subsequent data analysis. Finally, changes in oxygenated hemoglobin (relative to the 10 second baseline) in each of the 10 source-detector channels were exported for subsequent analysis by averaging across every 4 seconds of each 24 second block starting 4 seconds after block onset to account for the hemodynamic response lag (Miezen et al., 2000). This created a total of 5 time intervals for each block during the learning task.

The 10 source-detector channels were divided and averaged into four regions of interest for subsequent data analysis, with the two left-most channels corresponding to left dorsal-lateral PFC (dlPFC), the three center-left channels corresponding to left medial PFC (mPFC), the three center-right channels corresponding to right mPFC, and the two right-most channels corresponding to right dlPFC (see Figure 2a). These regions of interest were verified by estimating measurement sensitivity to these cortical regions (based on the positioning of the optode array in reference to standardized 10-20 coordinates as described above) using AtlasViewer NIRS image reconstruction tools (Figure 2b; see Aasted et al., 2015).

3.1.3.2.3 Eye blink rate measures

In addition to measuring frontal activity, we also measured infants' eye blink rate during learning, which is thought to be a physiological correlate of striatal dopamine activity (e.g., Colzato et al., 2009; Blin et al., 1990; Karson, 1983; Kleven and Koek, 1996; Shukla, 1985; Taylor et al., 1999). Evidence for this association comes from several sources. For example, the use of dopamine agonists and antagonists systematically increases and decreases eye blink rate in non-human primates (Kleven and Koek, 1996; Jutkiewicz and Bergman 2004; Elsworth et al., 1991; Karson, 1983). Eye blink rate is also altered in clinical populations associated with dopaminergic dysfunction. For example, increased eye blink rate is observed in schizophrenic patients (Freed et al., 1980; Mohr et al., 2005), who have increased striatal dopaminergic activity, whereas decreased eye blink rate is observed in Parkinson's disease patients (Blandini et al., 2000; Deuschel, and Goddemeier, 1988; Bodfish et al., 1995) and in recreational cocaine users (Colzato et al., 2008), two populations associated with substantial loss of nigrostriatal dopaminergic cells.

Numerous studies also provide evidence that eye blink rate is reliably altered during cognitive tasks that are associated with dopaminergic functioning. For example, in adults increased eye blink rate is associated with increased cognitive flexibility (e.g., flexibly switching between tasks), during positive compared to negative feedback learning (Slagter et al., 2015), and during working memory tasks (Dreisbach et al., 2005; Müller et al., 2007), whereas phasic decreases in eye blink rate are associated with increases in visual attention and cognitive load (Fukuda et al., 2005; Tada, 1978; Oh et al., 2012). Blink rate is also predictive of the effects of dopamine agonists on reinforcement learning (Cavanagh et al., 2014). While less is currently known about the relation between eye blink rate and dopaminergic functioning during infancy, eye blink rate is found to increase during feeding and presentation of novel stimuli in infants, which is also influenced by dopamine agonists (see Bacher and Smotherman, 2004 for a review). Notably, increased eye blink rate has also been observed when infants switch between higher-order rules during hierarchical rule learning tasks (Werchan et al., 2015).

We used this evidence to generate a specific prediction about the relationship between infants' eye blink rate and PFC activity to consider an exploratory analysis that could implicate involvement of dopamine-innervated frontostriatal circuitry during learning (note that we do not have the ability to directly image subcortical areas with NIRS). Infants' eye blink rate was measured during the learning task by manually reviewing the recorded videos frame by frame to identify eye blinks during the learning task. Eye closures were counted as a blink when both of infants' eyelids closed symmetrically for 100-500 ms. Eye closures due to coughs, sneezes, or yawns were not counted as valid blinks. We predicted that if hierarchical rule learning and generalization is supported by dopamine-innervated circuits between PFC and striatum in infants, then a relationship between infants' eye blink rate and PFC activity should be greatest during the Switch conditions, when infants must update learned rule sets into working memory, relative to the Stay conditions during the learning task. Moreover, we also predicted that infants' eye blink rate and PFC activity during rule switching should together predict infants' ability to make inferences about novel face-voice/toy-word pairings during the final inference test, as measured by infants' looking to the inconsistent relative to the consistent pairings.

3.1.4 Results

3.1.4.1 Behavioral Results

We examined infants' looking time between the pairings that were consistent versus inconsistent with the rule set structure during the final inference test (Figure 3a). We predicted that infants should look longer at the inconsistent relative to the consistent pairings if the hierarchical rule structure was learned. This test indicated that infants looked longer at the inconsistent pairing than at the consistent pairing, $t(36) = 1.975$, $p = .05$, replicating our earlier finding that infants created hierarchical rule sets and used these sets to generalize toy-word mappings across speaker contexts (Figure 3b).

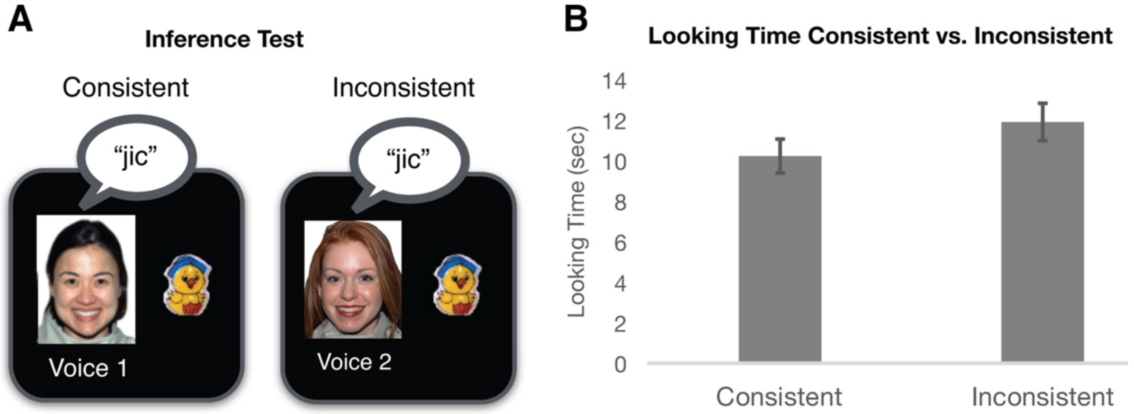

Figure 3. (A) During the inference test, infants' looking times to pairings that were consistent versus inconsistent with the learned rule structures was measured. (B) Infants looked significantly longer at the inconsistent pairing, providing evidence that they constructed hierarchical rule sets and used these sets to make inferences about novel pairings. Error bars reflect standard error of the mean.

3.1.4.2 Eye blink rate

We next examined infants' eye blink rate during the initial learning task. Specifically, we tested the prediction that infants' average eye blink rate per trial should be greatest during the second half of rule switching, after the rule sets are likely learned, based on prior findings showing that infants have increased eye blinks when higher-order rules switch after learning (Werchan et al., 2015). We conducted a repeated measures ANOVA using Block (Block 1 vs. Block 2) and Condition (Switch vs. Stay) as within-subject factors. This test revealed a Block by Condition interaction, $F(1,36) = 4.084$, $p = .051$. Subsequent within- and between-condition comparisons showed that this interaction was driven by increased eye blinks per trial during Switch 2 relative to Switch 1, $t(36) = 3.104$, $p = .004$ (Bonferroni corrected alpha set to .008, see Figure 4).

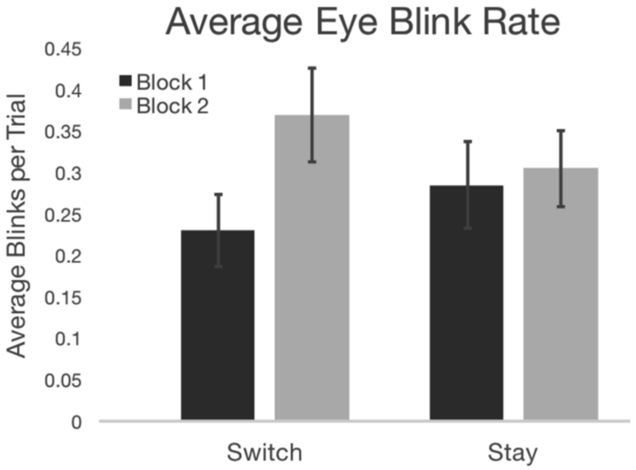

Figure 4. Infants' eye blink rate was significantly greater during the second rule switch block relative to the first rule switch block during learning. Error bars reflect standard error of the mean.

3.1.4.3 NIRS results

We examined differences in infants' cortical activity, as measured by changes in oxygenated hemoglobin (HbO_2), during each of the four conditions of the learning task (Stay 1, Stay 2, Switch 1, and Switch 2), and across the four regions of interest (left dlPFC, left mPFC, right dlPFC, and right mPFC; Figure 2a and 2b), and assessed whether this neural activity varied as a function of whether infants successfully generalized learned rules. Changes in HbO_2 responses were calculated by averaging the HbO_2 response across every 4 seconds of each 24 second block (starting 4 seconds after block onset to account for the hemodynamic response lag; Miezen et al., 2000). This created a total of 5 time intervals for each block. Changes in HbO_2 responses were measured relative to a 10 second baseline period prior to block onset for each condition, where the baseline period consisted of presenting a black screen with a white fixation cross. We conducted an omnibus repeated measures ANOVA with Time-Interval (average HbO_2 response across the five time intervals for each condition), Region (left dlPFC, left mPFC, right dlPFC, and right mPFC), and Condition (Stay 1, Stay 2, Switch 1, and Switch 2) as within-subject factors, and Learning-Score (infants' looking to the inconsistent – consistent test trials) as a continuous variable. This analysis revealed a significant 3-way interaction between Time-Interval, Region, and Condition, $F(36,1260) = 1.686$, $p = .007$, $\eta_p^2 = .046$. To determine which of the four regions of interest exhibited significant differences in activity in this 3-way interaction, we next conducted four separate repeated measures ANOVAs with Time-Interval and Condition as within-subject factors and Learning-Score as a continuous variable. This test (alpha set to a corrected .0125 value) revealed significant interactions between Time-Interval and Condition for right dlPFC only, $F(12,420) = 2.219$, $p = .01$, $\eta_p^2 = .060$. This interaction was not significant for left dlPFC, $F(12,420) = 0.798$, $p = .653$, $\eta_p^2 = .022$, left mPFC, $F(12,420) = 0.601$, $p = 0.842$, $\eta_p^2 = .017$, or for right mPFC, $F(12,420) = 0.575$, $p = 0.862$, $\eta_p^2 = .016$.

We next examined right dlPFC activity separately for all combinations of conditions from the previous interactions (Switch 1 versus Stay 1, Switch 1 versus Switch 2, Stay 1 versus Stay 2, Stay 2 versus Switch 2, Switch 1 versus Stay 2, and Stay 1 versus Switch 2). We compared the change in right dlPFC activity over time *between* conditions rather than *within* single conditions to test the prediction that switching between higher-order rules (in the Switch conditions) would evoke a greater increase in PFC activation than switching between simpler, lower-order rules (in the Stay conditions). We used a Bonferroni corrected alpha value set to .008 (.05/6). These analyses showed a significant 3-way interaction between Interval, Condition, and Learning-Score for right dlPFC in Stay 2 versus Switch 2 only, $F(4,140) = 5.328$, $p = .001$, $\eta_p^2 = .132$ (full results for all comparisons are reported in Table 1), which indicates that infants who had better learning and subsequent transfer of the rule structures also had a greater increase in right dlPFC activity across the Switch 2 block relative to the Stay 2 block (Figure 5 and Figure 6a). To verify this interpretation, we tested the precise relations among infants' Learning-Scores and infants' HbO_2 responses across the significant Conditions from the prior analyses (Switch 2 and Stay 2). To examine this, we performed two regression analyses using infants' right dlPFC activity during either Switch 2 or Stay 2 as predictor variables and infants' Learning-Scores as the dependent variable. Results indicated that right dlPFC activity during Switch 2 was trending towards predicting Learning-Scores, $F(5,31) = 2.081$, $p = .095$, $R^2 = .251$, but right dlPFC activity during Stay 2 was not a significant predictor of Learning-Scores, $F(5,31) = 1.324$, $p = .280$, $R^2 = .176$.

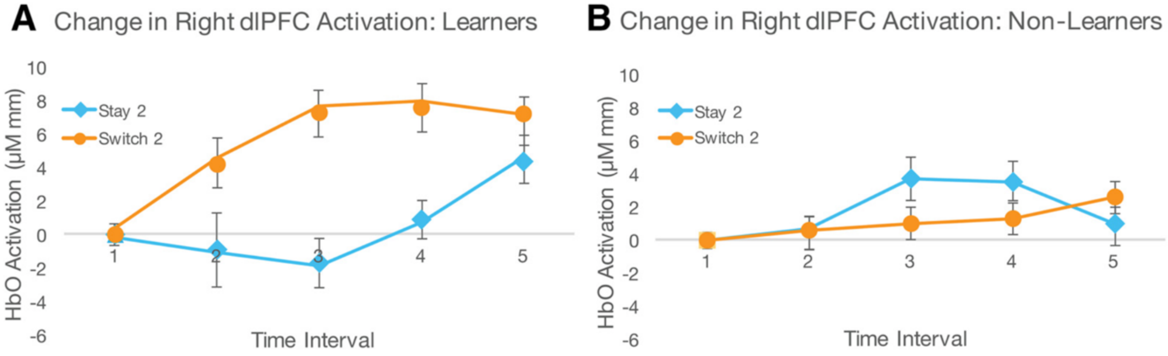

Figure 5. Right dlPFC activity was greater during the second Switch block relative to the second Stay block in infants who demonstrated better learning and generalization of the rule set structure, as evidenced by the significant interaction between Interval, Condition, and Learning Score. To visually illustrate this interaction, the baseline-corrected change in right dlPFC activity is shown in the 10 infants with the best Learning Scores (A) and the 10 infants with the worst Learning Scores (B). Error bars reflect standard error of the mean.

3.1.4.4 Interactions between Eye Blink Rate, PFC activity, and Learning

Next, we examined the interaction between increased PFC activity and infants' eye blink rate to consider exploratory evidence of dopamine-mediated frontostriatal involvement during learning. Specifically, we tested the prediction that a relationship between increased PFC activity and eye blink rate would be greatest when infants must switch between higher-order rules during the second rule switch block, after the rule sets are likely learned. To examine this, we calculated the interaction between the change in right dlPFC activity and eye blink rate by multiplying the change in right dlPFC activity (from the beginning of each block to the end of each block, which is captured in the prior interactions between Learning Score with Time Interval; see Figure 5), with the average eye blink rate for each of the four conditions. We then performed six repeated measures ANOVAs comparing differences between this interaction in all combinations of conditions (Switch 1 versus Stay 1, Switch 1 versus Switch 2, Stay 1 versus Stay 2, Stay 2 versus Switch 2, Stay 1 versus Switch 2, and Switch 1 versus Stay 2) using Learning-Score as a continuous variable. We used a Bonferroni corrected alpha value set to .008 (.05/6). These analyses indicated that there was a significant main effect of Condition in the interaction for Switch 2 relative to Switch 1 only, $F(1,35) = 8.447$, $p = .006$, $\eta_p^2 = .194$, which reflected the fact that right dlPFC activity and eye blink rate was greater in Switch 2 relative to Switch 1. Condition was only marginally significant in the interaction for Switch 1 relative to Stay 1, $F(1,35) = 7.344$, $p = .010$, $\eta_p^2 = .173$, reflecting that the interaction between right dlPFC activity and eye blink rate was greater in Stay 1 relative to Switch 1. There was not a significant effect of Condition for the interaction between right dlPFC activity and eye blink rate for Stay 1 versus Stay 2, $F(1,35) = 0.422$, $p = .520$, $\eta_p^2 = .012$, Stay 2 versus Switch 2, $F(1,35) = 0.728$, $p = .399$, $\eta_p^2 = .020$, Switch 1 versus Stay 2, $F(1,35) = 5.227$, $p = .028$, $\eta_p^2 = .184$, or Stay 1 versus Switch 2, $F(1,35) = 0.001$, $p = .981$, $\eta_p^2 = .069$.

We then conducted a regression analysis to determine whether the interaction between right dlPFC activity and eye blink rate during Switch 2 was related to transfer performance during the inference test. We used right dlPFC activity, eye blink rate, and the interaction between right dlPFC activity and eye blink rate during Switch 2 as predictor variables. This regression analysis indicated that the conjunction between eye blink rate and the change in right dlPFC activity significantly predicted transfer performance, $\beta = .885$, $p = .001$ (Figure 6b). We performed a control regression analysis using infants' learning score rank order as a dependent variable to ensure that the results were not due to any one infant with an unusually high or low learning score. The results indicated that the conjunction between eye blink rate and PFC activation remained marginally significant ($\beta = .512$, $p = .097$), and explained statistically more variance than either eye blink rate or PFC activation alone ($\beta = -.242$, $p = .371$ and $\beta = -.218$, $p = .300$, respectively). Taken together, these findings indicate that infants who had better learning and subsequent

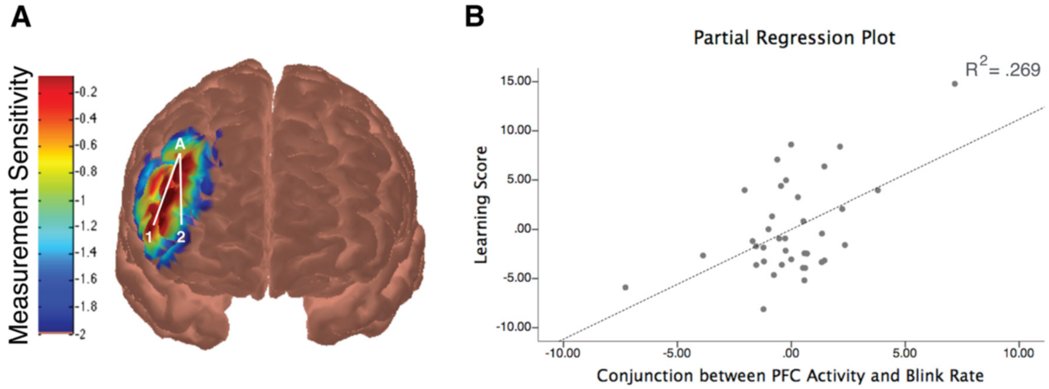

Figure 6. (A) Significant cortical activity was observed in channels over right dlPFC (measurement sensitivity map shown) during Switch 2 relative to Stay 2 in infants who demonstrated subsequent transfer of the rule structures. (B) Partial regression plot illustrating the relation between right dlPFC activity, eye blink rate, and transfer performance. Results from the regression analysis indicated that a conjunction between eye blink rate and the change in right dlPFC activity (from the beginning to the end of each block) predicted transfer performance during the subsequent inference test. In particular, infants with a higher blink rate combined with a greater change in right dlPFC activity showed better transfer performance.

transfer of the hierarchical rule structures also had increased right dlPFC activity combined with increased eye blinks during the second block of rule-switching, providing exploratory evidence for dopamine-innervated frontostriatal involvement during learning.

3.1.5 Discussion

Here we examined the neural circuitry supporting hierarchical rule learning in 8-month-old infants by using NIRS to record infants' PFC activity during a hierarchical rule learning task. Our behavioral results replicated previous findings and showed that infants spontaneously create hierarchical rule structures during incidental learning and use these structures to make generalizations in novel contexts (Werchan et al., 2015). Notably, however, our results showed that increases in PFC activation relate to this learning mechanism in 8-month-old infants, a skill that was previously assumed to emerge later in life in parallel with the protracted anatomical development of PFC. In particular, we provide novel evidence that infants had increased right dlPFC activity during the second half of learning when higher-order rules switched from one trial to the next, requiring infants to update the currently relevant rule representation in working memory, relative to when the higher-order rules stayed the same from one trial to the next. We also measured eye blink rate, which is thought to be a physiological correlate of striatal dopamine activity (e.g., Colzato et al., 2009; Blin et al., 1990; Karson, 1983; Kleven and Koek, 1996; Shukla, 1985; Taylor et al., 1999), and found that infants had increased eye blink rate during the second half of rule switching, replicating prior findings (Werchan et al., 2015). Critically, we also found novel evidence that infants' transfer of hierarchical rule structures was related to the conjunction of right dlPFC activity and eye blink rate, potentially indicative of dopamine-mediated frontostriatal circuitry in hierarchical rule learning in infancy.

Our finding that infants had increased eye blink rate during the second half of rule-switching suggests the potential involvement of dopamine-related processes during learning, based on prior findings indicating a relationship between increased eye blink rate and *tonic* dopamine firing (e.g., Colzato et al., 2009; Blin et al., 1990; Karson, 1983; Kleven and Koek, 1996; Shukla, 1985; Taylor et al., 1999). However, computational models suggest that *phasic* dopamine firing relates to reward prediction errors during learning, which drives learning of hierarchical rule structures (Collins and Frank, 2013; Frank and Badre, 2012). This raises the question of what increased eye blink rate reflects in relation to hierarchical rule learning in our current findings. One possibility is that tonic dopamine may reflect accumulation of phasic dopamine across learning (due to spillover, e.g., Niv et al., 2007); thus, as learning occurs, tonic dopamine levels may increase, leading to increased eye blink rate during the second switch block relative to the first. It is also possible that *either* tonic or phasic changes in dopamine translate to changes in eye blink rate. However, given that eye blink rate is an exploratory measure, particularly in infants and in relation to phasic changes in dopamine, more research is needed to corroborate these findings and come to a more conclusive understanding of the exact nature of eye blink rate in relation to learning in infants.

Our finding that infants had increased right dlPFC activity during rule switch trials in the second half of learning also aligns with prior electrophysiology work showing that ERPs over right dlPFC are related to switching between hierarchical rules during incidental hierarchical rule learning in adults (Collins et al., 2014). It also aligns with predictions that arise from neural network models of spontaneous hierarchical rule learning, which indicates that the network must first learn relevant mappings between contexts and rule structures before PFC reliably switches between the rule structures (Collins & Frank, 2013). Prior to this learning, the network model could update PFC representations even when a switch had not occurred, or it could fail to update when a switch did occur, and thus the model predicts no reliable difference in PFC activity between Switch and Stay blocks during early learning. Moreover, our finding that the conjunction between infants' dlPFC activity and eye blink rate during rule switch trials predicted behavioral performance during the inference test is consistent with findings in adults indicating that ERP responses over dlPFC also predict generalization of learned rule structures in novel contexts (Collins et al., 2014). These findings also align with prior work showing a more general role for dlPFC in the application and maintenance of abstract rules after initial learning (e.g., White and Wise, 1999; Goel and Dolan, 2000; Seger et al., 2000; Strange et al., 2001; Bunge, 2004).

An alternative possibility is that the increased dlPFC activity that was observed during rule switching might reflect cortical responses to changes in perceptual features, such as changes in the perceptual features of the faces, rather than cortical responses to changes in the higher-order rule from one trial to the next. However, if dlPFC activity reflected changes in perceptual features rather than changes in the higher-order rule, then dlPFC activity should be similar during the first and second switch blocks, which were perceptually equivalent to one another. Since significant dlPFC activity was only observed during the second rule switch block, after the rule structures were likely learned, it is unlikely that our results reflect cortical responses to changes in perceptual features. Further, the fact that right dlPFC activity combined with eye blink rate predicted transfer performance during the final inference test provides further support that our results reflect cortical responses to switching between higher-order rules.

Another important note is that in the PFC activation data, differences were found between Stay 2 and Switch 2, whereas differences were found between Switch 1 and Switch 2 for the eye blink rate data. These apparent discrepancies may reflect differences in the underlying processes captured by eye blink rate and PFC activity. For instance, if eye blink rate reflects dopamine-related processes, then differences in blink rate may gradually emerge as the rule sets are acquired through dopaminergic-based reward prediction learning, leading to differences between Switch 1 and Switch 2 over the course of learning. In contrast, if PFC activation reflects updating or switching between acquired rule sets, then differences in PFC activity during rule switch relative to rule stay blocks would likely only emerge after the rule sets are acquired, leading to differences between Stay 2 and Switch 2. These findings are supported by computational models, which suggest that (1) reward prediction errors are greatest early in learning (2) relevant mappings between contexts and rule structures must be learned before PFC can reliably switch between the rule structures (Collins & Frank, 2013).

Our results are somewhat surprising, however, given that PFC is not thought to reach maturity until adolescence (e.g., Huttenlocher, 1979), and rule learning and working memory processes are thought to have a similarly protracted course of development (e.g., Bunge and Zelazo, 2006; Crone et al., 2006). How then can we reconcile our current findings with what is known about the maturational state of PFC in infancy? One possibility is that PFC is functional and performs similar computations across the lifespan, but these computations are adapted for learning demands and inputs that are appropriate to infants in their unique ecological niche (Rovee-Collier and Cuevas, 2009). For instance, this general mechanism may help infants raised in bilingual environments separate languages without having to learn every word in each context, as supported by our current behavioral findings. Similarly, it may also help infants organize visual information in such a way that affords generalization in novel contexts (Werchan et al., 2015). As the learning demands present in an individual's ecological niche change across development, this PFC-supported learning mechanism (and hierarchical extensions thereof in anterior PFC) might then be adapted to support more complex learning demands in the environment, such as cognitive control of complex thought and action in more mature learners. Thus, previous studies that have observed protracted developmental trajectories of complex rule learning may broadly reflect mismatches between the niche-specific functions that these mechanisms are adapted to support and the functions that a particular task tests, rather than reflecting immaturity of PFC per se.

A further possibility is that adults might use a more advanced hierarchical nesting of frontostriatal circuits to perform such rule learning and transfer. While the hierarchical structure in the current task is analogous to that in the adult studies (Collins and Frank, 2013; Collins et al., 2014; Collins and Frank, 2016), the latter studies clearly impose more demands on using hierarchical rules to contextualize selection of arbitrary actions, where both the rule structures and the actions must be learned simultaneously, whereas the infant studies do not require such action control. Thus, infants might engage similar circuits and computations but which do not require as advanced processing across multiple stages along the rostrocaudal axis of PFC.

In summary, our results provide evidence that dlPFC is implicated in hierarchical rule learning in 8-month-old infants, supporting the formation of abstract rules that afford flexible behavior and generalization in novel environments. Our results showing a relationship between dlPFC activity and eye blink rate in predicting transfer performance also provides exploratory evidence for involvement of dopamine-innervated frontostriatal circuitry in hierarchical rule learning in infancy. In addition to increasing our general understanding of the fundamental learning mechanisms available to infants, these results also provide novel insights into the functional role of PFC and frontostriatal circuitry in learning and behavior during early life. In contrast to prior notions that PFC and complex rule learning mechanisms are slow developing, these results suggest an important role for PFC and frontostriatal circuitry in learning during early life. Future work is needed to examine how this basic learning mechanism and its supporting neural circuitry might then be adapted to support increasingly complex learning demands across development.

Chapter 4

4. A Rule Learning Account of Infants' Errors on Canonical Measures of Executive Functions

In the prior two chapters, I provided behavioral and neuroimaging evidence that the PFC supports learning and generalization of inputs that are relevant to infants in their environment. These studies support the hypothesis that PFC is adapted for learning demands and behaviors that are relevant to infants in their unique ecological niche. Yet, a wide body of research shows that young infants perform poorly on canonical measures of executive functions, such as commonly seen in infants under one year of age during the A-Not-B task (Piaget, 1954; Diamond, 2002). Although the A-Not-B task is typically used to test perseverative behavior and inhibitory control in infants, it is possible that infants might experience it as a series of events that they may attempt to organize into generalizable rule structures. In this view, infants' perseverative errors might reflect the byproduct of these domain-general rule learning mechanisms that act to organize environmental inputs into generalizable rule structures. Therefore, the aim of the following chapter was to test this hypothesis and examine the flexibility of an ecological approach to PFC development by testing whether a learning account of infant PFC function can account for the seemingly perseverative A-Not-B error (Piaget, 1994).

4.1 Adults Just Don't Understand: The A-Not-B Error Reflects Online Reinforcement Learning About Task Dynamics

4.1.1 Abstract

Recent findings highlight a role for the infant prefrontal cortex in hierarchical rule learning, rather than action inhibition. These findings raise questions about previous explanations for infants' A-not-B errors. Although the A-Not-B task is typically used to test perseverative behavior, infants might experience it as a series of events that can be organized into hierarchical rules. Nine-month-olds (N = 119) participated in three experiments designed to test this hypothesis. Results indicated that infants' reaching accuracy was better on trials that were *consistent* with a hierarchical structure, even if these trials required infants to inhibit a prepotent motor response. These data suggest that the A-Not-B error reflects an adaptive strategy that helps infants learn generalizable representations of their environment. *(118 words)*

4.1.2 Introduction

The prefrontal cortex (PFC) supports adaptive behavior through its role in maintaining and updating information into working memory to align with relevant goals for learning and action (Badre, 2008; E. K. Miller & Cohen, 2001; O'Reilly, 2006; Rougier et al., 2005). A recent approach to PFC development argues that the PFC performs these same computations across the lifespan, e.g. working memory updating, beginning in infancy. Rather than having a protracted developmental course, Werchan & Amso (2017) argue that PFC functional development involves continual adaptation of these core computations to meet the changing demands on cognition and behavior in infants' ever changing ecological niche, including changes in visual input, language, body proportions, and motor skills for example. In this view, the PFC might be better adapted to support learning during infancy, the time in life when the learning curving is steepest, rather than inhibitory control over motor or oculomotor action (Werchan, Collins, Frank, & Amso, 2015, 2016). However, PFC function in infancy is almost always tested and interpreted from the latter inhibitory control perspective (Diamond & Goldman-Rakic, 1989; M. H. Johnson et al., 1994). Here we ask whether a learning account of infant PFC function can account for the seemingly perseverative A-Not-B error (Piaget, 1994).

In the A-Not-B task, infants watch as an experimenter hides a desirable toy in one of two locations, and are then allowed to search for the toy after a brief delay. The toy is hidden in the same location (A) for multiple trials, after which it is switched and hidden in the alternate location (B). The A-Not-B error occurs when infants reach to the previous location when the hiding location is switched (Piaget, 1954). These perseverative errors are typically observed in infants 8-12 months of age (Diamond, 2002).

Infants' failures are attributed to a range of explanations. According to Piaget, the A-not-B error reflected a failure of object permanence. This explanation became increasingly unlikely as researchers showed that even 3.5- to 4-month-old infants were capable of object permanence (Baillargeon, 1987). Key studies in human and non-human primate infants showed that PFC development plays a critical role in success and failure on the A-not-B task (Diamond & Godman-Rakic, 1989). Munakata et al., (1997, 2001) argued for a graded representations account where action is dependent on strength of memory representations formed online during learning. This approach stresses a knowledge-action dissociation – the idea that infants know where the toy is, but cannot execute a response consistent with that knowledge (Munakata, 2001). Diamond & Goldman-Rakic, and later Munakata, thus agree that understanding reaching perseveration is important for understanding the A-not-B error. However, they disagree on the source, with Diamond & Goldman-Rakic (1989) arguing for poor inhibitory control over motor reaching and Munakata (1997, 2001) arguing for the idea that weak working memory representations can support some actions (looking) but not others (reaching).

Here we build on the ideas offered above that the A-not-B task involves PFC (Diamond & Goldman-Rakic, 1989) and that it involves online reinforcement learning about task dynamics (Munakata, 2001; Munakata, McClelland, Johnson, & Siegler, 1997). However, our approach is inspired by recent evidence demonstrating that the PFC supports learning during infancy through its role in organizing working memory representations and structuring inputs into generalizable rule sets (Werchan et al., 2015, 2016; Werchan & Amso, 2017). In this work, we found that 8-month-old infants can use the shape of a simple object as a higher-order context to organize working memory representations for simpler color-location rules (e.g., if the context is "square", then "red" predicts a cartoon in "location 1" and "blue" predicts a cartoon in "location 2"). We also found that infants similarly use face-voice contexts to organize different spoken labels for the same sets of objects. The benefit of structuring inputs in this way is that it then allows infants to make predictions about future events and generalize learning in new contexts. Additionally, it helps prevent catastrophic interference when new learning interferes with prior information learned in a different context. Thus, based on this adaptive learning account of PFC, we hypothesized that the A-not-B error may be a byproduct of infants' attempt to learn a higher-order rule structure that supports learning and generalization in new contexts.

This proposal appeals to research in hierarchical action and reinforcement learning, where the goal is to learn stimulus-action-outcome (S-A-O) contingencies that depend on a higher-order context (space, object, person). Empirical and computational evidence in adults show that this learning depends on nested dopamine-innervated frontostriatal circuitry (Badre & Frank, 2012; Collins, Cavanagh, & Frank, 2014; Collins & Frank, 2013; Michael J Frank & Badre, 2012), and fNIRS studies show that this learning mechanism engages similar frontostriatal circuitry in infants as in adults (Werchan et al., 2016).

Infants experience a particular sequence of events in the A-Not-B task. They interact with an experimenter (higher-order context), a fun toy (stimulus) is shown and hidden by the experimenter in well A, and they are then allowed to search (action) for the rewarding toy (outcome). This procedure is repeated for several trials. By the learning account of PFC offered here, rather than acting on a trial-by-trial basis with the singular goal of finding the hidden toy, infants may instead use the repeated A-event exposures to learn a hierarchical rule structure where a higher-order context, in this case the experimenter, organizes separate working memory representations for the toy-well-action rules. From this vantage point, what may seem like an error may be an attempt to learn how a higher-order context (the experimenter) organizes lower-order toy-well-action rules. Put simply, infants may be using contextual information offered during these event sequences to try to learn rules for where an experimenter is most likely to put a toy. Thus, although we may be interested in control over action in the A-not-B task, infants may experience it as a series of event sequences that can be organized into hierarchical rules to support learning and procurement of the toy reward.

This idea is broadly consistent with a key role for the PFC (Diamond & Rakic, 1989) and with a graded memory representations account relevant to learning (Munakata et al., 1997). It differs in that we predict that the A-not-B error reflects an otherwise adaptive attempt to learn hierarchical and generalizable event sequences in the environment (Figure 1), and a possibly meaningful strategy of negative transfer when the higher-order context (experimenter) remains the same when the hiding location is switched.

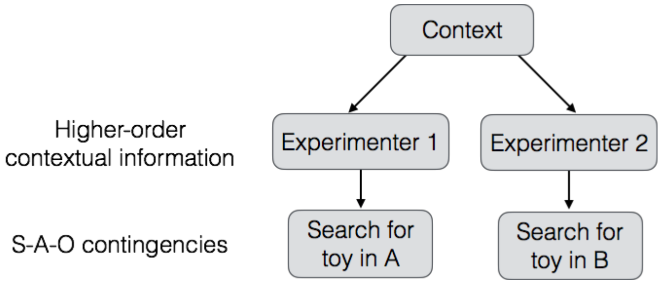

Figure 1. Higher-order contextual information specifies the set of lower-order rules (S-A-O contingencies). Infants may use experimenters as higher-order contexts to learn hierarchical and generalizable rule sets from event sequences, which are then used to guide search behavior.

To examine this hypothesis, we modified the A-Not-B task so that infants could use two experimenters as higher-order contexts to organize separate working memory representations for the lower-order toy-well-action rules. We tested 9-month-olds, an age-group who consistently show the A-not-B error (Diamond, 2002). In Experiments 1 and 2, we examined whether perseverative errors are reduced when a change in the hiding well is associated with a change in the experimenter higher-order context (Figure 2). We hypothesized that a change in the experimenter higher-order context would cue updating of a novel S-A-O rule into working memory. In Experiment 3 we tested the critical prediction that A trials, which require repeating a motor response, that are *inconsistent* with a hierarchical rule structure (where a higher-order context should have hidden the toy) should result in *worse* performance than B trials, which require switching a motor response, that are *consistent* with a hierarchical structure (Figure 7). In other words, if a switch in the higher-order context cues updating of a novel S-A-O rule into working memory, then this should facilitate accurate reaches on a trial that otherwise requires inhibiting a prepotent motor response. *(1249 words)*

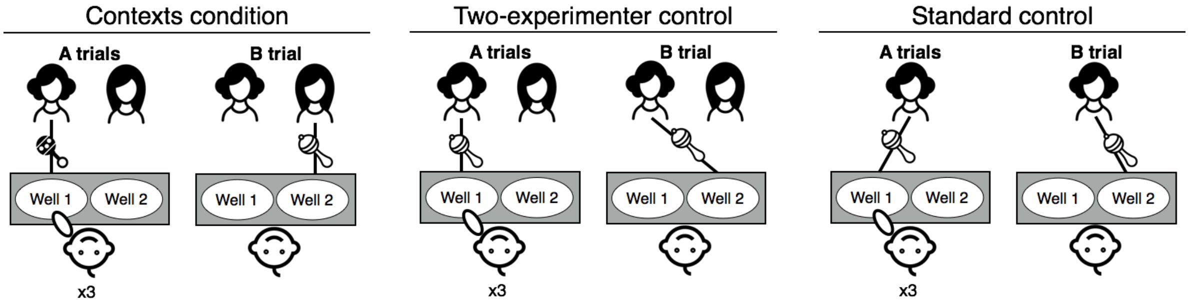

Figure 2. Examples of a testing block in each of the three conditions in Experiment 1. In the contexts condition, a change in the hiding well was associated with a change in the experimenter, whereas the experimenter remained the same across trials in the control conditions.

4.1.3 General Method

4.1.3.1 Participants

The final sample across all three experiments consisted of 119 nine-month-old infants (*M* = 9.24 months, *SD* = .84 months, 44 females, 40 males, 64 white non-Hispanic, 5 black, 7 Hispanic, 5 Asian, and 3 Mixed Race/Other). Infants were recruited from the state department of health birth records and through community advertisements. Infants were prescreened for premature birth (< 36 weeks), low birth weight (< 5 lb), or a history of serious health problems. The Brown University Institutional Review Board approved the study, and parental consent was obtained prior to testing.

4.1.3.2 Task Apparatus

Infants were seated on their parent's lap across a table from the experimenter(s). The A-Not-B apparatus was designed following Diamond & Goldman-Rakic (1989) and consisted of a small blue-felt-covered table that measured 60 cm (L) X 40 cm (W). Two wells were embedded within this table, which were 9 cm in diameter, 11.5 cm deep, and 30 cm apart from center to center. Red felt cloths were used to cover the wells, which measured 22 cm (L) X 15 cm (W). The height of the apparatus was such that the infant could see inside the hiding wells while they were seated on their parent's lap. The toys were attractive, brightly-colored squeaky toys.

4.1.3.3 Procedure

Each experimental session consisted of a familiarization period followed by two blocks of test trials, and was modeled after Diamond & Goldman-Rakic (1989) and Munakata (1998). The familiarization period was used to allow infants to become comfortable with the toys and experimenters. The two test blocks consisted of multiple repeat A trials followed by one switch B trial. Infants' reaching accuracy was the dependent measure. Parents were instructed to not encourage, guide, or correct the infant's actions during the study.

The familiarization period began with the experimenter introducing the toy to the infant, and allowing them to play with it to gain familiarity with the toys and experimenters. The experimenter then partially hid the toy and encouraged the infant to search for the toy. Next, the experimenter placed the toy in one of the wells and covered it with the red felt cloth, and allowed the infant to search for the toy immediately. The infant was encouraged to remove the cover and search for the toy if they did not do so independently. They were then praised and allowed to play with the rewarding toy. This practice trial was repeated in the other well.

After familiarization, infants received the test trials. Each trial began with the experimenter waving the toy centered over the task apparatus to capture the infant's attention. The experimenter then slowly placed the toy in the first well. If the infant looked away while the toy was being hidden, the experimenter repeated the hiding procedure. The experimenter then covered the well with the red felt cloth, and audibly counted to three to draw the infant's visual attention away from the well. After the three second delay, the experimenter pushed the apparatus within reaching distance for the infant, and allowed the infant to search for the toy. If the infant searched for the toy in the correct well, they were praised and rewarded by being allowed to play with the toy for a few seconds. If the infant did not search correctly, the experimenter showed the infant where the toy was hidden, but the infant was not praised or allowed to play with the toy. The experimenter repeated hiding the toy in the same well until the infant searched correctly three consecutive times (A trials). Next, a B trial occurred by hiding the toy in the other well using the same hiding procedure. This block of trials was then repeated by hiding the toy repeatedly in the second well until the infant searched correctly three consecutive times (A trials), and then switching back to the first hiding well (B trial). Thus, there was a total of two testing blocks. The left-right assignment of the first and second hiding wells was counterbalanced across subjects. Also note that since the toy's location is reversed more than one time in the present study, we use the term "A trial" to refer to trials when the toy is repeatedly hidden in the same location, and the term "B trial" to refer to trials when the hiding location switches to the other well.

4.1.3.4 Coding

All testing sessions were videotaped for subsequent coding by a trained observer. A reach was defined as an action that resulted in contact and removal of the red felt cloth. The first location that infants reached to on each trial was scored as the response. A subset of the videos (25%; *n* = 30) were rated by a second independent observer. Reliability of the ratings was 100%.

4.1.4 Experiment 1

4.1.4.1 Method

4.1.4.1.1 Participants

The final sample consisted of 63 nine-month-old infants (*M* = 9.24 months, *SD* = .84 months, 44 females, 40 males, 64 white non-Hispanic, 5 black, 7 Hispanic, 5 Asian, and 3 Mixed Race/Other).

Sample size was determined based on an a priori power analysis with a large effect size (*f* = .4) estimated from prior work (Werchan et al., 2015) at 80% power, which indicated that approximately 22 infants per condition would provide sufficient statistical power. Infants (*N* = 63) were randomly assigned to one of three conditions: a Standard Control (n = 21), a Two-Experimenter Control (n = 21), or a Contexts condition (n = 21). An additional 23 infants were tested, but excluded from the sample due to fussiness resulting in a failure to participate in the task (n = 19), experimenter error (n = 2) or parental interference (n = 2). Approximately equal numbers of infants were excluded from each condition (contexts condition: n = 9, two-experimenter control: n = 9, standard control: n = 5).

4.1.4.1.2 Procedure

We tested three between-subjects conditions – a contexts condition, a standard control, and a two-experimenter control (see Figure 2). The procedure was as described in the General Method section with the following modifications. In the contexts condition, two experimenters hid the toys, and a change in the toy's hiding location was associated with a change in the experimenter hiding the toy. Thus, experimenter 1 hid their toy in the first well until the infant searched correctly three consecutive times (A trials). Then, experimenter 2 hid their toy in the second well (B trial). They continued hiding it in the second well until the infant searched correctly three consecutive times (A trials), after which experimenter 1 again hid their toy in the first well (B trial). Note that each experimenter hid a unique toy in the contexts condition, which was counterbalanced across subjects. In the two-experimenter control, two experimenters were seated side-by-side. Experimenter 1 hid the toy in the first well until the infant searched correctly three consecutive times (A trials), and they then hid it in the second well (B trial). Experimenter 2 then hid their toy in the second well until the infant searched correctly three consecutive times (A trials), and they then hid it in the first well (B trial). Each experimenter hid a unique toy, as in the contexts condition. This control condition allowed us to test whether using two experimenters during the task impacts performance for reasons unrelated to associating each hiding well with a unique experimenter. In the standard control, only one experimenter hid a single toy across all A trials and B trials. The two experimenters in all conditions were young Caucasian females, and thus only differed in identity and not in race, age, or gender. Additionally, the same two experimenters hid the toys across all conditions, and the left-right seating assignments of the experimenters was counterbalanced across subjects.

4.1.4.2 Experiment 1 Results

We conducted an omnibus Mixed-Effects ANOVA with condition as a between-subjects variable, trial type (A trial, B trial) as a within-subjects variable, and reaching accuracy as the dependent measure. This resulted in main effects of trial type, $F(1,60) = 116.660$, $p < .001$, and condition, $F(2,60) = 3.105$, $p = .052$, and an interaction between trial type and condition, $F(2,60) = 6.518$, $p = .003$ (Figure 3). We followed up on this significant interaction with two one-way ANOVAs comparing reaching accuracy by condition on the A trials and B trials separately. There was no significant difference between conditions on A-trial accuracy, $F(2,60) = .474$, $p = .625$. However, this same analysis revealed a significant difference between conditions for B-trial accuracy, $F(2,60) = 5.654$, $p = .006$. Planned two-tailed t tests were used to analyze the source of these differences in B-trial accuracy. These tests revealed that infants performed significantly better on the contexts condition than both the standard control, $t(40) = 3.006$, $p = .005$, $d = .93$, and the two-experimenter control, $t(40) = 2.919$, $p = .006$, $d = .90$, and that there were no differences between the two control conditions, $t(40) = .458$, $p = .650$, $d = .15$.

We next tested whether B-trial reaching accuracy was significantly better than chance (50%) in each of the conditions using three one-sample two-tailed t tests. These tests indicated that B-trial reaching accuracy was not different from chance on the contexts condition, $t(20) = 1.142$, $p = .267$, $M = 57.1\%$, $SD = 28.6\%$, and accuracy was significantly *worse* than chance, indicating perseveration, in the standard control, $t(20) = -2.911$, $p =$

.009, *M* = 26.2%, *SD* = 37.5%, and in the two-experimenter control, *t*(20) = -2.961, *p* = .008, *M* = 31.0%, *SD* = 29.5%. Thus, these findings indicate that there is a reduction in perseverative behavior in the contexts condition relative to the two control conditions.

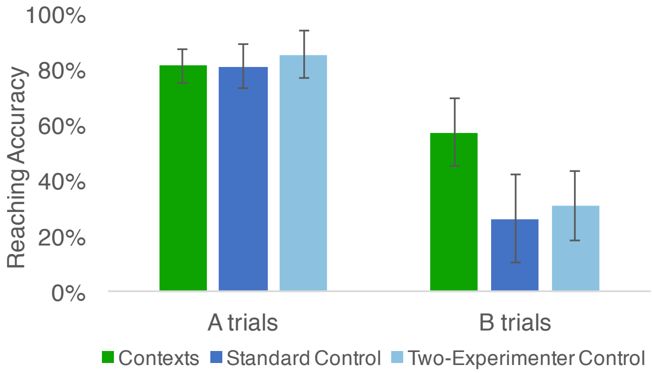

Figure 3. Infants performed significantly better on the B trials in the contexts condition in Experiment 1. There were no differences in A trial accuracy. Error bars reflect 95% confidence intervals.

4.1.4.3 Experiment 1 Discussion

Our findings provide initial support for our hypothesis that infants' perseverative errors might reflect an adaptive attempt to learn how a higher-order context (the experimenter) organizes toy-well-action rules. However, there are two alternative perception-level explanations of our results. First, using two different toys, which may help infants differentiate B trials from A trials, could have driven better performance on the contexts condition (Figure 2). A second possibility is that infants might have used the experimenter's left-right seating location to determine the correct well, as each experimenter was seated on the same side as the well that they hid the toy in (Figure 2). Thus, Experiment 2 aimed to replicate Experiment 1 while controlling for these alternative explanations. *(117 words)*

4.1.5 Experiment 2

4.1.5.1 Method

4.5.1.1 Participants

The sample in the contexts-same-toy condition consisted of 21 nine-month-old infants (*M* = 9.25 months, *SD* = .91 months, 10 females, 11 males, 17 white non-Hispanic, 1 Hispanic, 2 Asian, and 1 Mixed Race/Other). An additional 6 infants were tested, but excluded from the sample due to parental interference (n = 1) or fussiness (n = 5).

The sample in the contexts-location-change condition consisted of 21 nine-month-old infants (*M* = 9.21 months, *SD* = .83 months, 11 females, 10 males, 15 white non-Hispanic, 1 black, 2 Hispanic, 1 Asian, and 2 Mixed Race/Other). An additional 12 infants were tested, but excluded from the sample due to failure to complete the task due to fussiness.

4.1.5.1.2 Procedure

The procedure for Experiment 2 was identical to the contexts condition in Experiment 1 with minor modifications relevant to each condition. For the contexts-same-toy condition, rather than each of the two experimenters hiding a different toy, the two experimenters hid the same toy on each trial (Figure 4). Thus, experimenter 1 would hide the toy in the first well until the infant searched correctly three consecutive times (A trials). Then experimenter 2 would hide the same toy in the other well (B trial), and would continue hiding it until the infant searched correctly three times (A trials). Finally, experimenter 1 would hide the same toy in the first well (B trial).

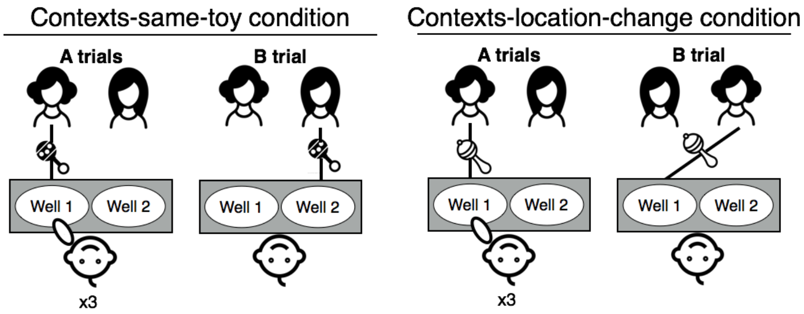

Figure 4. Examples of testing blocks in Experiment 2. In the contexts-same-toy condition, both experimenters hid the same toy. In the contexts-location-change condition, B trials were associated with a change in the location of the experimenter.

The modification for the contexts-location-change condition was such that on B trials, rather than the hiding location of the toy changing to the other well, the spatial location of the experimenter changed (i.e., the two experimenters swapped left-right seating locations; Figure 4). The experimenter and the hiding well remained the same, and the same toy was hidden across all trials. Thus, experimenter 1 would hide the toy in the first well until the infant searched correctly three consecutive times (A trials). Then, the experimenters swapped left-right seating locations, and experimenter 1 again hid the toy in the same well (B trial). Experimenter 2 then repeated this procedure in the second well.

4.1.5.2 Experiment 2 Results

4.1.5.2.1 Contexts-same-toy condition

We compared infants' reaching accuracy on the A trials to the B trials. A paired-samples two-tailed *t* test indicated that there was no significant difference in reaching accuracy between the A trials and the B trials, $t(20) = 1.523$, $p = .143$ (Figure 5). We next tested whether B-trial reaching accuracy was significantly better than chance (50%) using a one-sample two-tailed *t* test. This analysis indicated that accuracy was significantly better than chance, *t*(20) = 2.335, *p* = .030, *M* = 64.3%, *SD* = 28.0%. Note that B-trial reaching accuracy above chance is particularly notable given that B-trial accuracy is significantly below chance in standard versions of the task. A direct comparison of A trial and B trial performance on this condition relative to the 2 Experimenter control in Experiment 1 shows only that there is better performance on the B trial in the contexts-same-toy condition, *t*(40) = 3.755, *p* = .001, *d* = 1.15.

4.1.5.2.2 Contexts-location-change condition

We compared infants' reaching accuracy on the A trials to the B trials. A paired-samples *t* test indicated that there was no difference in reaching accuracy between the A trials and the B trials, *t*(20) = .994, *p* = .332 (Figure 6). We next tested whether B-trial reaching accuracy was significantly better than chance (50%) using a one-sample *t* test. This test indicated that accuracy was significantly better than chance, *t*(20) = 2.609, *p* = .017, *M* = 69.0%, *SD* = 33.5%. Again, a direct comparison of A trial and B trial performance on this condition relative to the 2 Experimenter standard control in Experiment 1 shows only that there is better B-trial performance in the contexts-location-change condition, *t*(40) = 3.915, *p* < .001, *d* = 1.20.

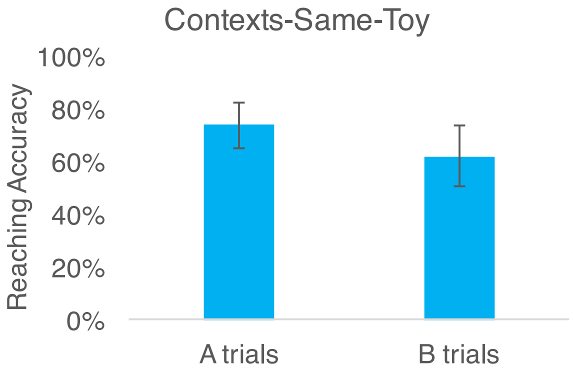

Figure 5. There was no difference in reaching accuracy on A trials and B trials. Error bars reflect 95% confidence intervals.

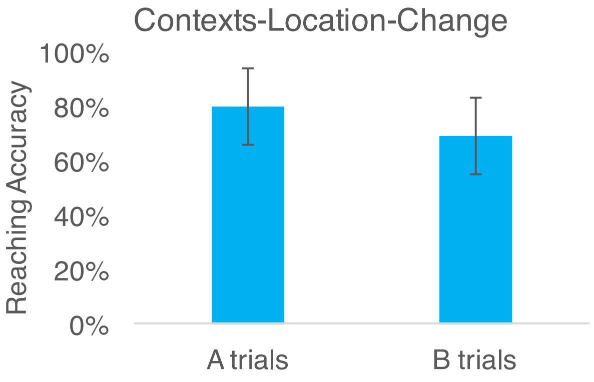

Figure 6. There was no difference in reaching accuracy on A trials and B trials. Error bars reflect 95% confidence intervals.

4.1.5.3 Experiment 2 Discussion

These findings replicate Experiment 1, and show that the improvement in performance on B trials was not driven from using two different toys and that it was not due to the experimenter providing a visual cue indicating where infants should search. Unlike the contexts condition in Experiment 1, performance on A and B trials was statistically equivalent in Experiment 2. Yet Figures 5 and 6 indicate differences that are worth interpreting. If infants are indeed using a learned hierarchical structure to guide reaching behavior, why do infants still make errors at all on B trials? One possibility is that there remains a cost to switching even at a lower-order level (as designated in Figure 1), though that cost is smaller than at a higher-order level and does not drive perseverative behavior (Collins & Frank, 2013; Monsell, 2003). Our data show a greater difference between A trial and B trial accuracy in the contexts condition in Experiment 1, where both the lower-order toy and the hiding location switched, relative to finding no statistical difference between A and B trial accuracy in both contexts conditions in Experiment 2, where the toy stayed the same. Theoretically, this may have reduced the lower-order switch cost.

Experiment 3 tests a critical prediction for the broad hypothesis of this work (Figure 7): Repeat *A trials* that are *inconsistent* with a hierarchical rule structure should result in worse performance than Switch *B trials* that are *consistent* with a rule hierarchical structure. An added benefit of the design of Experiment 3 is that it allows us to ask whether switching the lower order toy-location rule in the contexts condition is less costly than repeating that location when it is paired with a higher-order rule switch. *(287 words)*

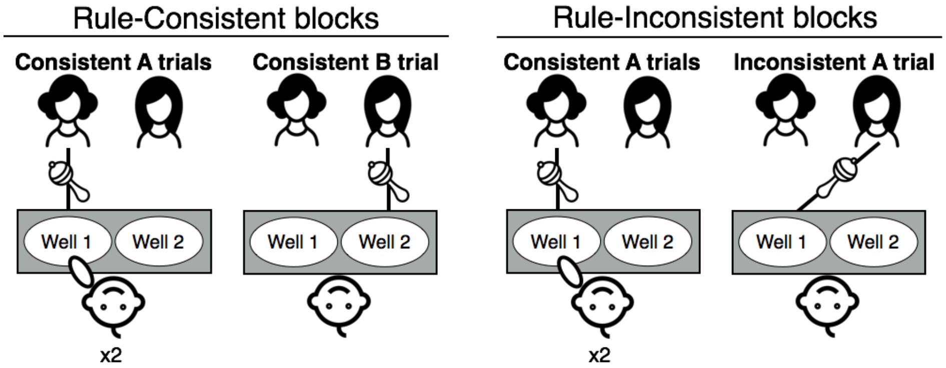

Figure 7. Schematic of the design in Experiment 3. Infants first received two rule-consistent blocks, which had two A trials followed by one B trial, both of which were consistent with the hierarchical rule structure. Infants then received two rule-inconsistent blocks, which had two consistent A trials followed by one inconsistent A trial that violated the hierarchical rule structure.

4.1.6 Experiment 3

4.1.6.1 Method

4.1.6.1.1 Participants

The sample consisted of 14 nine-month-old infants (*M* = 9.32 months, *SD* = .98 months, 10 females, 4 males, 8 white non-Hispanic, 3 black, 1 Hispanic, 1 Asian, and 1 Mixed Race/Other). An additional 4 infants were tested, but excluded from the sample due to fussiness (n = 3) or experimenter error (n = 1). Sample size was determined based on a power analysis with a large effect size (*d* = .90; estimated from Experiment 1), which indicated that 14 infants would provide sufficient statistical power (85%) for the within-subjects design.

4.1.6.1.2 Procedure

The hiding procedure described in the General Method section was used, and the testing session consisted of four blocks: two rule-consistent blocks followed by two rule-inconsistent blocks (Figure 7). In the first rule-consistent block, experimenter 1 hid the toy in the first well two times (rule-consistent A trials), after which experimenter 2 hid the same toy in the other well (rule-consistent B trial). This block was then repeated. After the two rule-consistent blocks, the first rule-inconsistent block occurred, during which experimenter 1 hid the toy in the first well two times (rule-consistent A trials), and then experimenter 2 hid the same toy in the same well (rule-inconsistent A trial). This block was then repeated in the other hiding well. Note that if an infant did not search in the correct location on the second rule-consistent A trial in any of the four testing blocks, it was repeated until the infant searched correctly. Also note that we used two rule-consistent A trials rather than three to shorten the testing session, given that meta-analyses indicate that the number of A trials does not impact perseverative errors when it ranges from one to three trials (e.g., Wellman, Cross, Bartsch, & Harris, 1986).

4.1.6.2 Experiment 3 Results

We compared infants' reaching accuracy by trial type (rule-consistent B trial, rule-consistent A trial, rule-inconsistent A trial) using a repeated-measures ANOVA, which indicated a significant difference in reaching accuracy by trial type, $F(2,26) = 8.881$, $p = .001$ (Figure 8). Follow-up paired-samples two-tailed t tests (Bonferroni corrected alpha = .017) indicated that infants had better performance on rule-consistent A trials relative to rule-inconsistent A trials, $t(13) = 4.039$, $p = .001$, and better performance on the rule-consistent B trials relative to the rule-inconsistent A trials, $t(13) = 2.511$, $p = .026$. There was no difference in performance between the rule-consistent A trials and rule-consistent B trials, $t(13) = 1.255$, $p = .232$. We next compared rule-consistent B trial, rule-consistent A trial, and rule-inconsistent A trial reaching accuracy to chance (50%) using three one-sample two-tailed *t* tests. These tests indicated that accuracy on the rule-consistent B trials was significantly better than chance, *t*(13) = 3.798, *p* = .002, *M* = 82.1%, *SD* = 31.6%, as well as accuracy on the rule-consistent A trials, *t*(13) = 9.455, *p* < .001, *M* = 89.6%, *SD* = 15.7%. However, accuracy on the rule-inconsistent B trials did not differ from chance, *t*(13) = .366, *p* = .720, *M* = 53.6%, *SD* = 36.5%.

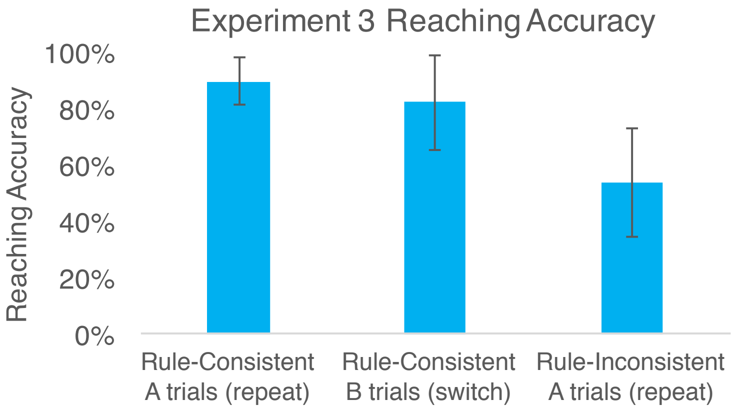

Figure 8. Infants performed significantly better on rule-consistent A trials (repeat) and rule-consistent B trials (switch) relative to rule-inconsistent A trials (repeat). Error bars reflect 95% confidence intervals

4.1.7 General Discussion

We tested a novel explanation of infant performance on the A-Not-B task inspired by a recent model of PFC functional development (Werchan & Amso, 2017) and recent work indicating a role for the infant PFC in the spontaneous structuring of inputs into nested rule structures that can be flexibly updated into working memory (Werchan et al., 2015, 2016). Structuring environmental inputs in this way provides several benefits to the infant learner. For instance, it allows the infant to make predictions about future events, and to generalize learning across similar contexts. Thus, based on an adaptive learning account of PFC function offered in both adult (Collins & Frank, 2013; Collins et al., 2014) and infant (Werchan et al., 2015, 2016) studies, we hypothesized that the A-not-B error may be the byproduct of infants' attempt to learn a higher-order rule structure during the A-Not-B task, and to use this structure to guide search behavior. Thus, what appears to be an error may reflect an adaptive attempt to learn how a higher-order context, in this case the experimenter, organizes separate working memory representations for the lower-order toy-well-action rules.

We tested nine-month-old infants, who consistently show the A-Not-B error, in a modified version of the A-Not-B task, in which a change in the toy's hiding well was associated with a change in the experimenter (higher-order context) hiding the toy. We found that infants had fewer perseverative errors when a change in the toy's hiding well was associated with a change in the experimenter hiding the toy. Follow-up experiments indicated that this reduction in perseverative errors persisted even when the same toy was hidden across A trials and B trials, and that the source of this improvement was not due to infants using the experimenter's left-right seating location as a surface-level visual feature that cues which left-right well to search for the toy in. The critical test of our hypothesis in Experiment 3 demonstrated that infants perform better on switch B trials, which require switching a reach from the A-well to the B-well, that are *consistent* with a learned hierarchical structure, relative to repeat A trials, which require repeating a reach to the A-well, that are *inconsistent* with a learned hierarchical structure. This is even as the switch B trials require infants to inhibit a prepotent motor response while the repeat A trials do not.

These findings suggest that infants' A-not-B errors, in part, reflect an adaptive response strategy as infants learn how a higher-order context organizes lower-order toy-well-action rules. When the higher-order context does not change when the hiding well changes, infants continue to align behavior with the learned hierarchical structure. This explanation also suggests that infants reaching to the A location on B trials may reflect negative transfer of a learned hierarchical structure on standard versions of the task (Perkins & Salomon, 1992), until such time that a new structure is learned.

Our findings are consistent with prior research examining how environmental or perceptual information impact infants' A-not-B errors. For instance, prior work shows that 10-month-old infants are more likely to make A-not-B errors when the experimenters use communicative cues during the task (Topal, Gergely, Miklosi, Erdohegyi, & Csibra, 2008). The use of communicative cues may make the experimenters more salient to the infant, which may create a stronger bias for infants to use the experimenter as a higher-order context. Other work shows that altering the posture of the infant on B trials results in fewer A-not-B errors (Smith, Thelen, Titzer, & McLin, 1999; Thelen, Schöner, Scheier, & Smith, 2001). Similarly, using a different-colored hiding cover on the B trials relative to the A trials results in better performance (Bremner, 1978). These changes in task dynamics or perceptual information on B trials may provide alternative or more distinctive higher-order contextual information that helps cue updating of a new S-A-O rule.

Prior work indicates that infants can use higher-order contextual information to organize inputs into hierarchical rule structures that can be flexibly updated into working memory as early as 8 months of age (Werchan et al., 2015). Related work demonstrates that infants can also learn hierarchical serial patterns from audiovisual event sequences, which they can then generalize to novel sequences (Lewkowicz, Schmuckler, & Mangalindan, 2018). Computational models indicate that the structuring of inputs into hierarchical rule structures of this sort is supported by nested frontostriatal circuitry (Collins & Frank, 2013). Neuroimaging research in both adults (Collins et al., 2014) and infants (Werchan et al., 2016) shows that this learning mechanism is related to increased PFC activation.

Our results expand upon these prior findings to provide a novel explanation of infants' perseverative behavior on the A-Not-B task. Specifically, we demonstrate that infants use relevant higher-order contextual information, such as the experimenter, to structure inputs into latent hierarchical rule structures that guide learning and behavior in the A-Not-B task. There are two important benefits to this type of hierarchical organization (Collins et al., 2014; Collins & Frank, 2013; Donoso et al., 2014). First, organizing the world hierarchically prevents catastrophic interference, meaning that separating information by higher-order contexts ensures that new information does not interfere with prior learning. Second, rule-sets are *latent* and can be transferred to novel contexts without having to learn them anew. Thus, what appears to be an error may reflect an adaptive strategy that supports learning and helps prevent catastrophic interference when new information conflicts with prior learning in different contexts.

In sum, our findings suggest that infants use higher-order contextual information to organize toy-well-action rules in the A-Not-B task. These results suggest that the A-not-B error is a byproduct of an adaptive system that helps infants learn stable and generalizable representations of their environment. Importantly, these findings also have broader impact on how we have typically studied the PFC and infant behavior more generally. We tend to impose our own constructs and cognitive biases on infants, and conduct studies that are inherently designed to look for failures and limitations (Karmiloff-Smith, 1995). Yet, to have a fuller understanding of the developing mind and brain, our findings illuminate the importance of considering infants as different organisms with unique goals for learning and behavior.

Chapter 5

5. Reciprocal Influences of PFC-Dependent Rule Learning on Downstream Sensory Systems

The prior chapters provide behavioral and neuroimaging evidence supporting the hypothesis that the PFC is adapted for learning and action appropriate to infants in their environment through its role in hierarchical rule learning. An additional prediction arising from an ecological approach to PFC development is that the PFC simultaneously modulates the functional response properties of posterior cortical regions over ontogenetic development to adapt for processing behaviorally-relevant information (Figure 1, p. 36). Importantly, this view predicts that these processes are reciprocal, such that the PFC supports the organization of multimodal inputs into generalizable rule structures, which in turn shapes developing information processing systems to best adapt for processing relevant information. This broad hypothesis aligns with existing views suggesting that the PFC supports the specialization and organization of posterior neural regions (M. H. Johnson, 2000, 2011; M. H. Johnson, Jones, & Gliga, 2015; Thatcher, 1992), as well as work indicating that PFC exerts modulatory control over neural activity within posterior regions in adults (Gilbert & Li, 2013; E. K. Miller & Cohen, 2001).

Thus, the aim of the fourth and final study was to empirically test this framework in infants using the known feedforward and feedback connectivity between PFC/visual cortex. To examine this, 9-month-old infants were presented with visual cue-reward pairs in a rule learning and generalization task, where simple visual features (color or shape) could act as higher-order contexts that organize visual inputs for learning. Critically, infants' visual attention biases to the features of color and shape were measured before and after the rule learning and generalization task. The aims were to test (1) whether infants use PFC to form rule structures that organize visual inputs for learning, and (2) whether this in turn influences what infants subsequently attend to in a novel visual context. The findings from this work will contribute to a better understanding of how the PFC influences developing cortical circuitry to adapt for processing information most relevant for learning and behavior.

5.1 Top-Down Knowledge Rapidly Acquired Through PFC-Dependent Rule Learning Biases Visual Attention in 9-Month-Old Infants

5.1.1 Abstract

Visual attention is an information-gathering mechanism that supports the emergence of complex perceptual and cognitive capacities. Yet, little is known about how the infant brain learns to direct attention to information that is most relevant for learning and behavior. Here we address this gap by examining how top-down knowledge rapidly acquired through PFC-dependent rule learning influences 9-month-old infants' visual attention in a novel visual context. In Experiment 1, we found that infants structured simple visual inputs into generalizable rules, which then biased attention towards behaviorally-relevant features in a novel visual context. In Experiment 2, we found that increased functional connectivity between the PFC and visual cortex was related to the efficacy of rule learning. These findings provide new insights into how the infant brain learns to flexibly select features from the cluttered visual world that support adaptive behavior.

5.1.2 Introduction

Visual attention is a fundamental capacity that enables infants to gather information and interact with their environment. Many prior studies have examined how visual attention influences learning and the emergence of more complex perceptual and cognitive capacities across infancy and childhood (Amso & Johnson, 2006; Cheng, Kaldy, & Blaser, 2019; S. P. Johnson, Amso, & Slemmer, 2003; Markant & Amso, 2013, 2016; Markant, Oakes, & Amso, 2015; Ross-Sheehy, Oakes, & Luck, 2011; Wu & Kirkham, 2010). Yet, less is known about how the infant brain *learns* to direct attention to information that is most relevant for learning and behavior in early life. The current paper addresses this gap in our understanding of the development of visual attention by examining how rapidly acquired top-down knowledge influences infants' attention to behaviorally-relevant information in a novel visual context.

Visual attention shows rapid developmental changes over the first year of postnatal life (Amso & Scerif, 2015; Oakes & Amso, 2018). Visual attention can either be bottom-up driven, based on saliency maps of the visual-spatial environment (e.g., Althaus & Mareschal, 2012; Amso, Haas, & Markant, 2014; Frank, Vul, & Johnson, 2009), or top-down driven. Top-down visual attention, which is guided based on prior experience or current behavioral goals, has been demonstrated by 4 months of age (M. H. Johnson & Vecera, 1996; Tummeltshammer & Amso, 2018; Tummeltshammer, Mareschal, & Kirkham, 2014; Werchan, Collins, Frank, & Amso, 2015). Attention is often thought of in terms of a "spotlight" that enhances processing of relevant information by biasing attention towards some stimuli over others that are simultaneously competing for attentional resources (Carrasco, 2011). This biasing occurs through two mechanisms: bottom-up sources that transmit sensory information from lower-order to higher-order cortical areas via feedforward cortical pathways, and top-down sources that carry information regarding current behavioral goals from higher-order to lower-order cortical areas via feedback cortical pathways (Amso & Scerif, 2015; Carrasco, 2011; Desimone & Duncan, 1995). Thus, these bottom-up and top-down mechanisms can shift the attentional spotlight as a function of both low-level stimuli characteristics and internal behavioral goals.

Previous work has shown that infants' attention can be biased by learning from the environment. For example, Markant, Oakes, and Amso (2016) used a spatial cueing task to bias 9-month-old infants' attention to own- or other-race faces. The authors found that infants discriminated faces in the focus of the attention bias, indicating that attention engagement influenced the efficacy of face discrimination. Other work has found that 6-month-old infants can rapidly extract top-down knowledge about spatial covariations from simple arrays, and then use this contextual knowledge to guide visual search (Tummeltshammer & Amso, 2018). Recently, we also showed that infants as young as 8-months of age can use the prefrontal cortex (PFC) to structure visual inputs into abstract rules that support learning in novel contexts (Werchan et al., 2015; Werchan, Collins, Frank, & Amso, 2016). In this work, we found that 8-month-old infants can use simple visual features, such as the shape of an object, as higher-order contexts to structure simpler color-location associations into hierarchically-organized rules (e.g., if the context is "square", then "red" predicts a cartoon in "location 1" and "blue" predicts a cartoon in "location 2"). Importantly, infants were able to use these rule structures to generalize learning in new contexts (e.g., to new shapes). Intuitively, directing attention to the relevant higher-order features that cue these rule representations would facilitate faster and more efficient generalization of these rules by facilitating the recognition and selection of relevant features cueing these rules in novel contexts. Thus, these findings suggest that PFC-dependent rule learning might be a mechanism that biases infants' attention towards features that are relevant for learning and behavior.

In adults, converging evidence from neuroimaging and anatomical studies has established the PFC as a source of top-down attention signals that modulate processing in early visual areas (Desimone & Duncan, 1995; Gilbert & Li, 2013; Noudoost, Chang, Steinmetz, & Moore, 2010; Shomstein & Gottlieb, 2016). The PFC is a higher-order area that is involved in encoding top-down knowledge about task-relevant goals and abstract rules that support flexible and goal-directed control of behavior (Badre, 2008; Cohen, Braver, & Brown, 2002; Kolb et al., 2012; Miller & Cohen, 2001; O'Reilly, 2006; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005). This region is also highly interconnected, sending and receiving long-range projections from nearly all sensory and motor systems, making it well-suited to modulate processing in posterior neural regions (Gilbert & Li, 2013; Miller & Cohen, 2001).

Evidence from axonal tract-tracing studies in monkeys reveal an intricate anatomical network of reciprocal corticocortical connections between areas of the PFC and extrastriate visual areas (Barbas, 2000; Petrides & Pandya, 2001; Ungerleider, Gaffan, & Pelak, 1989; Webster, Bachevalier, & Ungerleider, 1994). Functional interactions within these corticocortical connections are thought to be the basis for PFC modulation of neuronal activity in early visual areas (Baluch & Itti, 2011; Paneri & Gregoriou, 2017). In support of this hypothesis, functional neuroimaging studies in adult humans have shown that activity in PFC areas is correlated in a task specific manner with activity in posterior visual regions (Corbetta & Shulman, 2002; Gazzaley et al., 2007; Kastner & Ungerleider, 2000; Morishima et al., 2009; Rossi, Pessoa, Desimone, & Ungerleider, 2009; Taylor et al., 2007). Additionally, studies of patients with PFC lesions and studies using transcranial magnetic stimulation (TMS) to perturb PFC function provide direct causal evidence that the PFC exerts top-down modulatory control over processing in visual cortex in adults (Barceló, Suwazono, & Knight, 2000; Capotosto et al., 2009; Ruff et al., 2008; Taylor et al., 2007; Zanto, Rubens, Thangavel, & Gazzaley, 2011). Yet, it is unclear whether similar mechanisms operate in young infants with immature prefrontal and perceptual systems.

As such, the current study addresses the impact of PFC-dependent rule learning on influencing subsequent downstream visual processing in 9-month-old infants in two experiments. We tested 9-month-olds given that infants of this age are capable of both engaging in PFC-dependent rule learning mechanisms (Werchan et al., 2015, 2016) and engaging in top-down guidance of spatial attention (Amso & Johnson, 2006, 2008; Tummeltshammer & Amso, 2018). In our first experiment, we adapted a PFC-dependent rule learning paradigm previously used with infants (Werchan et al., 2015, 2016) to examine whether acquiring top-down knowledge influences infants' visual processing in a novel visual context. In our second experiment, we used functional near-infrared spectroscopy (fNIRS) to explore the neural correlates of these processes. Examining the impact of PFC-dependent rule learning on shaping visual processing may provide mechanistic insights into how the infant brain *learns* to efficiently direct attention to information that is most relevant for learning and behavior over ontogenetic development.

5.1.3 Experiment 1

In our first experiment, we used a behavioral paradigm to examine whether top-down knowledge acquired through abstract rule learning influences infants' downstream visual processing. To test this hypothesis, we presented 9-month-old infants with visual cue-reward pairs in a rule learning and generalization task modeled after prior work, where simple visual features (color or shape) can act as higher-order contexts that organize visual inputs for learning (Werchan et al., 2015). Critically, we used an attention bias priming task (Werchan et al., under revision) to measure infants' attention biases to the visual features of color and shape before and after the rule learning and generalization task. We predicted that individual differences in the efficacy of rule learning would correlate with the change in infants' attention bias to the relevant higher-order visual feature (shape or color) based on condition.

5.1.3.1 Experiment 1 Method

5.1.3.1.1 Participants

The final sample consisted of 40 nine-month-old infants (*M* = 9.45 months, *SD* = .67 months, 20 females, 20 males, 29 white non-Hispanic, 3 black, 5 Hispanic, 2 Asian, and 1 Mixed Race/Other). Sample size was determined based on an a priori power analysis with a large effect size (*d* = .4) estimated from prior work (Werchan et al., 2015) at 90% power, which indicated that approximately 20 infants per condition would provide sufficient statistical power. An additional 3 infants were tested but excluded from the final sample for failing to complete the experiment due to fussiness or crying. Infants were randomly assigned to a Shape Contexts condition (*N* = 20, *M* = 9.59 months, *SD* = .84

months) or a Color Contexts condition ($N = 20$, $M = 9.45$ months, $SD = .43$ months). Infants were recruited through community advertisements and through birth records from the state department of health. Infants were prescreened for premature birth (< 36 weeks), low birth weight (< 5lb), or a history of serious health problems. The Brown University Institutional Review Board approved the study, and parental consent was obtained prior to testing. Families were compensated for time and travel to our laboratory.

5.1.3.1.2 Eye Tracking Apparatus

Stimuli were presented via SMI Experiment Center software on a 24" monitor. Eye tracking was collected using an SMI REDn-Scientific apparatus (Teltow, Germany). Infants were seated on their parents' lap approximately 60 cm from the monitor. Before the study began, infants' point of gaze was calibrated by presenting five target stimuli, one in the middle of the monitor and one in each of the four corners of the monitor. The point of gaze was validated by presenting one stimulus in each of the four corners of the monitor. Calibration was repeated if deviations were greater than 2°. Areas of interest (AOIs) were defined in the native SMI software-analysis package BeGaze.

5.1.3.1.3 Procedure

The study consisted of (1) a baseline attention bias priming task, (2) a rule learning and generalization task, (3) a post-learning attention bias priming task.

5.1.3.1.3.1 Attention bias priming task. Infants received four trials during the attention bias priming task, which was used to measure infants' attention biases to the visual features of color and shape at baseline and after the rule learning task (Figure 1). During each trial, an attention-getting stimulus was first presented in the middle of the screen to center infants' point-of-gaze. The trial was initiated once the experimenter judged that the infant was looking at the attention-getter. The prime stimulus was presented in the center of the screen for 1,000-ms. The prime stimulus then disappeared and the test stimuli appeared simultaneously on the left and right sides of the screen for 2,000-ms. On each trial, one test stimulus matched the prime stimulus in shape but differed in color (shape-match item), and the other test stimulus matched the prime stimulus in color but differed in shape (color-match item). The right/left locations of the color-match and shape-match items were counterbalanced across trials, and the order of trials was randomized for each infant. Different sets of stimuli were used for the baseline and post-learning attention bias priming task, which was counterbalanced across infants.

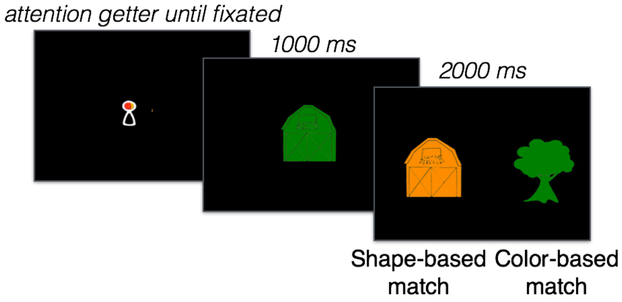

Figure 1. Example of one trial in the attention bias priming task, which was used to measure infants' attention biases to color and shape.

SMI BeGaze software was used to calculate the duration of looking to each test item by summing across all observed samples in which an infant's point of gaze fell within that item's AOI. The first 150-ms bin was excluded from the analysis for each test trial to account for the time required for infants to make a saccade away from the central prime stimulus toward either the left or right test item after the appearance of the test display.

Infants' color bias scores were then calculated for each trial by subtracting the duration of time spent looking at the shape-match item from the duration of time spent looking at the color-match item, and infants' shape bias scores were calculated for each trial by subtracting the duration of time spent looking at the color-match item from the duration of time spent looking at the shape-match item. We then used these baseline and post-learning color and shape bias scores to measure the change in infants' attention biases to color and to shape. To calculate the change in infants attention bias to color, we subtracted infants' baseline color bias scores from the post-learning color bias scores. Thus, positive color difference scores indicate a greater attention bias to color after learning relative to baseline, and difference scores near zero indicate no change in infants' attention biases. We also calculated the change in infants attention bias to shape analogously by subtracting infants' baseline shape bias scores from the post-learning shape bias scores. Thus, positive shape difference scores indicate a greater attention bias to shape after learning relative to baseline.

5.1.3.1.3.2 Rule learning and generalization task. During the rule learning task, infants were presented with eight different visual cue-reward pairs (Figure 2A). In the Shape Contexts condition, the central cues varied by four different shapes (the higher-order context), and by two different colors (the lower-order feature). In the Color Contexts condition, the cues varied by four different colors (the higher-order context), and by two different shapes (the lower-order feature). We used more shapes than colors in the Shape Contexts condition and more colors than shapes in the Color Contexts condition to help bias infants to learn a generalizable rule structure, rather than learning flat individual cue-reward associations (Collins & Frank, 2013). Infants were presented with a total of 16 trials, which were divided into two blocks of 8 trials. Thus, each of the eight cue-reward pairs were presented once per block. The presentation order of the trials was randomized in each block.

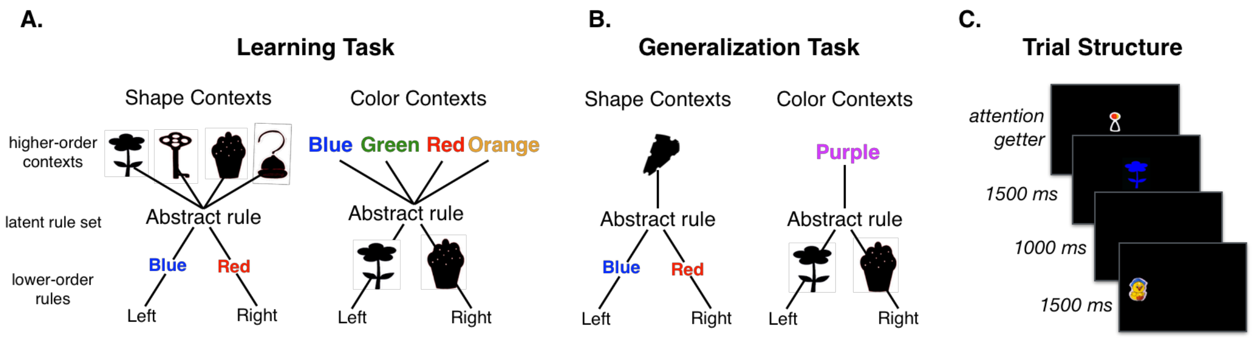

Figure 2. Example of the hierarchical structure used during the learning task (A) and generalization task (B). Each trial consisted of the central cue, followed by a cartoon reward presented on the right or left of the screen (C).

After the learning task, infants were then presented with two novel cue-reward pairings during a generalization task (Figure 2B). In the Shape Contexts condition, the cue-reward pairings were characterized by having a novel higher-order shape feature, but the same lower-order color-location rules as in the previous learning task. In the Color Contexts condition, the pairings were characterized by having a novel higher-order color feature, but the same lower-order shape-reward rules as in the previous learning task. Infants were presented with a total of 8 trials in a randomized order.

At the start of each trial during the learning and generalization tasks, infants' point of gaze was centered by presenting an attention-getting stimulus in the middle of the screen. Once the experimenter judged that the infant was looking at the central attention-getter, the trial was initiated. During each trial, the central cue was first displayed in the center of the screen for 1,500-ms. The cue then disappeared, and after a 1,000-ms delay a cartoon reward appeared on the left or right of the screen for 1,500-ms (Figure 2C).

Infants' eye movement reaction times from the onset of the central cue to the arrival at the reward location were recorded as an index of learning. We averaged across every four consecutive trials in the learning task to create four learning bins, and we defined learning as a decrease in infants' eye movement reaction times across learning bins. Generalization scores were calculated by subtracting the average reaction times during the generalization task from the average reaction times during the generalization task. Thus, positive values indicate generalization of the learned rule, and values near or less than zero indicate a lack of generalization.

5.1.3.2 Experiment 1 Results

We first examined infants eye movement reaction times during the rule learning task. We conducted a repeated-measures ANOVA with Learning Bin (1, 2, 3, 4) as a within-subjects variable, Condition (Shape Contexts, Color Contexts) as a between-subjects variable, and infants' eye movement reaction times as the dependent variable. This analysis revealed a main effect of Learning Bin, $F(3,114) = 6.578, p < .001$, and a main effect of Condition, $F(1,38) = 7.907, p = .008$. These findings indicate that infants learned to predict the left-right reward location with trial exposure. They also show that infants in the Color Contexts condition showed faster learning than infants in the Shape Contexts condition, as evidenced by infants having faster reaction times across all bins of learning in the Color Contexts condition (M = 2573.70-ms, SD = 41.14-ms) relative to the Shape Contexts condition (M = 2770.03-ms, SD = 42.80-ms).

We next examined how infants' baseline attention biases impacted learning. At baseline, we found that all infants had a significantly greater attention bias to color information over shape information, $t(39) = 3.416$, $p = .001$. We also conducted a two-tailed independent samples t test to verify that there were no differences in baseline attention biases to color between the Color Contexts condition (M = 108.35-ms, SD =

62.03-ms) and the Shape Contexts condition (*M* = 179.83-ms, *SD* = 57.65-ms). This test revealed no differences between conditions, $t(38) = 0.844$, $p = .404$. Next, we analyzed how infants' baseline attention biases related to individual differences in learning. We generated a learning score by subtracting infants' average eye movement reaction times during the last bin of learning from the first bin of learning. We then examined correlations between learning scores and infants' baseline attention biases to color and shape. Results indicated that attention biases to shape were positively correlated with learning scores in the Shape Contexts condition, $r(20) = 0.481$, $p = .032$, and attention biases to color were positively correlated with learning in the Color Contexts condition, $r(20) = 0.531$, $p = .016$. These results indicate that baseline attention biases influence the efficacy of learning, such that greater color-based attention biases supports learning of rules cued by color, and greater shape-based attention biases supports learning of rules cued by shape.

We then examined whether infants generalized these newly-formed rules to novel shape or color contexts. We conducted a repeated-measures ANOVA with Task (Learning, Generalization) as a within-subjects variable, Condition (Shape Contexts, Color Contexts) as a between-subjects variable, and infants' eye movement reaction times as the dependent variable. Results indicated that all infants as a group had faster reaction times in generalization relative to learning as evidenced by a main effect of Task, $F(1,38) = 23.010$, $p < .001$. We also found a main effect of Condition, $F(1,38) = 20.598$, $p < .001$, and an interaction between Task and Condition, $F(1,38) = 8.349$, $p = .006$. We followed up on this interaction by conducting two two-tailed paired-samples *t* tests comparing reaction times during learning and generalization separately for each condition. These analyses revealed a significant difference between learning and generalization in the Color Contexts condition, *t*(19) = 7.647, *p* < .001, but not in the Shape Contexts condition, *t*(19) = 1.103, *p* = .284 (Figure 3). Thus, these analyses provide evidence that, as a group, infants generalized a newly-formed rule to a novel context, and that infants in the Color Contexts condition showed better learning and generalization than infants in the Shape Contexts condition. This finding may be due to the greater baseline attention bias to color demonstrated by all infants, which may have facilitated rapid learning and generalization in the Color Contexts condition.

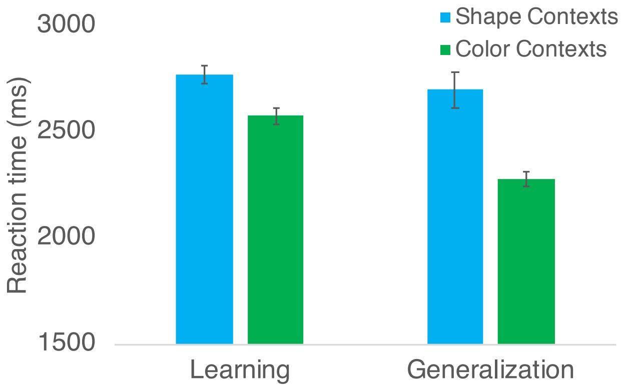

Figure 3. Infants' eye movement reaction times during the learning and generalization tasks in the Shape Contexts and Color Contexts conditions.

Finally, we examined whether these newly-formed rule structures influenced what visual features infants subsequently attended to in a novel visual context. Our results revealed that individual differences in infants' generalization scores was correlated with the change in infants' attention bias to shape in the Shape Contexts condition, *r*(20) = 0.525, *p* = .018 (Figure 4A), and to color in the Color Contexts condition, *r*(20) = 0.619, *p* = .004 (Figure 4B). These results provide behavioral evidence that learning abstract rules influences infants' visual processing by biasing attention to relevant higher-order visual feature categories that can be used to support learning in novel contexts.

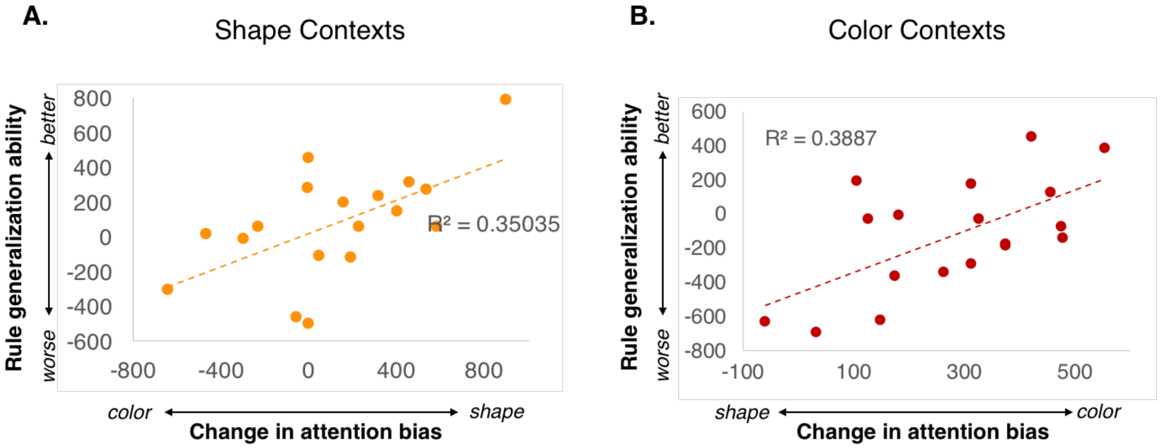

Figure 4. The relation between infants' generalization ability and changes in attention biases to shape information in the Shape Contexts condition (A) and to color information in the Color Contexts Condition (B).

5.1.4 Experiment 2

In Experiment 1, we found behavioral evidence indicating that rapidly acquired top-down knowledge influences feature-based attention in infants. In adults, top-down visual attention is thought to be mediated by functional interactions between the PFC and visual cortex (Gilbert & Li, 2013; Baluch & Itti, 2011; Paneri & Gregoriou, 2017). Therefore, in a second experiment, we examined how individual differences in infants' cortical activity during rule learning relates to the efficacy of generalization and changes in infants' attention biases. We tested an independent sample of infants using the same experimental paradigm as in Experiment 1, and we used functional near-infrared spectroscopy (fNIRS) to measure infants' frontal and occipital cortical activity during learning. We predicted that increased PFC/visual cortex connectivity during learning would relate to both the efficacy of infants' rule learning and generalization, as well as changes in infants' attention biases to the relevant higher-order context.

5.1.4.1 Experiment 2 Method

5.1.4.1.1 Participants

The final sample consisted of 19 nine-month-old infants (*M* = 9.48 months, *SD* = .20 months, 12 females, 7 males, 16 white non-Hispanic, 2 Hispanic, and 1 black). An additional 5 infants were tested, but their data were discarded due to equipment malfunction (*n* = 2) and fussiness or crying resulting in failure to complete the experiment (*n* = 3).

5.1.4.1.2 Eye Tracking Apparatus

Stimuli were presented via SMI Experiment Center software on a 24" monitor. All eye tracking procedures were identical to those described in the Method section of Experiment 1.

5.1.4.1.3 Procedure

The exact same experimental methods described in Experiment 1 were used in Experiment 2, with the following exceptions: each of the three blocks during the rule learning and generalization task was preceded by a 10-s white fixation cross on a black background to allow the hemodynamic response to return to baseline prior to the start of each block. Additionally, we only tested infants in the Shape Contexts condition.

5.1.4.1.4 fNIRS Recording

Infants' frontal and visual cortical activity was recorded during the learning task using a TechEn CW6 NIRS system with wavelengths set at 695 and 830 nm. Raw signals were continuously sampled at 50 Hz. An array consisting of 9 optodes (3 sources and 6 detectors, resulting in 6 source-detector channels) with an interoptode separation of 3 cm was placed over infants' visual association cortex and frontal brain regions. The array was fixed on sturdy, flexible plastic to ensure that the distance between the sources and detectors remained constant at 3 cm. The optode array was attached inside of an adjustable neoprene headband to secure the optodes to the scalp. The array was placed over infants' scalps using standardized coordinates corresponding to the right and left lateral PFC (F3/F4 in the 10-20 international EEG system) and visual association cortex (O2 in the 10-20 international EEG system). This positioning aligns with the 10–20 coordinates used for localizing frontal and visual cortex activation in prior fNIRS work with infants (Bortfeld, Wruck, & Boas, 2007; Emberson, Richards, & Aslin, 2015; Werchan, Baumgartner, Lewkowicz, & Amso, 2018; Werchan et al., 2016).

After recording, the fNIRS data were preprocessed prior to analyses using HomER 2.0 software (Huppert, Diamond, Franceschini, & Boas, 2009). We first digitally band-pass filtered the raw signals at 0.01– 0.1 Hz to remove systematic physiological and motion artifacts (Homae et al., 2010; White et al., 2009). We then calculated the change in optical density for each wavelength relative to the 10-s baseline prior to block onset, during which a black screen with a white fixation cross was presented. Next, we used the modified Beer-Lambert law to calculate changes in the concentration of oxygenated and deoxygenated hemoglobin from the changes in optical density. Afterwards, we screened for motion artifacts by identifying signal fluctuations 5M over a 0.5-s range in each channel (Emberson et al., 2015; Lloyd-fox et al., 2009). Finally, changes in oxygenated hemoglobin (relative to the 10-s baseline) in each of the 6 source-detector channels were exported for subsequent analysis by averaging across every 4-s of each 32-s block starting 4-s after block onset to remove serial autocorrelation and to eliminate the need to make assumptions about the shape of the hemodynamic response in subsequent analyses. We limited analyses to the period starting 4 s after stimulus onset based on previous studies that have seen that this is the typical delay in the hemodynamic response function (HRF) initiation in infants (Taga & Asakawa, 2007; Werchan et al., 2016). This created a total of 7 time intervals for each block during the learning task.

The 6 source-detector channels were divided and averaged into three regions of interest for subsequent data analysis, with the two left frontal channels corresponding to left lateral PFC, the two right frontal channels corresponding to right lateral PFC, and the two posterior occipital channels corresponding to visual association cortex. These regions of interest were verified by estimating measurement sensitivity to these cortical regions (based on the positioning of the optode array referenced to standardized 10-20 coordinates as described above) using AtlasViewer NIRS image reconstruction tools (Aasted et al., 2015; Figure 5). We restricted our analyses to cortical activations during the first block of learning, given that there were increased motion artifacts and reduced data quality in the second block of learning, and since the PFC tends to be implicated more in early relative to late stages of learning (Kelly & Garavan, 2005).

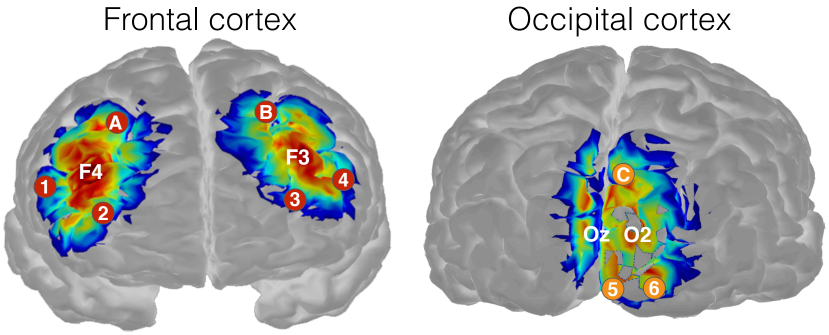

Figure 5. The fNIRS optode array was placed over frontal and occipital regions. Image reconstruction (see Aasted et al., 2015) shows the estimated cortical regions recorded during the study. Warmer colors indicate regions with higher measurement sensitivity, reflecting better ability to detect small signal changes relative to cooler colored regions.

5.1.4.2 Experiment 2 Results

5.1.4.2.1 Behavioral results

We conducted a repeated-measures ANOVA with Learning Bin (1, 2, 3, 4) as a within-subjects variable and infants' eye movement reaction times as the dependent variable. This analysis revealed a trending main effect of Learning Bin, $F(3,51) = 1.715$, $p = .176$, indicating that infants learned to predict the left-right reward locations with trial exposure. Moreover, we also found that infants generalized this abstract rule to novel contexts, as shown by faster reaction times during generalization relative to learning, $t(18) = 4.327$, $p < .001$ (Figure 6). Critically, individual differences in infants' generalization scores were also correlated with the change in infants' attention bias to shape from baseline to post-test, $r(19) = 0.653$, $p = .002$ (Figure 7). Together, these results replicate the behavioral findings from Experiment 1 in an independent sample of infants.

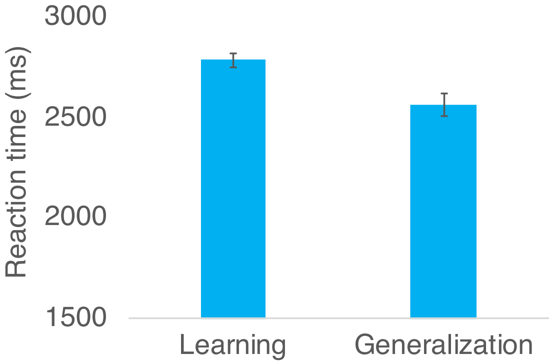

Figure 6. Infants' eye movement reaction times during the learning and generalization tasks in Experiment 2.

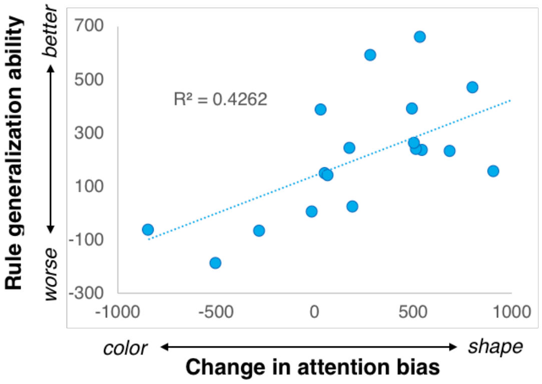

Figure 7. Relation between infants' generalization ability and change in attention bias in Experiment 2.

5.1.4.2.2 fNIRS results

We next examined how infants' cortical activity related to individual differences in rule learning and changes in infants' attention biases. We first conducted an omnibus repeated-measures ANOVA on infants' cortical activations using Region (left PFC, right

PFC, Visual Cortex) and Time Block (seven 4-s bins) as within-subjects factors and infants' generalization scores and attention bias change as continuous variables. This test revealed significant interactions between Region, generalization scores, and the change in infants' shape bias, Wilks' lambda = .535, *F*(2,14) = 6.091, *p* = .013, and between Region, Time Block, generalization scores, and the change in infants' shape bias, Wilks' lambda = .033, *F*(12,4) = 9.904, *p* = .020. We then conducted planned Helmert contrasts comparing the right and left PFC activations to the visual cortex, and the right PFC to the left PFC. These analyses indicated that the right and left PFC differed from the visual cortex in the Region x Generalization score interaction, *F*(1,15) = 6.572, *p* = .022, as well as in the Region x Shape Bias change interaction, *F*(1,15) = 25.656, *p* < .001, and in the Region x Generalization score x Shape Bias change interaction, *F*(1,15) = 9.275, *p* = .008. There were no differences between the right and left PFC for any of these interactions, all *F*s < 2.228, all *p*s > .156.

We then collapsed across the right and left PFC and followed up on these interactions by examining activations across Time Block separately by region, including infants' generalization scores and attention bias change scores as continuous variables. Results indicated that visual cortex activations interacted with the change in infants' attention bias to shape, Wilks' lambda = .205, *F*(6,10) = 6.447, *p* = .005, and with infants' generalization scores and the change in infants' attention bias to shape, Wilks' lambda = .349, *F*(6,10) = 3.109, *p* = .055. These interactions indicate that infants with better generalization scores and greater changes in their attention biases to shape had higher visual cortex activation during learning. For the PFC, we similarly found that cortical activations interacted with infants' generalization scores, Wilks' lambda = .267, *F*(6,10) =

4.579, *p* = .017, indicating that infants with higher PFC activation showed better generalization and had a greater change in attention biases to shape.

We next examined how functional connectivity between the PFC and visual cortex during learning related to individual differences in rule learning and generalization. Following the methods in prior fNIRS work exploring task-based functional connectivity in infants (Homae, Watanabe, Nakano, & Taga, 2011; Keehn et al., 2013), we calculated a Pearson's *r* value for each infant by temporally correlating the PFC activations (collapsing across the left and right PFC) with the visual association cortex activations across the seven averaged time bins during the rule learning task. We then analyzed how this functional connectivity value related to infants' generalization scores. Results indicated that higher functional connectivity was related to better generalization performance, *r*(19) = 0.643, *p* = .003 (Figure 8), but did not correlate with the change in infants' attention bias to shape, *r*(19) = 0.309, *p* = .199. This adds further support, in combination with the behavioral data, that the change in infants' attention bias to the relevant higher-order feature was related to the efficacy of rule learning and generalization.

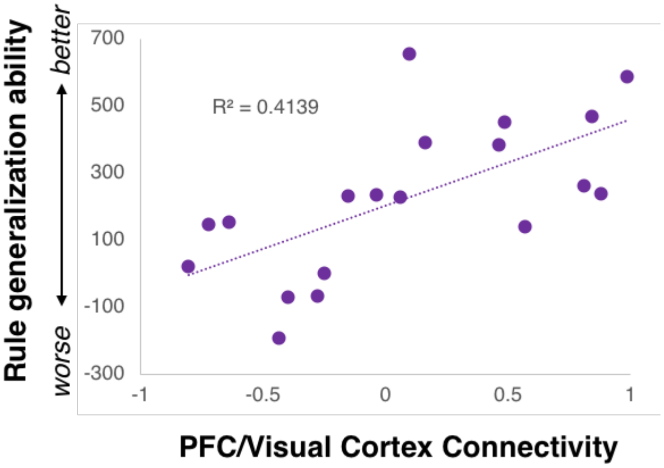

Figure 8. Relation between infants' generalization ability and PFC/visual cortex functional connectivity during learning.

5.1.5 General Discussion

Visual attention is a fundamental capacity that supports the flexible selection of information based on relevant rules and goals that guide adaptive behavior across time and contexts. In adults, top-down visual attention is thought to be mediated by functional interactions between the PFC and visual cortex (Gilbert & Li, 2013; Baluch & Itti, 2011; Paneri & Gregoriou, 2017). Yet, it is unclear whether similar mechanisms operate in infants with immature frontal and visual systems. We recently showed that 8-month-old infants can use the PFC to structure visual inputs for learning (Werchan et al., 2016). In addition, other work has found that infants are capable of using top-down knowledge to guide visual search as young as 6 months of age (Tummeltshammer & Amso, 2018). As such, in the current study we examined whether infants can use top-down knowledge rapidly acquired through PFC-dependent rule learning mechanisms to modulate visual activity for processing behaviorally-relevant stimuli.

In our first experiment, we found clear evidence that rule learning influenced infants' visual processing in a novel visual context. Specifically, our results indicated that infants used a simple visual feature (color or shape) as a higher-order context to organize visual inputs into rules for learning. Critically, rule learning resulted in a change in infants' attention biases to the higher-order visual feature category (color or shape) that cued these rule structures. In addition, we also saw that individual differences in infants' baseline attention biases correlated with the efficacy of rule learning, such that infants with a stronger attention bias to color showed better learning of rules that use color as a higher-order context. These results provide behavioral evidence that learning abstract rules shapes infants' visual processing and attentional biases to relevant information in a novel visual context.

In our second experiment, we examined the neural underpinnings of these processes by using fNIRS to record frontal and visual cortex activity during learning. We found that infants' with better generalization performance and greater changes in attention biases to shape showed higher PFC activation during learning. Mirroring these findings, we also found that higher visual cortex activation during learning was related to greater changes in infants' attention biases and better generalization performance. These findings suggest that greater PFC modulation over visual cortex during learning might support rule learning and generalization, which in turn leads to a greater attention bias to the relevant higher-order feature cueing these rule structures. In support of this interpretation, our results also revealed that individual differences in infants' functional connectivity between the PFC and visual cortex during learning was correlated with the efficacy of rule learning and generalization, such that infants with higher functional connectivity showed better learning and generalization. These results mirror prior findings showing dorsolateral PFC involvement in similar rule learning tasks in infants (Werchan et al., 2016) and adults (Collins, Cavanagh, & Frank, 2014), as well as prior work implicating functional interactions between the PFC and visual cortex in modulating visual processing in adults (Corbetta & Shulman, 2002; Gazzaley et al., 2007; Kastner & Ungerleider, 2000; Morishima et al., 2009; Rossi et al., 2009; Taylor et al., 2007).

These results are somewhat surprising, particularly given that resting state functional connectivity in human infants is dominated by short-range intra-cortical connections relative to long-range intercortical connections (Fransson, Åden, Blennow, &

Lagercrantz, 2011; Gao et al., 2011). Yet, despite long-range functional connectivity being relatively immature in infancy, these long-range anatomical connections are in place at birth (Goldman-Rakic, 1987), and functional connectivity within long range corticocortical connections is evident by 6-9 months of age (Fransson et al., 2007). In addition, a recent study reported similar behavioral findings showing that top-down knowledge guides visual search in simple spatial arrays in infants as young as 6 months of age (Tummeltshammer & Amso, 2018). Thus, our behavioral and neuroimaging findings provide preliminary evidence that corticocortical connections between the PFC and visual cortex are capable of modulating learning and guiding visual processing in infants. These findings add to growing evidence showing that infants can use top-down knowledge to guide visual processing, as well as add new insights into the neural circuitry that supports these processes in infants. However, it is important to note that while examining functional connectivity analysis is informative for describing the neural networks that may be involved in top-down modulation of visual processing in infants, it cannot be used to make statements of causality or directionality.

A secondary finding from this study was that infants showed better learning and generalization of abstract rules when color was used as a higher-order context relative to when shape was used as a higher-order context. In addition, we also observed that all infants demonstrated a stronger baseline attention bias to color relative to shape, and moreover, infants' baseline attention biases correlated with the efficacy of rule learning. We interpret these results as evidence that infants' learning and generalization of abstract rules may be largely influenced by infants' pre-existing attention biases, such that greater baseline attention biases to color facilitates better learning of rule structures that are cued by color. As infants' attention biases shift towards a stronger shape-based attention bias towards the end of the first year of life (Werchan et al., in press), our results would predict that infants would begin to show better learning of abstract rules cued by shape information. More generally, this result also raises the possibility that infants' baseline attention biases and visual processing may scaffold PFC function and learning more broadly. This idea is supported by growing evidence indicating that attention biases impact the encoding and maintenance of information in working memory (Astle et al., 2015; Barnes, Woolrich, Baker, Colclough, & Astle, 2016; Shimi, Kuo, Astle, Nobre, & Scerif, 2014) as well as the efficacy of rule switching (Diamond & Kirkham, 2005; Erb, Moher, Song, & Sobel, 2017). It also aligns with views suggesting that the quality of feedforward visual processing constrains PFC function and executive attention processes (Amso & Scerif, 2015; Werchan & Amso, 2017). Future work is needed to more precisely determine the influence of attention biases and the quality of feedforward visual processing more generally on PFC function and learning in infancy.

Our findings also raise the possibility that similar processes might support the emergence of more complex cognitive and perceptual systems that help the developing child achieve an adaptive fit with their unique ecological niche. For instance, while infants show a preference to look at faces from birth (e.g., Johnson et al., 1991), behavioral and neural specialization for face processing continues to develop well into adolescence (e.g., Aylward et al., 2005; Golarai, Liberman, & Grill-spector, 2015; Golarai, Liberman, Yoon, Grill-spector, & Bunge, 2010; Peelen, Glaser, Vuilleumier, & Eliez, 2009). It is possible that as infants gain more experience learning that faces and other social stimuli are relevant cues for learning and behavior, similar mechanisms may support functional organization for face processing over ontogenetic development through PFC feedback connectivity. This broad hypothesis also aligns with existing views suggesting that the PFC supports the specialization and organization of posterior neural regions (M. H. Johnson, 2000, 2011; M. H. Johnson, Jones, & Gliga, 2015; Thatcher, 1992; Werchan & Amso, 2017). Thus, the findings reported here raise exciting avenues for future research.

In sum, our findings show that infants can rapidly acquire top-down knowledge using PFC-dependent rule learning mechanisms, and that this top-down knowledge subsequently influences infants' feature-based attention in a novel visual context. This initial demonstration of top-down knowledge influencing visual processing in infants may help infants flexibly select features from the cluttered visual world that support adaptive behavior and guide learning in new contexts. These findings provide new mechanistic insights into how the infant brain learns to efficiently direct attention to information that is most relevant for learning and behavior over ontogenetic development.

Chapter 6: Conclusions

In this dissertation, I proposed a novel framework for PFC functional development inspired by ecological approaches to examining developmental change. Rather than having a protracted developmental course, I argue that the PFC continuously adapts its computations to accommodate the demands present in the changing ecological niche of the growing child. As such, the primary aim of this dissertation was to test a key prediction of this model: that the PFC is adapted to support *learning* during infancy through its role in the formation of flexible rule structures. I examined this prediction across four empirical studies designed to test whether the infant PFC contributes to hierarchical rule learning, a mechanism that supports learning and generalization.

In my first study, I examined whether 8-month-old infants spontaneously organize visual and auditory inputs into hierarchical rule structures that support learning and generalization across contexts (Werchan et al., 2015). I exploited a simple visual paradigm and used eye tracking to show that infants used the shape of an object as a higher-order context to organize simpler color-location rules, which they then generalized to novel shape contexts. In a second experiment, I used this same hierarchical structure to show that this learning mechanism is domain general and helps infants learn object-label mappings across different speaker contexts. Thus, these experiments provide behavioral evidence that the PFC appears to be adapted to support learning demands relevant to infants.

In my second study, I used functional near-infrared spectroscopy to examine the role of the PFC in hierarchical rule learning and generalization in infants (Werchan, Collins, Frank, & Amso, 2016). Prior computational and neuroscience work indicate that this type of learning is governed by frontrostriatal reinforcement learning mechanisms

(Collins & Frank, 2013). Using an adapted version of the tasks from my prior study, I applied fNIRS recording over PFC and behavioral eye blink rates, an exploratory measure of striatal dopamine activity (Colzato et al., 2009; Karson, 1983; Taylor et al., 1999; Dreisbach et al., 2005; Müller et al., 2007), to understand whether these same mechanisms operate in infancy. I found that infants had increased dorsolateral PFC activity and increased eye blink rate when they were updating rules into working memory. Critically, the conjunction of infants' dorsolateral PFC activity and eye blink rate during learning predicted infants' generalization ability during the subsequent inference task. These findings indicate that hierarchical rule learning mechanisms involve PFC and frontostriatal circuitry in infants as found previously in adults (e.g., Collins et al., 2014).

In my third study, I examined the flexibility of this framework by considering whether a learning account of PFC function can explain infants' errors on canonical measures of executive functions. Traditionally, PFC development and executive functions in infancy have primarily been studied using the A-Not-B error. Thus, I tested whether hierarchical rule learning mechanisms can explain the classic A-Not-B error in 9-month-old infants (Werchan & Amso, in review). I hypothesized that infants might use a higher-order context during the A-not-B task, in this case the experimenter, to organize separate working memory representations for toy-location-action rules. To test this hypothesis, infants participated in a modified A-not-B task where the experimenter changed when the toy's hiding location changed. I found that infants did not perseverate when a change in the hiding location was paired with a change in the experimenter. Critically, infants performed better on *switch* B trials that were *consistent* with the hierarchical structure, relative to *repeat* A trials that were *inconsistent* with it, even as the switch B trials required infants to inhibit a prepotent response.

In the fourth and final study, I examined the reciprocal impact of these rule learning mechanisms on influencing what infants subsequently attend to in a novel visual context. This work was designed to test an additional prediction from my ecological model: that the PFC modulates the functional response properties of posterior neural regions to adapt for processing information relevant to the developing child in their environment. I tested infants on a rule learning and generalization task, which was designed such that infants could use the visual features of shape or color as higher-order contexts to organize visual inputs into abstract rules for learning. Critically, I measured the change in infants' visual attention biases to color and shape before and after the rule learning task. In a first experiment, I found that infants structured simple visual inputs into generalizable rules, which then biased attention towards behaviorally-relevant visual features. In a second experiment, I used fNIRS to examine the neural correlates of these processes, and found that increased functional connectivity between the PFC and visual cortex was related to the efficacy of rule learning. These findings support the prediction that the PFC modulates the response properties of posterior cortical regions to support processing of behaviorally relevant information. Moreover, they also provide new insights into how the infant brain learns to flexibly select features from the cluttered visual world that support adaptive behavior.

These studies provide support for the prediction that the PFC is adapted to support learning during infancy, arguably the time in life when the learning curve is steepest. Future work is needed to explore how these systems might scaffold the emergence of increasingly complex concepts and cognitive representations over development, and how this in turn shapes learning and cortical organization to adapt to the changing mind, body, and environment. For instance, the ecological model proposed in this dissertation argues that the PFC's profuse connectivity with other neural regions makes it highly adaptable, but it also makes it highly vulnerable to atypical developmental trajectories. Thus, this model predicts that deviations in species-expected environments and early disruptions in perceptual and motor abilities will change the feedforward input to the PFC, and ultimately, the top-down influence that the PFC exerts in response. Thus, one aim of future work is to examine how changes in the environment, as well as the internal capacity to interact with the environment (through changes in attention, behavioral control, motor skills, and verbal ability), has cascading effects that impacts learning and adaptation of the mind and brain to the environment more generally. A second related aim of future work is to explore how changes in reciprocal connections between the PFC and posterior brain regions impacts attention, learning and memory.

Taken together, this research demonstrates that the infant PFC is adapted to support learning through its role in helping structure the "buzzing, blooming confusion" of environmental input into abstract rule structures that support adaptive behavior across time and contexts. This research also adds a fundamental learning mechanism to what is known about the neurocognitive toolbox available to infants to help make sense of the cluttered and rapidly changing multisensory world. Importantly, these findings also have broader impact for considering how we have typically studied the PFC and infant behavior more generally. As scientists, we often impose our own constructs and cognitive biases on infants and young children, and conduct studies that are inherently designed to look for failures and limitations (Karmiloff-Smith, 1995). Yet, this research sheds light on the importance of considering infants and young children as unique organisms with different goals for learning and behavior to have a fuller understanding of the dynamic mechanisms that contribute to complexity in thought and action as children grow and develop.

Bibliography

Aasted, C. M., Cooper, R. J., Petkov, M. P., Boas, D. A., Aasted, C. M., Yücel, M. A., … Tsuzuki, D. (2015). Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. *Neurophotonics*, 2(2). https://doi.org/10.1117/1.NPh.2.2.020801

Ahmed, S. P., Bittencourt-Hewitt, A., & Sebastian, C. L. (2015). Neurocognitive bases of emotion regulation development in adolescence. *Developmental Cognitive Neuroscience*, 15, 11–25. https://doi.org/10.1016/j.dcn.2015.07.006

Althaus, N., & Mareschal, D. (2012). Using saliency maps to separate competing processes in infant visual cognition. Child Development, 83(4), 1122–1128.

Amso, D., & Johnson, S. P. (2006). Learning by Selection: Visual Search and Object Perception in Young Infants. *Developmental Psychology*, 42(6), 1236–1245. https://doi.org/10.1037/0012-1649.42.6.1236

Amso, D., & Johnson, S. P. (2008). Development of Visual Selection in 3- to 9-Month-Olds: Evidence From Saccades to Previously Ignored Locations. *Infancy*, 13(6), 675–686. https://doi.org/10.1080/15250000802459060

Amso, D., & Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, 16(10), 606–619. https://doi.org/10.1038/nrn4025

Amso, D., Haas, S., & Markant, J. (2014). An Eye Tracking Investigation of Developmental Change in Bottom-up Attention Orienting to Faces in Cluttered Natural Scenes. *PLoS ONE*, 9(1), 1–7. https://doi.org/10.1371/journal.pone.0085701

Amso, D., Haas, S., McShane, L., & Badre, D. (2014). Working memory updating and the development of rule-guided behavior. *Cognition*, 133(1), 201–210. https://doi.org/10.1016/j.cognition.2014.06.012

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., … Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature*, 501(7465), 97–101. https://doi.org/10.1038/nature12486

Armbruster, D. J. N., Ueltzhöffer, K., Basten, U., & Fiebach, C. J. (2012). Prefrontal Cortical Mechanisms Underlying Individual Differences in Cognitive Flexibility and Stability. *Journal of Cognitive Neuroscience*, 24(12), 2385–2399. https://doi.org/10.1162/jocn_a_00286

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, 10(6), 410–422. https://doi.org/10.1038/nrn2648

Astle, D. E., Luckhoo, H., Woolrich, M., Kuo, B., Nobre, A. C., & Scerif, G. (2015). The Neural Dynamics of Fronto-Parietal Networks in Childhood Revealed using Magnetoencephalography. Cerebral Cortex, 25, 3868–3876. https://doi.org/10.1093/cercor/bhu271

Aylward, E. H., Park, J. E., Field, K. M., Parsons, A. C., Richards, T. L., Cramer, S. C., & Meltzoff, A. N. (2005). Brain Activation during Face Perception: Evidence of a Developmental Change. Journal of Cognitive Neuroscience, 17(2), 308–319.

Bacher, L. F., & Smotherman, W. P. (2004). Spontaneous eye blinking in human infants: a review. *Developmental psychobiology*, *44*(2), 95-102.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5), 193–200. https://doi.org/10.1016/j.tics.2008.02.004

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 19, 2082–2099. https://doi.org/10.1162/jocn.2007.19.12.2082

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, 10(9), 659–669. https://doi.org/10.1038/nrn2667

Badre, D., & Frank, M. J. (2012). Mechanisms of Hierarchical Reinforcement Learning in Cortico--Striatal Circuits 2: Evidence from fMRI. *Cerebral Cortex*, 22, 527–536. https://doi.org/10.1093/cercor/bhr117

Badre, D., & Wagner, A. D. (2004). Selection, Integration, and Conflict Monitoring. *Neuron*, 41(3), 473–487. https://doi.org/10.1016/S0896-6273(03)00851-1

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron*, 66(2), 315–326. https://doi.org/10.1016/j.neuron.2010.03.025

Baillargeon, R. (1987). Object Permanence in 3 1 / 2- and 4 1 / 2-Month-Old Infants. *Developmental Psychology*, 23(5), 655–664.

Baird, A., Kagana, J., Gaudetteb, T., Walza, K. A., Hershlaga, N., & Boas, D. A. (2002). Frontal Lobe Activation during Object Permanence: Data from Near-Infrared Spectroscopy. *NeuroImage*, 16(4), 1120–1126. https://doi.org/10.1006/nimg.2002.1170

Baluch, F., & Itti, L. (2011). Mechanisms of top-down attention. Trends in

Neurosciences, 34(4), 210–224. https://doi.org/10.1016/j.tins.2011.02.003

Barbas, H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. Brain Research Bulletin, 52(5), 319–330.

Barbas, H., & García-Cabezas, M. Á. (2016). How the prefrontal executive got its stripes. *Current Opinion in Neurobiology*, 40, 125–134. https://doi.org/10.1016/j.conb.2016.07.003

Barceló, F., Suwazono, S., & Knight, R. T. (2000). Prefrontal modulation of visual processing in humans. Nature Neuroscience, 3(4), 399–403. https://doi.org/10.1038/73975

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia*, 35(10), 1373–1380.

Barnes, J. J., Woolrich, M. W., Baker, K., Colclough, G. L., & Astle, D. E. (2016). Electrophysiological measures of resting state functional connectivity and their relationship with working memory capacity in childhood. Developmental Science, 19(1), 19–31. https://doi.org/10.1111/desc.12297

Barres, B. A., & Raff, M. C. (1993). Proliferation of oligodendrocyte precursor cells depends on electrical activity in axons. *Nature*, 361(6409), 258–260. https://doi.org/10.1038/361258a0

Baruth, J. M., Casanova, M. F., Sears, L., & Sokhadze, E. (2010). Early-stage visual processing abnormalities in high-functioning autism spectrum disorder (ASD). *Translational Neuroscience*, 1(2), 177–187. https://doi.org/10.2478/v10134-010-0024-9

Bell, M. A., & Fox, N. A. (1992). The relations between frontal brain electrical activity and cognitive development during infancy. *Child Development*, 63(5), 1142–63. https://doi.org/10.3758/CABN.1.2.137

Blandini F, Nappi G, Tassorelli C, Martignoni E (2000) Functional changes of the basal ganglia circuitry in Parkinson's disease. *Progress in Neurobiology* 62:63-88.

Blin, O., Masson, G., Azulay, J. P., Fondarai, J., & Serratrice, G. (1990). Apomorphine-induced blinking and yawning in healthy volunteers. *British journal of clinical pharmacology*, *30*(5), 769-773.

Bodfish, J. W., Powell, S. B., Golden, R. N., & Lewis, M. H. (1995). Blink rate as an index of dopamine function in adults with mental retardation and repetitive behavior disorders. *American Journal of Mental Retardation,* 99:335-344.

Bortfeld, H., Wruck, E., & Boas, D. A. (2007). Assessing infants ' cortical response to speech using near-infrared spectroscopy. *NeuroImage*, 34, 407–415. https://doi.org/10.1016/j.neuroimage.2006.08.010

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in cognitive sciences*, *12*(5), 201-208.

Bourgeois, J. P., Goldman-Rakic, P. S., & Rakic, P. (1994). Synaptogenesis in the prefrontal cortex of rhesus monkeys. *Cerebral Cortex*, 4(1), 78–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8180493

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In Control of cognitive processes: Attention and performance (pp. 713–737).

Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences*, 106(18), 7351–7356. https://doi.org/10.1073/pnas.0808187106

Bremner, J. G. (1978). Spatial errors made by infants: Inadequate spatial cues or evidence of egocentrism? *British Journal of Psychology*, 69(1), 77–84.

Brito, N. H., & Noble, K. G. (2014). Socioeconomic status and structural brain development. *Frontiers in Neuroscience*, 8(SEP), 1–12. https://doi.org/10.3389/fnins.2014.00276

Bunge SA (2004) How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cognitive Affective Behavioral Neuroscience*, 4:564-579.

Bunge, S. A. (2003). Neural Circuits Subserving the Retrieval and Maintenance of Abstract Rules. *Journal of Neurophysiology*, 90(5), 3419–3428. https://doi.org/10.1152/jn.00910.2002

Bunge, S. A., & Zelazo, P. D. (2006). A Brain-Based Account of the Development of Rule Use in Childhood. *Current Directions in Psychological Science*, 15(3), 118–121. Retrieved from http://www.jstor.org/stable/20183092

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315(5820), 1860–1862. https://doi.org/10.1126/science.1138071

Cahalane, D. J., Charvet, C. J., & Finlay, B. L. (2012). Systematic, balancing gradients in neuron density and number across the primate isocortex. *Frontiers in Neuroanatomy*, 6(July), 28. https://doi.org/10.3389/fnana.2012.00028

Callaghan, B. L., & Richardson, R. (2011). Maternal separation results in early emergence of adult-like fear and extinction learning in infant rats. *Behavioral Neuroscience*, 125(1), 20–28. https://doi.org/10.1037/a0022008

Callaghan, B. L., & Richardson, R. (2012). The effect of adverse rearing environments on persistent memories in young rats: removing the brakes on infant fear memories. *Translational Psychiatry*, 2(7), e138. https://doi.org/10.1038/tp.2012.65

Callaghan, B. L., & Richardson, R. (2014). Early emergence of adult-like fear renewal in the developing rat after chronic corticosterone treatment of the dam or the pups. *Behavioral Neuroscience*, 128(5), 594–602. https://doi.org/10.1037/bne0000009

Callaghan, B. L., & Tottenham, N. (2016). The neuro-environmental loop of plasticity: A cross-species analysis of parental effects on emotion circuitry development following typical and adverse caregiving. *Neuropsychopharmacology*, 41(1), 163–76. https://doi.org/10.1038/npp.2015.204

Callaghan, B. L., Sullivan, R. M., Howell, B., & Tottenham, N. (2014). The international society for developmental psychobiology Sackler symposium: Early adversity and the maturation of emotion circuits-A cross-species analysis. *Developmental Psychobiology*, 56(8), 1635–1650. https://doi.org/10.1002/dev.21260

Canfield, R. L., & Kirkham, N. Z. (2001). Infant cortical development and the prospective control of saccadic eye movements. *Infancy*, 2(2), 197–211. https://doi.org/10.1207/S15327078IN0202_5

Canfield, R.L. & Haith, M.M. (1991). Active expectations in 2- and 3-month-old infants: complex event sequences. *Developmental Psychology*, 27, 198-208.

Capotosto, P., Babiloni, C., Romani, G. L., Corbetta, M., San, C., & Cassino, R. (2009). Frontoparietal Cortex Controls Spatial Attention through Modulation of

Anticipatory Alpha Rhythms. The Journal of Neuroscience, 29(18), 5863–5872. https://doi.org/10.1523/JNEUROSCI.0539-09.2009

Caroni, P., Donato, F., & Muller, D. (2012). Structural plasticity upon learning: regulation and functions. *Nature Reviews Neuroscience*, 13(7), 478–490. https://doi.org/10.1038/nrn3258

Carrasco, M. (2011). Visual attention: the past 25 years. Vision Research, 51(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Cavanagh J. F., Wiecki T. V., Kochar A., Frank M. J. (2014) Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology-General*, 143:1476-1488.

Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. *Current Opinion in Behavioral Sciences*, 1, 23–31. https://doi.org/10.1016/j.cobeha.2014.08.001

Chatham, C. H., Frank, M. J., & Badre, D. (2014). Corticostriatal Output Gating during Selection from Working Memory. *Neuron*, 81(4), 930–942. https://doi.org/10.1016/j.neuron.2014.01.002

Cheng, C., Kaldy, Z., & Blaser, E. (2019). Focused attention predicts visual working memory performance in 13- month-old infants: A pupillometric study. Developmental Cognitive Neuroscience, 36. https://doi.org/10.1016/j.dcn.2019.100616

Cheng, W., Palaniyappan, L., Li, M., Kendrick, K. M., Zhang, J., Luo, Q., … Feng, J. (2015). Voxel-based, brain-wide association study of aberrant functional connectivity in schizophrenia implicates thalamocortical circuitry. *Npj*

Schizophrenia, 1(1), 15016. https://doi.org/10.1038/npjschz.2015.16

Chiew, K. S., & Braver, T. S. (2017). *Context processing and control in the human brain: from gating models to dual mechanisms*. In T. Egner (Ed.), The Wiley Handbook of Cognitive Control (pp. 143–166). Chichester, West Sussex: John Wiley & Sons.

Chomsky, N. (1988). *Aspects of the Theory of Syntax* (Vol. 11). MIT press.

Chrysikou, E. G., Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2011). The other side of cognitive control: Can a lack of cognitive control benefit language and cognition? *Topics in Cognitive Science*, 3(2), 253–256. https://doi.org/10.1111/j.1756-8765.2011.01137.x

Chrysikou, E. G., Weber, M. J., & Thompson-Schill, S. L. (2014). A matched filter hypothesis for cognitive control. *Neuropsychologia*, 62, 341–355. https://doi.org/10.1016/j.neuropsychologia.2013.10.021

Chugani H. T. (1996) *Neuroimaging of developmental nonlinearity and developmental pathologies*. In: Developmental neuroimaging: mapping the development of brain and behavior (Thatcher RW, Lyon GR, Rumsey J, Krasnegor N, eds.), pp. 187–95. San Diego, CA: Academic Press.

Chugani, H. T., & Phelps, M. E. (1986). Maturational changes in cerebral function in infants determined by 18FDG positron emission tomography. *Science*, 231(4740), 840–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3945811

Cocchi, L., Halford, G. S., Zalesky, A., Harding, I. H., Ramm, B. J., Cutmore, T., … Mattingley, J. B. (2014). Complexity in relational processing predicts changes in functional brain network dynamics. *Cerebral Cortex*, 24(9), 2283–2296.

https://doi.org/10.1093/cercor/bht075

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Current Opinion in Neurobiology*, 12(2), 223–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12015241

Collins A. G. E., & Frank M. J. (2016) Neural signature of latent structure learning predicts hierarchical clustering and transfer of rule sets. *Cognition* 152: 160-169.

Collins, A. G. E., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, 120(1), 190–229. https://doi.org/10.1037/a0030852

Collins, A. G. E., Cavanagh, J. F., & Frank, M. J. (2014). Human EEG Uncovers Latent Generalizable Rule Structure during Learning. *Journal of Neuroscience*, 34(13), 4677–4685. https://doi.org/10.1523/JNEUROSCI.3900-13.2014

Colzato L. S., Slagter H. A., van den Wildenberg W. P., & Hommel B. (2009) Closing one's eyes to reality: Evidence for a dopaminergic basis of psychoticism from spontaneous eye blink rates. *Personality and Individual Differences*, 46:377-380.

Colzato L. S., van den Wildenberg W. P., & Hommel B. (2008) Reduced spontaneous eye blink rates in recreational cocaine users: evidence for dopaminergic hypoactivity. *PLoS One,* 3:e3461.

Conel, J. L. (1939). *The postnatal development of the human cerebral cortex*. Vol. 1. The cortex of the newborn.

Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, 17(7), 438–449. https://doi.org/10.1038/nrn.2016.43

Corbetta, M., & Shulman, G. L. (2002). Control of Goal-Directed and Stimulus-Driven Attention in the Brain. Nature Reviews Neuroscience, 3, 201–215. https://doi.org/10.1038/nrn755

Courchesne, E., Chisum, H. J., Townsend, J., Cowles, A., Covington, J., Egaas, B., … Press, G. A. (2000). Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology*, 216(3), 672–682. https://doi.org/10.1148/radiology.216.3.r00au37672

Crone E. A., Wendelken C., Donohue S., van Leijenhorst L., & Bunge S. A. (2006) Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences USA*, 103:9315-9320.

Crone, E. A., & Steinbeis, N. (2017). Neural perspectives on cognitive control development during childhood and adolescence. *Trends in Cognitive Sciences*, 21(3), 205–215. https://doi.org/10.1016/j.tics.2017.01.003

Csibra, G., Tucker, L. A., & Johnson, M. H. (2001). Differential frontal cortex activation before anticipatory and reactive saccades in infants. *Infancy*, 2(2), 159–174. https://doi.org/10.1207/S15327078IN0202_3

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9), 415–423. https://doi.org/10.1016/S1364-6613(03)00197-9

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, 66(1), 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Davidson, R. J., & Fox, N. A. (1982). Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants. *Science*, 218(4578), 1235–1237. https://doi.org/10.1126/science.7146906

Davidson, R. J., & Fox, N. A. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology*, 98(2), 127–131. https://doi.org/10.1037/0021-843X.98.2.127

Dawson, G., Panagiotides, H., Klinger, L. G., & Hill, D. (1992). The role of frontal lobe functioning in the development of infant self-regulatory behavior. *Brain and Cognition*, 20(1), 152–175. https://doi.org/10.1016/0278-2626(92)90066-U

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional Neuroimaging of Speech Perception in Infants. *Science*, 298(5600), 2013–2015. https://doi.org/10.1126/science.1077066

Demerens, C., Stankoff, B., Logak, M., Anglade, P., Allinquant, B., Couraud, F., … Lubetzki, C. (1996). Induction of myelination in the central nervous system by electrical activity. *Proceedings of the National Academy of Sciences of the United States of America*, 93(18), 9887–92. https://doi.org/10.1016/j.tics.2008.02.004

Deoni, S. C. L., Dean, D. C., O'Muircheartaigh, J., Dirks, H., & Jerskey, B. A. (2012). Investigating white matter development in infancy and early childhood using myelin water faction and relaxation time mapping. *NeuroImage*, 63(3), 1038–1053. https://doi.org/10.1016/j.neuroimage.2012.07.037

Deoni, S. C. L., Mercure, E., Blasi, A., Gasston, D., Thomson, A., Johnson, M., … Murphy, D. G. M. (2011). Mapping Infant Brain Myelination with Magnetic Resonance Imaging. *Journal of Neuroscience*, 31(2), 784–791. https://doi.org/10.1523/JNEUROSCI.2106-10.2011

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. Annual Review of Neuroscience, 18, 193–222.

Deuschel G, Goddemeier C (1998) Spontaneous and reflex activity of facial muscles in dystonia, Parkinson's disease and in normal subjects. *Journal of Neurology, Neurosurgery, and Psychiatry*, 64:320–324.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development*, 56(4), 868–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4042750

Diamond, A. (2001). Looking closely at infants' performance and experimental procedures in the A-not-B task. *Behavioral and Brain Sciences*, 24(1), 38–41. https://doi.org/10.1017/S0140525X01253916

Diamond, A. (2002). *Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry*. In Principles of Frontal Lobe Function (pp. 466–503). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195134971.003.0029

Diamond, A., & Goldman-Rakic, P. S. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research*, 74(1), 24–40. https://doi.org/10.1007/BF00248277

Diamond, A., & Kirkham, N. Z. (2005). Not Quite as Grown-Up as We Like to Think. Psychological Science, 16(4), 291–297.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333(6045), 959–964. https://doi.org/10.1126/science.1204529.Interventions

Diaz, A., & Bell, M. A. (2012). Frontal EEG asymmetry and fear reactivity in different contexts at 10 months. *Developmental Psychobiology*, 54(5), 536–545. https://doi.org/10.1002/dev.20612

Dombrowski, S. M., Hilgetag, C. C., & Barbas, H. (2001). Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. *Cerebral Cortex*, 11(10), 975–88.

Domenech, P., & Koechlin, E. (2015). Executive control and decision-making in the prefrontal cortex. *Current Opinion in Behavioral Sciences*, 1(5796), 101–106. https://doi.org/10.1016/j.cobeha.2014.10.007

Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science*, 344(6191), 1481–1486. https://doi.org/10.1126/science.1252254

Doria, V., Beckmann, C. F., Arichi, T., Merchant, N., Groppo, M., Turkheimer, F. E., … Edwards, D. A. (2010). Emergence of resting state networks in the preterm human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46), 20015–20020. https://doi.org/10.1073/pnas.1007921107

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. a T., … Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104(26), 11073–8. https://doi.org/10.1073/pnas.0704320104

Dreisbach, G., Müller, J., Goschke, T., Strobel, A., Schulze, K., Lesch, K. P., & Brocke, B. (2005). Dopamine and cognitive control: the influence of spontaneous eyeblink rate and dopamine gene polymorphisms on perseveration and distractibility. *Behavioral neuroscience*, *119*(2), 483-490.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, 2(November), 820–829. https://doi.org/10.1038/35097575

Duncan, J., & Miller, E. K. (2002). *Cognitive Focus through Adaptive Neural Coding in the Primate Prefrontal Cortex*. In Principles of Frontal Lobe Function (Vol. 63, pp. 278–291). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195134971.003.0018

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10), 475–483. https://doi.org/10.1016/S0166-2236(00)01633-7

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *Journal of Neurophysiology*, 83(3), 1733–50. https://doi.org/10.1152/jn.00634.2002

Elman, J., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*.

Cambridge, MA: MIT Press/Bradford Books.

Elsworth, J. D., Lawrence, M. S., Roth, R. H., Taylor, J. R., Mailman, R. B., Nichols, D. E., ... Redmond, D. E. (1991) D1 and D2 dopamine receptors independently regulate spontaneous blink rate in the vervet monkey. *Journal of Pharmacology and Experimental Therapeutics*, 259:595-600.

Eluvathingal, T. J. (2006). Abnormal Brain Connectivity in Children After Early Severe Socioemotional Deprivation: A Diffusion Tensor Imaging Study. *Pediatrics*, 117(6), 2093–2100. https://doi.org/10.1542/peds.2005-1727

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015) Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences USA*, 112:9585-9590.

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. Proceedings of the National Academy of Sciences, 112(31), 9585–9590. https://doi.org/10.1073/pnas.1510343112

Erb, C. D., Moher, J., Song, J., & Sobel, D. M. (2017). Cognitive control in action: Tracking the dynamics of rule switching in 5- to 8-year-olds and adults. Cognition, 164, 163–173. https://doi.org/10.1016/j.cognition.2017.03.023

Fair, D. A., Dosenbach, N. U. F., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., … Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences of the United States of America*, 104(33), 13507–12. https://doi.org/10.1073/pnas.0705843104

Fallon, S. J., Williams-Gray, C. H., Barker, R. A., Owen, A. M., & Hampshire, A. (2013). Prefrontal dopamine levels determine the balance between cognitive stability and flexibility. *Cerebral Cortex*, 23(2), 361–369. https://doi.org/10.1093/cercor/bhs025

Farah, M. J., Betancourt, L., Shera, D. M., Savage, J. H., Giannetta, J. M., Brodsky, N. L., … Hurt, H. (2008). Environmental stimulation, parental nurturance and cognitive development in humans. *Developmental Science*, 11(5), 793–801. https://doi.org/10.1111/j.1467-7687.2008.00688.x

Farah, M. J., Shera, D. M., Savage, J. H., Betancourt, L., Giannetta, J. M., Brodsky, N. L., … Hurt, H. (2006). Childhood poverty: Specific associations with neurocognitive development. *Brain Research*, 1110(1), 166–174. https://doi.org/10.1016/j.brainres.2006.06.072

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 110(41), 16616–16621. https://doi.org/10.1073/pnas.1315235110

Fellows, L. K., & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal learning paradigm. *Brain*, 126(8), 1830–1837. https://doi.org/10.1093/brain/awg180

Field, T., Fox, N. A., Pickens, J., & Nawrocki, T. (1995). Relative right frontal EEG activation in 3- to 6-month-old infants of “depressed” mothers. *Developmental Psychology*, 31(3), 358–363. https://doi.org/10.1037/0012-1649.31.3.358

Figueras, B., Edwards, L., & Langdon, D. (2008). Executive function and language in deaf children. *Journal of Deaf Studies and Deaf Education*, 13(3), 362–377.

https://doi.org/10.1093/deafed/enm067

Finlay, B. L., & Uchiyama, R. (2015). Developmental mechanisms channeling cortical evolution. *Trends in Neurosciences*, 38(2), 69–76. https://doi.org/10.1016/j.tins.2014.11.004

Finlay, B. L., & Uchiyama, R. (2017). *The timing of brain maturation, early experience, and the human social niche*. In J. Kaas (Ed.), Evolution of Nervous Systems (pp. 123–148). Cambridge, MA: Elsevier. https://doi.org/10.1016/B978-0-12-804042-3.00134-2

Fox, N. A. (1991). If it's not left, it's right. *American Psychologist*, 46(8), 863–872.

Fox, N. A. (1994). Dynamic cerebral processes underlying emotion regulation. *Monographs of the Society for Research in Child Development*, 59(2–3), 152–66. https://doi.org/10.1162/jocn_a_00286

Fox, N. A., & Bell, M. A. (1990). Electrophysiological Indices of Frontal Lobe Development. *Annals of the New York Academy of Sciences*, 608(1 The Developme), 677–704. https://doi.org/10.1111/j.1749-6632.1990.tb48914.x

Fox, N. A., & Davidson, R. J. (1987). Electroencephalogram asymmetry in response to the approach of a stranger and maternal separation in 10-month-old infants. *Developmental Psychology*, 23(2), 233–240. https://doi.org/10.1037/0012-1649.23.2.233

Frank, M. C., Slemmer, J. A., Marcus, G. F., & Johnson, S. P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Developmental science*, *12*(4), 504-509.

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, 110(2), 160–170. https://doi.org/10.1016/j.cognition.2008.11.010.Development

Frank, M. J., & Badre, D. (2012). Mechanisms of Hierarchical Reinforcement Learning in Corticostriatal Circuits 1: Computational Analysis. *Cerebral Cortex*, 22, 509–526. https://doi.org/10.1093/cercor/bhr114

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(2), 137-160.

Frankenhuis, W. E., & Del Giudice, M. (2012). When do adaptive developmental mechanisms yield maladaptive outcomes? *Developmental Psychology*, 48(3), 628–642. https://doi.org/10.1037/a0025629

Frankenhuis, W. E., & Panchanathan, K. (2011). Balancing sampling and specialization: an adaptationist model of incremental development. *Proceedings of the Royal Society: Biological Sciences*, 278(1724), 3558–3565. https://doi.org/10.1098/rspb.2011.0055

Fransson, P., Åden, U., Blennow, M., & Lagercrantz, H. (2011). The functional architecture of the infant brain as revealed by resting-state fMRI. *Cerebral Cortex*, 21(1), 145–154. https://doi.org/10.1093/cercor/bhq071

Fransson, P., Skiöld, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., & Aden, U. (2007). Resting-state networks in the infant brain. *Proceedings of the National Academy of Sciences of the United States of America*, 104(39), 15531–6. https://doi.org/10.1073/pnas.0704380104

Freed, W. J., Kleinman, J. E., Karson, C. N., Potkin, S. G., Murphy, D. L., & Wyatt, R. J. (1980) Eye-blink rates and platelet monoamine oxidase activity in chronic schizophrenic patients. *Biological Psychiatry*, 15:329-332.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, 137(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Fukuda, K., Stern, J. A., Brown, T. B., & Russo, M. B. (2005) Cognition, blinks, eye-movements, and pupillary movements during performance of a running memory task. *Aviation, Space, and Environmental Medicine*, 76:C75-C85.

Fuster, J. M. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology*, 31(3/5), 373–385. https://doi.org/10.1023/A:1024190429920

Fuster, J. M. (2008). *The Prefrontal Cortex* (4th Edition). London: Academic Press.

Gao, W., Gilmore, J. H., Giovanello, K. S., Smith, J. K., Shen, D., Zhu, H., & Lin, W. (2011). Temporal and spatial evolution of brain network topology during the first two years of life. *PLoS ONE*, 6(9). https://doi.org/10.1371/journal.pone.0025278

Gazzaley, A., Rissman, J., Cooney, J., Seibert, T., Clapp, W., & Esposito, M. D. (2007). Functional Interactions between Prefrontal and Visual Association Cortex Contribute to Top-Down Modulation of Visual Processing. Cerebral Cortex, 17, 125–135. https://doi.org/10.1093/cercor/bhm113

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, *16*(3), 486-490.

Gee, D. G., Gabard-Durnam, L. J., Flannery, J., Goff, B., Humphreys, K. L., Telzer, E. H., … Tottenham, N. (2013). Early developmental emergence of human amygdala-prefrontal connectivity after maternal deprivation. *Proceedings of the National Academy of Sciences*, 110(39), 15638–15643. https://doi.org/10.1073/pnas.1307893110

Gee, D. G., Humphreys, K. L., Flannery, J., Goff, B., Telzer, E. H., Shapiro, M., … Tottenham, N. (2013). A Developmental Shift from Positive to Negative Connectivity in Human Amygdala-Prefrontal Circuitry. *Journal of Neuroscience*, 33(10), 4584–4593. https://doi.org/10.1523/JNEUROSCI.3446-12.2013

Genesee, F., & Nicoladis, E. (2007). Bilingual first language acquisition. *Handbook of language development*, 324-342.

Ghisleni, C., Bollmann, S., Poil, S.-S., Brandeis, D., Martin, E., Michels, L., … Klaver, P. (2015). Subcortical Glutamate Mediates the Reduction of Short-Range Functional Connectivity with Age in a Developmental Cohort. *Journal of Neuroscience*, 35(22), 8433–8441. https://doi.org/10.1523/JNEUROSCI.4375-14.2015

Gibson, E. J., Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. USA: Oxford University Press.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., … Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, 2(10), 861–863. https://doi.org/10.1038/13158

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature*

Reviews Neuroscience, 14(5), 350–363. https://doi.org/10.1038/nrn3476

Gilmartin, M. R., Balderston, N. L., & Helmstetter, F. J. (2014). Prefrontal cortical regulation of fear learning. *Trends in Neurosciences*, 37(8), 445–464. https://doi.org/10.1016/j.tins.2014.05.004

Goel V, Dolan RJ (2000) Anatomical segregation of component processes in an inductive inference task. *Journal of Cognitive Neuroscience*, 12:110-119.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, a C., … Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21), 8174–9. https://doi.org/10.1073/pnas.0402680101

Golarai, G., Liberman, A., & Grill-spector, K. (2015). Experience Shapes the Development of Neural Substrates of Face Processing in Human Ventral Temporal Cortex. Cerebral Cortex, 1–16. https://doi.org/10.1093/cercor/bhv314

Golarai, G., Liberman, A., Yoon, J. M. D., Grill-spector, K., & Bunge, S. A. (2010). Differential development of the ventral visual cortex extends through adolescence. Frontiers in Human Neuroscience, 3, 1–19. https://doi.org/10.3389/neuro.09.080.2009

Goldman-Rakic, P. S. (1987). Development of cortical circuitry and cognitive function. *Child Development*, 58(3), 601–22. Retrieved from http://journals.sagepub.com/doi/10.1111/j.0963-7214.2006.00419.x

Goldman-Rakic, P. S., Bourgeois, J. P., & Rakic, P. (1997). *Synaptic substrate of cognitive development: Lifespan analysis of synaptogenesis in the prefrontal* *cortex of the nonhuman primate*. In P. S. Krasnegor, N.A., Lyon, G.R., Goldman-Rakic (Ed.), Development of the prefrontal cortex: Evolution, neurobiology, and behavior (pp. 27–48). Baltimore, MD: Paul H. Brooks.

Gooch, D., Thompson, P. M., Nash, H. M., Snowling, M. J., & Hulme, C. (2016). The development of executive function and language skills in the early school years. *Journal of Child Psychology and Psychiatry*, 57(2), 180–187. https://doi.org/10.1111/jcpp.12458

Graham, S. A., & Diesendruck, G. (2010). Fifteen-month-old infants attend to shape over other perceptual properties in an induction task. *Cognitive Development*, *25*(2), 111-123.

Grossmann, T., & Johnson, M. H. (2010). Selective prefrontal cortex responses to joint attention in early infancy. *Biology Letters*, 6(4), 540–543. https://doi.org/10.1098/rsbl.2009.1069

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C. E., & Csibra, G. (2008). Early cortical specialization for face-to-face communication in human infants. *Proceedings of the Royal Society: Biological Sciences*, 275(1653), 2803–2811. https://doi.org/10.1098/rspb.2008.0986

Grossmann, T., Lloyd-Fox, S., & Johnson, M. H. (2013). Brain responses reveal young infants' sensitivity to when a social partner follows their gaze. *Developmental Cognitive Neuroscience*, 6, 155–161. https://doi.org/10.1016/j.dcn.2013.09.004

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, 13(2), 65–73. https://doi.org/10.1016/j.tics.2008.11.003

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, 11(9), 651–659. https://doi.org/10.1038/nrn2897

Hackman, D. A., Gallop, R., Evans, G. W., & Farah, M. J. (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science*, 18(5), 686–702. https://doi.org/10.1111/desc.12246

Hadland, K. A. (2003). The Anterior Cingulate and Reward-Guided Selection of Actions. *Journal of Neurophysiology*, 89(2), 1161–1164. https://doi.org/10.1152/jn.00634.2002

Hair, N. L., Hanson, J. L., Wolfe, B. L., & Pollak, S. D. (2015). Association of child poverty, brain development, and academic achievement. *JAMA Pediatrics*, 53706(9), 1–8. https://doi.org/10.1001/jamapediatrics.2015.1475

Hanson, J. L., Adluru, N., Chung, M. K., Alexander, A. L., Davidson, R. J., & Pollak, S. D. (2013). Early neglect is associated with alterations in white matter integrity and cognitive functioning. *Child Development*, 84(5), 1566–1578. https://doi.org/10.1111/cdev.12069

Hanson, J. L., Knodt, A. R., Brigidi, B. D., & Hariri, A. R. (2015). Lower structural integrity of the uncinate fasciculus is associated with a history of child maltreatment and future psychological vulnerability to stress. *Development and Psychopathology*, 27(4 Pt 2), 1611–9. https://doi.org/10.1017/S0954579415000978

Harden, K. P., Turkheimer, E., & Loehlin, J. C. (2007). Genotype by Environment Interaction in Adolescents' Cognitive Aptitude. *Behavior Genetics*, 37(2), 273–

283. https://doi.org/10.1007/s10519-006-9113-4

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry*, 53(1), 37–45. https://doi.org/10.1111/j.1469-7610.2011.02430.x

Hensch, T. K., Fagiolini, M., Mataga, N., Stryker, M. P., Baekkeskov, S., & Kash, S. F. (1998). Local GABA Circuit Control of Experience-Dependent Plasticity in Developing Visual Cortex. *Science*, 282(5393), 1504–1508. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2851625

Hill, J., Inder, T., Neil, J., Dierker, D., Harwell, J., & Van Essen, D. C. (2010). Similar patterns of cortical expansion during human development and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29), 13135–40. https://doi.org/10.1073/pnas.1001229107

Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T., & Hübener, M. (2009). Experience leaves a lasting structural trace in cortical circuits. *Nature*, 457(January), 313–317. https://doi.org/10.1038/nature07487

Holtmaat, A., & Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience*, 10(10), 759–759. https://doi.org/10.1038/nrn2721

Holtmaat, A., Wilbrecht, L., Knott, G. W., Welker, E., & Svoboda, K. (2006). Experience-dependent and cell-type-specific spine growth in the neocortex. *Nature*, 441(7096), 979–983. https://doi.org/10.1038/nature04783

Homae, F., Watanabe, H., Nakano, T., & Taga, G. (2011). Large-scale brain networks underlying language acquisition in early infancy. *Frontiers in Psychology*, 2, 1–

14. https://doi.org/10.3389/fpsyg.2011.00093

Homae, F., Watanabe, H., Otobe, T., Nakano, T., Go, T., Konishi, Y., & Taga, G. (2010). Development of Global Cortical Networks in Early Infancy. *The Journal of Neuroscience*, 30(14), 4877–4882. https://doi.org/10.1523/JNEUROSCI.5618-09.2010

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., & Polkey, C. E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, 16(3), 463–478. https://doi.org/10.1162/089892904322926791

Humphreys, K. L., Lee, S. S., Telzer, E. H., Gabard-Durnam, L. J., Goff, B., Flannery, J., & Tottenham, N. (2015). Exploration-exploitation strategy is dependent on early experience. *Developmental Psychobiology*, 57(3), 313–321. https://doi.org/10.1002/dev.21293

Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER : a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*, 48(10), D280–D298.

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain research*, *163*(2), 195-205.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology*, 387(2), 167–178. https://doi.org/10.1002/(SICI)1096-9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z

Hwang, K., Velanova, K., & Luna, B. (2010). Strengthening of top-down frontal cognitive control networks underlying the development of inhibitory control: a functional magnetic resonance imaging effective connectivity study. *Journal of Neuroscience*, 30(46), 15535–15545. https://doi.org/10.1523/JNEUROSCI.2825-10.2010

James, K. H., Jones, S. S., Smith, L. B., & Swain, S. N. (2014). Young Children's Self-Generated Object Views and Object Recognition. *Journal of Cognition and Development*, 15(3), 393–401. https://doi.org/10.1080/15248372.2012.749481

Johnson, C. M., Peckler, H., Tai, L.-H., & Wilbrecht, L. (2016). Rule learning enhances structural plasticity of long-range axons in frontal cortex. *Nature Communications*, 7, 10785. https://doi.org/10.1038/ncomms10785

Johnson, M. H. (1990). Cortical Maturation and the Development of Visual Attention in Early Infancy. *Journal of Cognitive Neuroscience*, 2(2), 81–95.

Johnson, M. H. (2000). Functional Brain Development in Infants: Elements of an Interactive Specialization Framework. *Child Development*, 71(1), 75–81. https://doi.org/10.1111/1467-8624.00120

Johnson, M. H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, 1(1), 7–21. https://doi.org/10.1016/j.dcn.2010.07.003

Johnson, M. H., & De Haan, M. (2015). *Developmental cognitive neuroscience: An introduction*. John Wiley & Sons.

Johnson, M. H., & Vecera, S. P. (1996). Cortical differentiation and neurocognitive development: The parcellation conjecture. *Behavioural Processes*, 36(2), 195–212. https://doi.org/10.1016/0376-6357(95)00028-3

Johnson, M. H., Jones, E. J. H., & Gliga, T. (2015). Brain adaptation and alternative developmental trajectories. *Development and Psychopathology*, 27(2015), 425–442. https://doi.org/10.1017/S0954579415000073

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of Visual Orienting in Early Infancy: Contingency Learning, Anticipatory Looking, and Disengaging. *Journal of Cognitive Neuroscience*, 3(4), 335–344. https://doi.org/10.1162/jocn.1991.3.4.335

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1994). Facilitation of Saccades toward a Covertly Attended Location in Early Infancy. *Psychological Science*, 5(2), 90–93.

Johnson, S. B., Riis, J. L., & Noble, K. G. (2016). State of the Art Review: Poverty and the Developing Brain. *Pediatrics*, 137(4), e20153075–e20153075. https://doi.org/10.1542/peds.2015-3075

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy : Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, 100(18).

Jones, N. A., Field, T., & Almeida, A. (2009). Right frontal EEG asymmetry and behavioral inhibition in infants of depressed mothers. *Infant Behavior and Development*, 32(3), 298–304. https://doi.org/10.1016/j.infbeh.2009.04.004

Jones, N. A., Field, T., Fox, N. A., Lundy, B., & Davalos, M. (1997). EEG activation in 1-month-old infants of depressed mothers. *Development and Psychopathology*, 9(January), 491–505. https://doi.org/10.1017/S0954579497001260

Jutkiewicz EM, Bergman J (2004) Effects of dopamine D1 ligands on eye blinking in monkeys: efficacy, antagonism, and D1/D2 interactions. *Journal of Pharmacology and Experimental Therapeutics*, 311:1008-1015.

Karasik, L. B., Tamis-Lemonda, C. S., & Adolph, K. E. (2011). Transition from crawling to walking and infants' actions with objects and people. *Child Development*, 82(4), 1199–1209. https://doi.org/10.1111/j.1467-8624.2011.01595.x

Karmiloff-Smith, A. (1995). *Beyond modularity: A developmental perspective on cognitive science*. MIT Press.

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, 45(1), 56–63. https://doi.org/10.1037/a0014506

Karson, C. N. (1983). Spontaneous eye-blink rates and dopaminergic systems. *Brain*, *106*(3), 643-653.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of Visual Attention in the Human Cortex. *Annual Review of Neuroscience*, 23, 315–341.

Keehn, B., Wagner, J. B., Tager-flusberg, H., Nelson, C. A., Uddin, L. Q., & Dinstein, I. (2013). Functional connectivity in the first year of life in infants at-risk for autism : a preliminary near-infrared spectroscopy study. *Frontiers in Human Neuroscience*, 7, 1–10. https://doi.org/10.3389/fnhum.2013.00444

Kelly, A. M. C., & Garavan, H. (2005). Human Functional Neuroimaging of Brain Changes Associated with Practice. *Cerebral Cortex*, 15, 1089–1102. https://doi.org/10.1093/cercor/bhi005

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. S.

(2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7), 940–947. https://doi.org/nn1724 [pii]

Kharitonova, M., Martin, R. E., Gabrieli, J. D. E., & Sheridan, M. A. (2013). Cortical gray-matter thinning is associated with age-related improvements on executive function tasks. *Developmental Cognitive Neuroscience*, 6, 61–71. https://doi.org/10.1016/j.dcn.2013.07.002

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35-B42.

Kleven, M. S., & Koek, W. (1996). Differential effects of direct and indirect dopamine agonists on eye blink rate in cynomolgus monkeys. *Journal of Pharmacology and Experimental Therapeutics*, *279*(3), 1211-1219.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience*, 14(1), 1–10. https://doi.org/10.1162/089892902317205276

Knickmeyer, R. C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J. K., … Gilmore, J. H. (2008). A Structural MRI Study of Human Brain Development from Birth to 2 Years. *The Journal of Neuroscience*, 28(47), 12176–12182. https://doi.org/10.1523/JNEUROSCI.3479-08.2008

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648), 1181-1185.

Koechlin, E. (2016). Prefrontal executive function and adaptive behavior in complex environments. *Current Opinion in Neurobiology*, 37(9), 1–6. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0959438815001786

Kolb, B., Mychasiuk, R., Muhammad, A., Li, Y., Frost, D. O., & Gibb, R. (2012). Experience and the developing prefrontal cortex. *Proceedings of the National Academy of Sciences*, 109(Supplement_2), 17186–17193. https://doi.org/10.1073/pnas.1121251109

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. *Science*, 336(6077), 95–98. https://doi.org/10.1126/science.1216930

Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, J., … Fink, G. R. (2005). Development of attentional networks: An fMRI study with children and adults. *NeuroImage*, 28(2), 429–439. https://doi.org/10.1016/j.neuroimage.2005.06.065

Kuhn, L. J., Willoughby, M. T., Vernon-Feagans, L., & Blair, C. B. (2016). The contribution of children's time-specific and longitudinal expressive language skills on developmental trajectories of executive function. *Journal of Experimental Child Psychology*, 148, 20–34. https://doi.org/10.1016/j.jecp.2016.03.008

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2000). Niche construction, biological evolution, and cultural change. *The Behavioral and Brain Sciences*, 23(1), 131-46; discussion 146-75. https://doi.org/10.1016/j.cobeha.2014.10.007

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, *3*(3), 299-321.

Lawson, G. M., Duda, J. T., Avants, B. B., Wu, J., & Farah, M. J. (2013). Associations between children's socioeconomic status and prefrontal cortical thickness. *Developmental Science*, 16(5), 641–652. https://doi.org/10.1111/desc.12096

Lee, A., Ratnarajah, N., Tuan, T. A., Chen, S. A., & Qiu, A. (2015). Adaptation of Brain Functional and Structural Networks in Aging. *PLOS ONE*, 10(4), e0123462. https://doi.org/10.1371/journal.pone.0123462

Lehrman, D. S. (1953). A critique of Konrad Lorenz's theory of instinctive behavior. *Quarterly Review of Biology*, 337–363.

Lew, S. E., & Tseng, K. Y. (2014). Dopamine Modulation of GABAergic Function Enables Network Stability and Input Selectivity for Sustaining Working Memory in a Computational Model of the Prefrontal Cortex. *Neuropsychopharmacology*, 39(13), 3067–3076. https://doi.org/10.1038/npp.2014.160

Lewkowicz, D. J., Schmuckler, M. A., & Mangalindan, D. M. J. (2018). Learning of hierarchical serial patterns emerges in infancy. *Developmental Psychobiology*, (February), 1–13. https://doi.org/10.1002/dev.21614

Li, L., Men, W.-W., Chang, Y.-K., Fan, M.-X., Ji, L., & Wei, G.-X. (2014). Acute Aerobic Exercise Increases Cortical Activity during Working Memory: A Functional MRI Study in Female College Students. *PLoS ONE*, 9(6), e99222. https://doi.org/10.1371/journal.pone.0099222

Lloyd-Fox, S., Blasi, A., Everdell, N., Elwell, C. E., & Johnson, M. H. (2011). Selective Cortical Mapping of Biological Motion Processing in Young Infants. *Journal of Cognitive Neuroscience*, 23(9), 2521–2532. https://doi.org/10.1162/jocn.2010.21598

Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., & Johnson, M. H. (2009) Social perception in infancy: a near infrared spectroscopy study. *Child Development*, 80:986-999.

Lövdén, M., Bodammer, N. C., Kühn, S., Kaufmann, J., Schütze, H., Tempelmann, C., … Lindenberger, U. (2010). Experience-dependent plasticity of white-matter microstructure extends into old age. *Neuropsychologia*, 48(13), 3878–3883. https://doi.org/10.1016/j.neuropsychologia.2010.08.026

Luciana, M. (2003). Practitioner review: Computerized assessment of neuropsychological function in children: Clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44(5), 649–663. https://doi.org/10.1111/1469-7610.00152

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, 36(3), 273–293. https://doi.org/10.1016/S0028-3932(97)00109-7

Luciana, M., & Nelson, C. A. (2002). Assessment of Neuropsychological Function Through Use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-Year-Old Children. *Developmental Neuropsychology*, 22(3), 595–624. https://doi.org/10.1207/S15326942DN2203_3

Luna, B., & Sweeney, J. A. (2004). *Cognitive development: functional magnetic resonance imaging studies*. In R. M. Keshavan, Matcheri S., Kennedy, James L., Murray (Ed.), Neurodevelopment and Schizophrenia (pp. 45–68). Cambridge University Press.

Lusby, C. M., Goodman, S. H., Bell, M. A., & Newport, D. J. (2014). Electroencephalogram patterns in infants of depressed mothers. *Developmental Psychobiology*, 56(3), 459–473. https://doi.org/10.1002/dev.21112

Mackey, A. P., Finn, A. S., Leonard, J. A., Jacoby-Senghor, D. S., West, M. R., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2015). Neuroanatomical Correlates of the Income-Achievement Gap. *Psychological Science*, 26(6), 925–933. https://doi.org/10.1177/0956797615572233

Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, *18*(5), 387-391.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*(5398), 77-80.

Markant, J., & Amso, D. (2013). Selective Memories: Infants' Encoding is Enhanced in Selection Via Suppression. *Developmental Science*, 16(6), 926–940. https://doi.org/10.1111/desc.12084.Selective

Markant, J., & Amso, D. (2016). The Development of Selective Attention Orienting is an Agent of Change in Learning and Memory Efficacy. *Infancy*, 21(2), 154–176.

Markant, J., Oakes, L. M., & Amso, D. (2015). Visual Selective Attention Biases Contribute to the Other-Race Effect Among 9-Month-Old Infants. *Developmental Psychobiology*, 355–365. https://doi.org/10.1002/dev.21375

Markham, J. a., & Greenough, W. T. (2004). Experience-driven brain plasticity: beyond the synapse. *Neuron Glia Biology*, 1(04), 351. https://doi.org/10.1017/S1740925X05000219

McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, 43(1), 2–15. https://doi.org/10.1016/S0018-506X(02)00024-7

McLaughlin, K. A., Sheridan, M. A., Winter, W., Fox, N. A., Zeanah, C. H., & Nelson, C. A. (2014). Widespread reductions in cortical thickness following severe early-life deprivation: A neurodevelopmental pathway to attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 76(8), 629–638. https://doi.org/10.1016/j.biopsych.2013.08.016

McNab, F., & Klingberg, T. (2007). Prefrontal cortex and basal ganglia control access to working memory. *Nature neuroscience*, *11*(1), 103-107.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000) Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11:735-759.

Miller, B. T., & D'Esposito, M. (2005). Searching for "the Top" in Top-Down Control. *Neuron*, 48(4), 535–538. https://doi.org/10.1016/j.neuron.2005.11.002

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Minagawa-Kawai, Y., Matsuoka, S., Dan, I., Naoi, N., Nakamura, K., & Kojima, S. (2009). Prefrontal activation associated with social attachment: Facial-emotion recognition in mothers and infants. *Cerebral Cortex*, 19(2), 284–292. https://doi.org/10.1093/cercor/bhn081

Mohr, C., Sandor, P. S., Landis, T., Fathi, M., & Brugger, P. (2005) Blinking and schizotypal thinking. *Journal of Psychopharmacology,* 19:513-520.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Morishima, Y., Akaishi, R., Yamada, Y., Okuda, J., Toma, K., & Sakai, K. (2009). Task-specific signal transmission from prefrontal cortex in visual selective attention. *Nature Neuroscience*, 12(1), 85–91. https://doi.org/10.1038/nn.2237

Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, 40(3), 255–265. https://doi.org/10.1002/dev.10033

Müller, J., Dreisbach, G., Brocke, B., Lesch, K. P., Strobel, A., & Goschke, T. (2007) Dopamine and cognitive control: The influence of spontaneous eyeblink rate, DRD4 exon III polymorphism and gender on flexibility in set-shifting. *Brain Research*, 1131:155-162.

Müller, J., Dreisbach, G., Goschke, T., Hensch, T., Lesch, K. P., & Brocke, B. (2007). Dopamine and cognitive control: the prospect of monetary gains influences the balance between flexibility and stability in a set-shifting paradigm. *European Journal of Neuroscience*, *26*(12), 3661-3668.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, 5(7), 309–315. https://doi.org/10.1016/S1364-6613(00)01682-X

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review*, 104(4), 686–713. https://doi.org/10.1037/0033-295X.104.4.686

Muñoz-Cuevas, F. J., Athilingam, J., Piscopo, D., & Wilbrecht, L. (2013). Cocaine-induced structural plasticity in frontal cortex correlates with conditioned place preference. *Nature Neuroscience*, 16(10), 1367–1369. https://doi.org/10.1038/nn.3498

Nepomnaschy, P., & Flinn, M. (2009). *Early life influences on the ontogeny of the neuroendocrine stress response in the human child*. In P. B. Ellison, Peter Thorpe, Gray (Ed.), Endocrinology of Social Relationships (pp. 364–382). Harvard University Press.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507-520.

Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., … Sowell, E. R. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, 18(5), 773–778. https://doi.org/10.1038/nn.3983

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, 10(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, 8(1), 74–

87. https://doi.org/10.1111/j.1467-7687.2005.00394.x

Noudoost, B., Chang, M. H., Steinmetz, N. A., & Moore, T. (2010). Top-down control of visual attention. *Current Opinion in Neurobiology*, 20(2), 183–190. https://doi.org/10.1016/j.conb.2010.02.003

O'Reilly, R. C. (2006). Biologically Based Computational Models of High-Level Cognition. *Science*, 314(5796), 91–94. https://doi.org/10.1126/science.1127242

Oakes, L. M., & Amso, D. (2018). *Development of Visual Attention*. In J. T. Wixted (Ed.), Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience (Fourth Edi, pp. 1–33). John Wiley & Sons. https://doi.org/10.1002/9781119170174.epcn401

Ochsner, K. N., & Gross, J. J. (2008). Cognitive Emotion Regulation: Insights from Social Cognitive and Affective Neuroscience. *Current Directions in Psychological Science*, 17(2), 153–158. https://doi.org/10.1111/j.1467-8721.2008.00566.x

Oh, J., Jeong, S. Y., & Jeong, J. (2012). The timing and temporal patterns of eye blinking are dynamically modulated by attention. *Human Movement Science*, 31:1353-1365.

Olesen, P. J., Nagy, Z., Westerberg, H., & Klingberg, T. (2003). Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network. *Cognitive Brain Research*, 18(1), 48–57. https://doi.org/10.1016/j.cogbrainres.2003.09.003

Paneri, S., & Gregoriou, G. G. (2017). Top-Down Control of Visual Attention by the Prefrontal Cortex. Functional Specialization and Long-Range Interactions.

Frontiers in Neuroscience, 11, 1–16. https://doi.org/10.3389/fnins.2017.00545

Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, 9(2), 60–68. https://doi.org/10.1016/j.tics.2004.12.008

Pearson, B. Z., Fernandez, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language learning*, *43*(1), 93-120.

Peelen, M. V, Glaser, B., Vuilleumier, P., & Eliez, S. (2009). Differential development of selectivity for faces and bodies in the fusiform gyrus. *Developmental Science*, 16–25. https://doi.org/10.1111/j.1467-7687.2009.00916.x

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, 21(1), 178–85. https://doi.org/10.3758/s13423-013-0466-4

Perkins, D. N., & Salomon, G. (1992). *Transfer of Learning*. In International Encyclopedia of Education (2nd ed., pp. 6452–6457).

Petanjek, Z., Judas, M., Simic, G., Rasin, M. R., Uylings, H. B. M., Rakic, P., & Kostovic, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 108(32), 13281–13286. https://doi.org/10.1073/pnas.1105108108

Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tétreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of child language*, *28*(02), 453-496.

Petrides, M., & Pandya, D. N. (2001). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, 16, 291–310. https://doi.org/10.1046/j.1460-9568.2002.02090.x

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Piaget, J. (1954). *The construction of reality in the child*. New York, NY: Basic Books.

Pitts, S. M., & Horvitz, J. C. (2000). Similar Effects of D1/D2 Receptor Blockade on Feeding and Locomotor Behavior. *Pharmacology Biochemistry and Behavior*, *65*(3), 433-438.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Rabinowicz, T., de Courten-Myers, G. M., Petetot, J. M., Xi, G., & de los Reyes, E. (1996). Human cortex development: estimates of neuronal numbers indicate major loss late during gestation. *Journal of Neuropathology and Experimental Neurology*, 55(3), 320–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8786390

Rakic, P. (2002). Pre- and post-developmental neurogenesis in primates. *Clinical Neuroscience Research*, 2(1–2), 29–39. https://doi.org/10.1016/S1566-

2772(02)00005-1

Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N., & Goldman-Rakic, P. S. (1986). Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science*, 232(4747), 232–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3952506

Reiss, A. L., Abrams, M. T., Singer, H. S., Ross, J. L., & Denckla, M. B. (1996). Brain development, gender and IQ in children. A volumetric imaging study. *Brain*, 5(12), 1763–74.

Richards, J. E. (2000). Localizing the development of covert attention in infants with scalp event-related potentials. *Developmental Psychology*, 36(1), 91–108. https://doi.org/10.1037/0012-1649.36.1.91

Romanski, L. M., & Goldman-Rakic, P. S. (2002). An auditory domain in primate prefrontal cortex. *Nature Neuroscience*, 5(1), 15–16. https://doi.org/10.1038/nn781

Rosa, E. C., Dickinson, D., Apud, J., Weinberger, D. R., & Elvevåg, B. (2010). COMT Val158Met polymorphism, cognitive stability and cognitive flexibility: an experimental examination. *Behavioral and Brain Functions*, 6(1), 53. https://doi.org/10.1186/1744-9081-6-53

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2011). Exogenous attention influences visual short-term memory in infants. *Developmental Science*, 14(3), 490–501. https://doi.org/10.1111/j.1467-7687.2010.00992.x

Rossi, A. F., Pessoa, L., Desimone, R., & Ungerleider, L. G. (2009). The prefrontal cortex and the executive control of attention. *Experimental Brain Research*, 192, 489–497. https://doi.org/10.1007/s00221-008-1642-z

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20), 7338–7343. https://doi.org/10.1073/pnas.0502455102

Rovee-Collier, C., & Cuevas, K. (2009). Multiple memory systems are unnecessary to account for infant memory development: An ecological model. *Developmental Psychology*, 45(1), 160–174. https://doi.org/10.1037/a0014538

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, 42(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Ruff, C. C., Bestmann, S., Blankenburg, F., Bjoertomt, O., Josephs, O., Weiskopf, N., … Driver, J. (2008). Distinct Causal Influences of Parietal Versus Frontal Areas on Human Visual Cortex: Evidence from Concurrent TMS-fMRI. *Cerebral Cortex*, 18, 817–827. https://doi.org/10.1093/cercor/bhm128

Ruff, H. A. (1981). Effect of context on infants' responses to novel objects. *Developmental Psychology*, 17(1), 87–89. https://doi.org/10.1037//0012-1649.17.1.87

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, 8(9), 410–417. https://doi.org/10.1016/j.tics.2004.07.009

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Saito, Y., Aoyama, S., Kondo, T., Fukumoto, R., Konishi, N., Nakamura, K., … Toshima, T. (2007). Frontal cerebral blood flow change associated with infant-directed speech. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 92, F113–F116. https://doi.org/10.1136/adc.2006.097949

Schneirla, T. C. (1957). *The concept of development in comparative psychology*. In D. B. Harris (Ed.), The concept of development. (pp. 78–108). Minneapolis: University of Minnesota Press.

Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, 26(2), 100–107. https://doi.org/10.1016/S0166-2236(02)00037-1

Seger, C. A., Poldrack, R. A., Prabhakaran, V., Zhao, M., Glover, G. H., & Gabrieli, J. D. (2000) Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia*, 38, 1316-1324.

Selemon, L. D. (2013). A role for synaptic plasticity in the adolescent development of executive function. *Translational Psychiatry*, 3(3), e238. https://doi.org/10.1038/tp.2013.7

Shankle, W. R., Rafii, M. S., Landing, B. H., & Fallon, J. H. (1999). Approximate doubling of numbers of neurons in postnatal human cerebral cortex and in 35 specific cytoarchitectural areas from birth to 72 months. *Pediatric and Developmental Pathology*, 2(3), 244–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10191348

Shaw, P., Greenstein, D., Lerch, J., Clasen, L. S., Lenroot, R., Gogtay, N., … Giedd, J. N.

(2006). Intellectual ability and cortical development in children and adolescents. *Nature*, 440(7084), 676–679. https://doi.org/10.1038/nature04513

Shaw, P., Kabani, N. J., Lerch, J., Eckstrand, K., Lenroot, R., Gogtay, N., … Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. The *Journal of Neuroscience*, 28(14), 3586–3594. https://doi.org/10.1523/JNEUROSCI.5309-07.2008

Sheridan, M. A., Fox, N. A., Zeanah, C. H., McLaughlin, K. A., & Nelson, C. A. (2012). Variation in neural development as a result of exposure to institutionalization early in childhood. *Proceedings of the National Academy of Sciences of the United States of America*, 109(32), 12927–32. https://doi.org/10.1073/pnas.1200041109

Sheridan, M. A., Kharitonova, M., Martin, R. E., Chatterjee, A., & Gabrieli, J. D. E. (2014). Neural Substrates of the Development of Cognitive Control in Children Ages 5–10 Years. *Journal of Cognitive Neuroscience*, 26(8), 1840–1850. https://doi.org/10.1162/jocn_a_00597

Shimi, A., Kuo, B., Astle, D. E., Nobre, A. C., & Scerif, G. (2014). Age Group and Individual Differences in Attentional Orienting Dissociate Neural Mechanisms of Encoding and Maintenance in Visual STM. *Journal of Cognitive Neuroscience*, 26(4), 864–877. https://doi.org/10.1162/jocn

Shokri-Kojori, E., Motes, M. a., Rypma, B., & Krawczyk, D. C. (2012). The Network Architecture of Cortical Processing in Visuo-spatial Reasoning. *Scientific Reports*, 2, 1–7. https://doi.org/10.1038/srep00411

Shomstein, S., & Gottlieb, J. (2016). Spatial and non-spatial aspects of visual attention:

Interactive cognitive mechanisms and neural underpinnings. *Neuropsychologia*, 92, 9–19. https://doi.org/10.1016/j.neuropsychologia.2016.05.021

Shukla, D. (1985) Blink rate as clinical indicator. *Neurology*, 35:286.

Shultz, T. R., Rivest, F., Egri, L., Thvierge, J.-P., & Dandurand, F. (2007). Could knowledge-based neural learning be useful in developmental robotics? The case of KBCC. *International Journal of Humanoid Robotics*, 04(02), 245–279. https://doi.org/10.1142/S0219843607001035

Sidman, R. L., & Rakic, P. (1973). Neuronal migration, with special reference to developing human brain: a review. *Brain Research*, 62(1), 1–35. https://doi.org/10.1016/0006-8993(73)90617-3

Simmonds, D. (2015). Protracted development of brain systems underlying working memory in adolescence: a longitudinal study. Retrieved from http://d-scholarship.pitt.edu/25045/

Slagter HA, Georgopoulou K, Frank MJ (2015) Spontaneous eye blink rate predicts learning from negative, but not positive, outcomes. *Neuropsychologia* 71:126-132.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the Context of Acting: The Task Dynamics of the A-Not-B Error. *Psychological Review*, 106(2), 235–260.

Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition*, 116(2), 155–167. https://doi.org/10.1016/j.cognition.2010.04.007

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology*, 46(1), 129–138. https://doi.org/10.1037/a0014618

Spear, N. E. (1984). *The future study of learning and memory from a psychobiological perspective*. In A. Sarris, V., Parducci (Ed.), Perspectives in psychological experimentation (pp. 87–103). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Sporns, O. (2013). The human connectome: Origins and challenges. *NeuroImage*, 80, 53–61. https://doi.org/10.1016/j.neuroimage.2013.03.023

Stevens, B., Porta, S., Haak, L. L., Gallo, V., & Fields, R. D. (2002). Adenosine: a neuron-glial transmitter promoting myelination in the CNS in response to action potentials. *Neuron*, 36(5), 855–68. https://doi.org/10.1016/S0896-6273(02)01067-X

Stiles, J., & Jernigan, T. L. (2010). The basics of brain development. *Neuropsychology Review*, 20(4), 327–348. https://doi.org/10.1007/s11065-010-9148-4

Strange, B. A., Henson, R. N. A., Friston, K. J., & Dolan, R. J. (2001) Anterior prefrontal cortex mediates rule learning in humans. *Cerebral Cortex,* 11:1040-1046.

Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition*, 20(1), 8–23. https://doi.org/10.1016/0278-2626(92)90059-U

Supekar, K., Musen, M., & Menon, V. (2009). Development of large-scale functional brain networks in children. *PLoS Biology*, 7(7). https://doi.org/10.1371/journal.pbio.1000157

Tada, H. (1977) Spontaneous blinking during a visual tracking performance. *Fukushima Journal of Medical Science*, 25, 91-100.

Tadic, V., Pring, L., Dale, N., Tadić, V., Pring, L., & Dale, N. (2009). Attentional processes in young children with congenital visual impairment. *British Journal of Developmental Psychology*, 27(2), 311–330. https://doi.org/10.1348/026151008X310210

Taga, G., & Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *NeuroImage*, 36, 1246–1252. https://doi.org/10.1016/j.neuroimage.2007.04.037

Taylor, J. R., Elsworth, J. D., Lawrence, M. S., Sladek Jr, J. R., Roth, R. H., & Redmond Jr, D. E. (1999). Spontaneous blink rates correlate with dopamine levels in the caudate nucleus of MPTP-treated monkeys. *Experimental neurology*, *158*(1), 214-220.

Taylor, P. C. J., Nobre, A. C., Rushworth, M. F. S., Hospital, J. R., Way, H., & Ox, O. (2007). FEF TMS Affects Visual Cortical Activity. *Cerebral Cortex*, 17, 391–399. https://doi.org/10.1093/cercor/bhj156

Thatcher, R. W. (1992). Cyclic cortical reorganization during early childhood. *Brain and Cognition*, 20(1), 24–50. https://doi.org/10.1016/0278-2626(92)90060-Y

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24, 1–86.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition Without Control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science,* 18(5), 259–263. https://doi.org/10.1111/j.1467-8721.2009.01648.x

Topal, J., Gergely, G., Miklosi, A., Erdohegyi, A., & Csibra, G. (2008). Infants'

Perseverative Search Errors Are Induced by Pragmatic Misinterpretation. *Science*, 321(September), 1831–1835.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009) The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168, 242–249.

Tsuchida, A., Doll, B. B., & Fellows, L. K. (2010). Beyond reversal: a critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. *Journal of Neuroscience*, 30(50), 16868–16875. https://doi.org/10.1523/JNEUROSCI.1958-10.2010

Tummeltshammer, K. S., & Amso, D. (2018). Top-down contextual knowledge guides visual attention in infancy. *Developmental Science*, 21, 1–9. https://doi.org/10.1111/desc.12599

Tummeltshammer, K. S., Mareschal, D., & Kirkham, N. Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Development*, 85(5), 1981–1994. https://doi.org/10.1111/cdev.12239

Turkewitz, G., & Kenny, P. A. (1982). Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychobiology*, 15(4), 357–368. https://doi.org/10.1002/dev.420150408

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14(6), 623–8. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

Ungerleider, L. G., Gaffan, D., & Pelak, V. S. (1989). Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys. *Experimental Brain Research*, 76, 473–484.

Ursache, A., & Noble, K. G. (2016). Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology*, 53(1), 71–82. https://doi.org/10.1111/psyp.12547

van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, 17(12), 683–696. https://doi.org/10.1016/j.tics.2013.09.012

Van Essen, D. C., & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, 6(C), 370–375. https://doi.org/10.1016/0166-2236(83)90167-4

Von Hofsten, C. (2004). An action perspective on motor development. *Trends in cognitive sciences*, *8*(6), 266-272.

Wake, H., Lee, P. R., & Fields, R. D. (2011). Control of Local Protein Synthesis and Initial Events in Myelination by Action Potentials. *Science*, 333(6049), 1647–1651. https://doi.org/10.1126/science.1206998

Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of Inferior Temporal Areas TEO and TE with Parietal and Frontal Cortex in Macaque Monkeys. *Cerebral Cortex*, 4(5), 470–483.

Wellman, H. M., Cross, D., Bartsch, K., & Harris, P. L. (1986). Infant Search and Object Permanence: A Meta-Analysis of the A-Not-B Error. *Monographs of the Society for Research in Child Development*, 51(3), 1–67.

Wendelken, C., Ferrer, E., Whitaker, K. J., & Bunge, S. A. (2016). Fronto-parietal network reconfiguration supports the development of reasoning ability. *Cerebral*

Cortex, 26(5), 2178–2190. https://doi.org/10.1093/cercor/bhv050

Wendelken, C., Munakata, Y., Baym, C., Souza, M., & Bunge, S. A. (2012). Flexible rule use: Common neural substrates in children and adults. *Developmental Cognitive Neuroscience*, 2(3), 329–339. https://doi.org/10.1016/j.dcn.2012.02.001

Werchan, D. M., & Amso, D. (2017). A Novel Ecological Account of Prefrontal Cortex Functional Development. *Psychological Review*, 124(6), 720–739.

Werchan, D. M., Baumgartner, H. A., Lewkowicz, D. J., & Amso, D. (2018). The origins of cortical multisensory dynamics: Evidence from human infants. *Developmental Cognitive Neuroscience*, 34, 75–81. https://doi.org/10.1016/j.dcn.2018.07.002

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2015) Spontaneous Hierarchical Rule Learning in 8-Month-Old Infants. *Psychological Science*, 26(6), 805–815. https://doi.org/10.1177/0956797615571442

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2016). Role of Prefrontal Cortex in Learning and Generalizing Hierarchical Rules in 8-Month-Old Infants. *Journal of Neuroscience*, 36(40), 10314–10322. https://doi.org/10.1523/JNEUROSCI.1351-16.2016

Werchan, D. M., Lynn, A., Kirkham, N. Z., & Amso, D. (in press). The Emergence of Object-Based Visual Attention in Infancy: A Role for Family Socioeconomic Status and Competing Visual Features. *Infancy*.

White I. M., & Wise, S. P. (1999) Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research*, 126, 315-335.

White, B. R., Snyder, A. Z., Cohen, A. L., Petersen, S. E., Raichle, M. E., Schlaggar, B. L., & Culver, J. P. (2009). Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. *NeuroImage*, 47, 148–156. https://doi.org/10.1016/j.neuroimage.2009.03.058

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. A. (2008) Hemodynamic response to featural changes in the occipital and inferior temporal cortex in infants: a preliminary methodological exploration. *Developmental Science, 11*, 361-370.

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., Armstrong, J., & Boas, D. (2009) Hemodynamic changes in the infant cortex during the processing of featural and spatiotemporal information. *Neuropsychologia*, 47:657-662.

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., Boas, D. A. (2005) Using near-infrared spectroscopy to assess neural activation during object processing in infants. *Journal of Biomedical Optics, 10*, 1010-1019.

Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*, 107, 118–136. https://doi.org/10.1016/j.jecp.2010.04.014

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, *16*(5), 372-377.

Yang, G., Pan, F., & Gan, W. B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature*, 462(7275), 920–924. https://doi.org/10.1038/nature08577\rnature08577 [pii]

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244–262. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0010027712001369

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience*, 14(5), 656–661. https://doi.org/10.1038/nn.2773.Causal

Zecevic, N. (1998). Synaptogenesis in layer I of the human cerebral cortex in the first half of gestation. *Cerebral Cortex*, 8(3), 245–252. https://doi.org/10.1093/cercor/8.3.245

Zecevic, N., Bourgeois, J. P., & Rakic, P. (1989). Changes in synaptic density in motor cortex of rhesus monkey during fetal and postnatal life. *Developmental Brain Research, 50*(1), 11–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2582602

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., … Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3), vii-137. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14723273