



Top-down knowledge rapidly acquired through abstract rule learning biases subsequent visual attention in 9-month-old infants

Werchan D.M. *, Amso D.

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, 02912 United States

ARTICLE INFO

Keywords:

Infants
PFC
Rule learning
Visual attention
Feature-Based attention
Learning

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 whether learning a hierarchical rule structure, where there is a higher-order feature that organizes visual inputs into predictable sequences, subsequently biases 9-month-old infants' visual attention to the higher-order visual feature. In Experiment 1, we found that individual differences in infants' ability to structure simple visual inputs into generalizable rules was related to the change in infants' attention biases towards higher-order features. In Experiment 2, we found that increased functional connectivity between the PFC and visual cortex was related to the efficacy of rule learning. Moreover, Granger causality analyses provided exploratory evidence that increased functional connectivity reflected PFC influence over visual cortex. These findings provide new insights into how the infant brain learns to flexibly select features from the cluttered visual world that were previously relevant for learning and behavior.

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 and Johnson, 2006; Cheng et al., 2019; Johnson et al., 2003; Markant and Amso, 2013b, 2016; Markant et al., 2015; Ross-Sheehy et al., 2011; Wu and Kirkham, 2010). Yet, less is known about how rapidly acquired information subsequently biases attentional selection in infancy. Prior work has shown that infants are capable of learning hierarchical rule structures, which involve a higher-order context governing stimulus-response associations. The term "higher-order context" is drawn from the hierarchical reinforcement learning literature (Collins et al., 2014; Collins and Frank, 2013; Donoso et al., 2014; Frank and Badre, 2012), and can be a stimulus, space, object, or person that organizes inputs into abstract rule structures. Here we asked whether abstract rule learning is a mechanism by which the infant brain learns which competing features of a cluttered visual environment are informative for subsequent attentional selection in a novel environment.

Visual attention shows rapid developmental changes over the first year of postnatal life (Amso and Scerif, 2015; Oakes and Amso, 2018).

Visual attention can either be bottom-up driven, based on saliency maps of the visual-spatial environment (e.g., Althaus and Mareschal, 2012; Amso et al., 2014; Frank et al., 2009a), 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 (Johnson and Vecera, 1996; Tummelshammer and Amso, 2018; Tummelshammer et al., 2014; Werchan et al., 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 and Scerif, 2015; Carrasco, 2011; Desimone and Duncan, 1995; Gilbert and Li, 2013). 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 established relationships between attention and learning in infancy. For example, Markant, et al. (2016) used a spatial

* Corresponding Author at: Postal Address: 190 Thayer St. Providence, RI, 02912 United States.

E-mail address: Denise_Werchan@Brown.edu (D.M. Werchan).

cueing task to bias 9-month-old infants' attention to either own- or other-race faces. The authors found that infants discriminated faces in the focus of the attention bias, regardless of race, 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 and Amso, 2018).

Prior work has also shown that infants as young as 8-months of age recruit the PFC to organize visual inputs into abstract rules that support flexible learning in novel contexts (Werchan et al., 2015, 2016). These studies found that 8-month-old infants can use simple visual features, such as the shape of an object, as higher-order contexts to structure inputs into stimulus-response rules. Importantly, infants were able to generalize these rules to support flexible learning in new contexts (e.g., to new shapes). These data indicate that PFC is online and acting as an important information gathering mechanism in infancy. As in any developing system, the data also showed that there is substantial individual variability in infants' learning and generalization of abstract rules (Werchan et al., 2015, 2016), which was also associated with individual differences in PFC activation during learning (Werchan et al., 2016). The present work was designed to explore the value of this mechanism to infant learners. We reasoned that structuring visual inputs into abstract rules may assign attentional priority to the higher-order features that cue these rule structures. The prediction would be that infants who learn that a higher-order feature is relevant for organizing stimulus-response associations into abstract rules should be biased to subsequently select that visual feature for processing in novel situations.

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 and Duncan, 1995; Gilbert and Li, 2013; Noudoost et al., 2010; Shomstein and Goltieb, 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 et al., 2002; Kolb et al., 2012; Miller and Cohen, 2001; O'Reilly, 2006; Rougier et al., 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 and Li, 2013; Miller and 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 cortex (Barbas, 2000; Petrides and Pandya, 2001; Ungerleider et al., 1989; Webster et al., 1994). Functional interactions within these corticocortical connections are thought to be the basis for PFC modulation of neuronal activity in early visual areas (Baluch and Itti, 2011; Paneri and 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 and Shulman, 2002; Gazzaley et al., 2007; Kastner and Ungerleider, 2000; Morishima et al., 2009; Rossi et al., 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 et al., 2000; Capotosto et al., 2009; Ruff et al., 2008; Taylor et al., 2007; Zanto et al., 2011). Yet, it is unclear whether similar mechanisms operate in young infants.

As such, the current study examines whether top-down knowledge rapidly acquired through abstract rule learning influences subsequent downstream visual attention in 9-month-old infants. We tested 9-month-olds, given that infants of this age are capable of both engaging in abstract rule learning mechanisms that involve the PFC (Werchan et al., 2015, 2016) and engaging in top-down guidance of spatial attention (Amso and Johnson, 2006, 2008; Tummeltshammer and Amso, 2018). In a first experiment, we used a behavioral paradigm to test whether

abstract rule learning subsequently biases infants' attention to task-relevant visual information. To examine this question, we first measured infants' baseline visual attention biases to color and shape information using an attention bias priming task (adapted from Werchan et al., 2019). During this task, infants were primed with an object image, after which they immediately saw two test items presented side-by-side (Fig. 1). One test item matched the prime in color but differed in shape (color-match item) and the other test item matched the prime in shape but differed in color (shape-match item). We measured the distribution of infants' looking to the color-match and shape-match items as an index of their attention biases. Importantly, the timing parameters of this task were based on similar tasks in the infant visual attention literature (see Werchan et al., 2019). For instance, spatial cueing paradigms, where a brief cue elicits an excitation toward a stimulus with short delays, would suggest that a cue followed immediately by a probe stimulus should benefit or facilitate the visual feature that is currently excited over the short cue-target duration (Markant and Amso, 2013a; Posner, 1980; Richards, 2000). The same timing parameters have been used in infant negative priming studies, where attended featural information is enhanced with short delays between prime and probe target displays (Amso and Johnson, 2005, 2008; Tipper, 1985). Based on this literature, we interpret greater looking as evidence that that feature had been selected for attentional priority.

After the attention bias priming task, we then presented infants with several visual stimulus-response pairings during an abstract rule learning task, where infants could use visual feature of color or shape as a higher-order context to organize stimulus-response inputs into abstract rules for action. In a Shape Contexts condition of the abstract rule learning task, infants could use shape as a higher-order context to organize visual inputs into simpler color-response rules (Fig. 2A). In a Color Contexts condition, infants could use color as a higher-order context to organize visual inputs into simpler shape-response rules (Fig. 2B). Importantly, we measured the efficacy of infants' abstract rule learning by assessing their ability to generalize these rules to novel shapes or colors after initial learning. Finally, after the rule learning task we then re-measured infants' attention biases to both color and shape to examine whether abstract rule learning influenced infants' attention biases. We predicted that individual differences in the efficacy of infants' generalization of abstract rules would relate to the degree of change in infants' attention biases to the visual features used as higher-order contexts (color or shape).

In a second experiment, we attempted to replicate and extend the behavioral results of Experiment 1, and also incorporated functional near-infrared spectroscopy (fNIRS) to explore the role of the PFC and visual cortex in our tasks. Recall that neuroimaging and anatomical studies establish the PFC as a source of top-down modulatory control over processing in early visual areas (e.g., Barceló et al., 2000; Capotosto et al., 2009; Ruff et al., 2008; Taylor et al., 2007; Zanto et al., 2011). Thus, we examined directed functional connectivity between the PFC and visual cortex during rule learning as a correlational measure of PFC modulation of visual attention. We predicted that individual differences in functional connectivity would relate to the efficacy of abstract rule learning, as well as the subsequent change in infants' attention biases to task-relevant visual features. Examining the impact of abstract rule learning on shaping visual selection may provide mechanistic insights into how the infant brain *learns* to efficiently direct attention to previously relevant visual information in a cluttered environment with many competing options.

1. Experiment 1

In our first experiment, we used a behavioral paradigm to examine whether top-down knowledge acquired through abstract rule learning influences downstream visual attention in 9-month-old infants. We first measured infants' baseline attention biases to both color and shape using an attention bias priming task (Werchan et al., 2019). Infants were

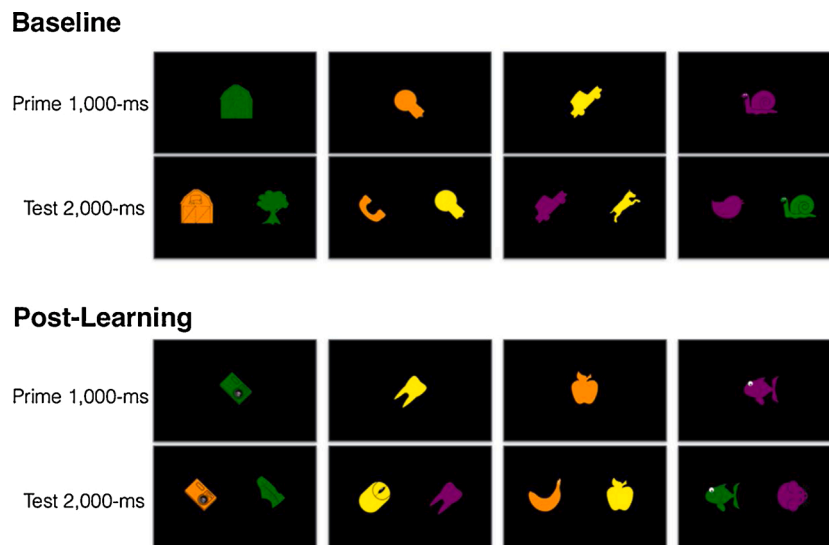


Fig. 1. Schematic of all trials used in the baseline and post-learning attention bias priming task, which was used to measure infants' attention biases to color and shape.

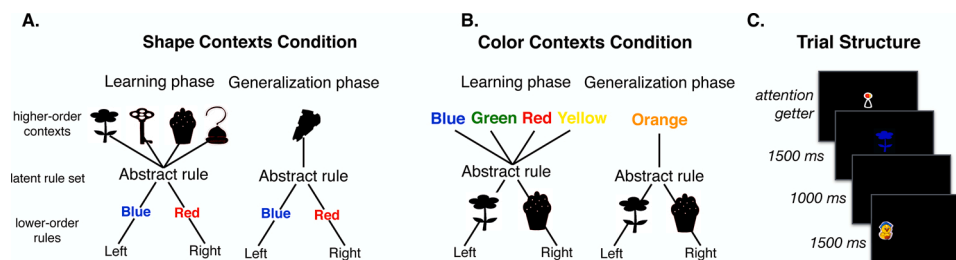


Fig. 2. Example of the hierarchical structure used during the abstract rule learning task for the Shape Contexts condition (A) and the Color Contexts condition (B). Each trial consisted of the central stimulus, followed by a cartoon reward presented on the right or left of the screen (C).

then randomly assigned to a Shape Contexts or a Color Contexts condition during a rule learning and generalization task. During this task, we presented infants with visual stimulus-response pairs, where simple visual features (color or shape) can act as higher-order contexts that organize visual inputs for learning. Specifically, infants could use the visual feature of shape as a higher-order context to organize inputs into simpler color-response rules in a Shape Contexts condition (Fig. 2A). In an analogous Color Contexts condition, infants could use the visual feature of color as a higher-order context to organize visual inputs into shape-response rules (Fig. 2B). Finally, we re-measured infants' attention biases to both color and shape. We predicted that individual differences in the efficacy of rule learning would correlate with changes in infants' attention biases to visual features that act as higher-order contexts to organize inputs for learning.

2. Experiment 1 method

2.1. Participants

The final sample consisted of 40 nine-month-old infants ($M = 9.45$ months, $SD = 0.67$ months, 20 females, 20 males, 29 white non-Hispanic, 3 black, 5 Hispanic, 2 Asian, and 1 Mixed Race/Other). Infants were randomly assigned to a Shape Contexts condition ($N = 20$, $M = 9.59$ months, $SD = 0.84$ months) or a Color Contexts condition ($N = 20$, $M = 9.45$ months, $SD = 0.43$ months). Sample size was determined based on an a priori power analysis with a large effect size ($d = 0.4$) 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 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 (< 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. Families were compensated for time and travel to our laboratory.

2.2. 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.

2.3. Procedure

The study consisted of (1) a baseline attention bias priming task, (2) a rule learning task, (3) a post-learning attention bias priming task. All infants received the same attention bias priming task at baseline and after learning. This task measured infants' attention biases to both color and shape. Infants were randomly assigned to either a Shape Contexts

condition or a Color contexts condition for the abstract rule learning task. Different colors and shapes were used in the attention bias priming task than the lower-order colors and shapes used in the rule learning and generalization tasks (see Fig. 1 for all stimuli used in the attention bias priming task, and Fig. 2 for all stimuli used in the rule learning and generalization task).

2.3.1. 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 (Fig. 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 2000-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.

2.3.2. Rule learning task

The rule learning task consisted of a learning phase followed by a generalization phase. During the learning phase, infants were presented with several visual stimulus-response pairs in a Shape Contexts condition (Fig. 2A) or a Color Contexts condition (Fig. 2B), where the centrally-presented stimulus would predict a cartoon reward that subsequently appeared on the right or left of the screen. In the Shape Contexts condition, the central stimuli varied by the visual feature of shape, which infants could use as a higher-order context to organize inputs into a set of simpler color-response rules (e.g., red-colored stimuli predict a cartoon reward appearing on the left of the screen, blue-colored stimuli predict a cartoon reward appearing on the right of the screen). In the analogous Color Contexts condition, the central stimuli varied by color, which infants could use to organize inputs into a set of simpler shape-response rules (e.g., flower-shaped stimuli predict a cartoon reward appearing on the left of the screen, cupcake-shaped stimuli predict a cartoon reward appearing on the right of the screen).

At the start of each trial during the learning and generalization phase, 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 stimulus was first displayed in the center of the screen for 1,500-ms. The stimulus 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 (Fig. 2C). Infants were presented with two blocks of eight 4,000-ms stimulus-response trials, for a total of 16 trials (~1-min), during the learning phase. The number of trials was determined by prior studies using similar spatiotemporal learning paradigms in infants (e.g., Johnson et al., 1991, 1994; Tummelshammer and Amso, 2018; Werchan et al., 2015). The presentation order of the trials was randomized in each of the two blocks of the learning phase.

After the learning phase, infants were then presented with two novel stimulus-response pairings during a generalization phase to assess whether infants learned and generalized an abstract rule (i.e., one that is not tied to the specific context in which it was learned in; Fig. 2). In the Shape Contexts condition, infants were presented with a novel shape that had the same color-response rules as in the previous learning phase. In the Color Contexts condition, infants were presented with a novel color that had the same shape-response rules as in the previous learning phase. Infants were presented with a total of 8 trials in a randomized

order. The same trial timing parameters were used during the generalization phase as in the learning phase.

3. Data processing

3.1. Attention bias priming task

We examined infants' attention biases by measuring the distribution of infants' looking to the color-match and the shape-match test items. 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 period was excluded from the analysis for each test trial to account for the time required for infants to make a saccade away from the prime stimulus toward either the left or right-test item after the appearance of the test display.

Our measure of interest was the change in infants' attention biases after the abstract rule learning, task relative to each infant's own baseline bias. We operationalized infants' attention biases as the difference in total duration of looking (out of 2,000-ms) at the shape-match relative to the color-match items a t-test (see also Werchan et al., 2019). Thus, we calculated infants' shape bias 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. Infants' color bias was the inverse of their shape bias. We then calculated the change in infants' attention biases by subtracting infants' baseline attention bias scores from the post-learning attention bias scores. Thus, positive difference scores for color biases indicate a greater attention bias to color after learning relative to baseline, and positive difference scores for shape biases indicate a greater attention bias to shape after learning relative to baseline. Difference scores near zero indicate no change in infants' attention biases after learning relative to baseline. We verified that infants' attention bias scores at baseline and post-learning were normally distributed through visual inspection and statistically using the Kolmogorov-Smirnov test for normality, all $ps > 0.127$.

3.2. Abstract rule learning task

We examined rule learning and generalization performance by measuring infants' eye movement reaction times from trial onset (i.e., when the central stimulus first appeared on the screen) to the time it took to arrive at the correct left/right reward location using SMI BeGaze software. We calculated eye movement reaction times using methods from prior similar studies that interpret decreases in eye movement reaction times as evidence of learning in infants under 12 months of age (Amso and Johnson, 2006; Frank et al., 2009b; Kirkham et al., 2007; Marcus et al., 2007, 1999; Markant and Amso, 2013a; Tummelshammer and Amso, 2018; Tummelshammer and Kirkham, 2013; Werchan et al., 2015). Trials with eye movement reaction times slower than two standard deviations from the median for each infant were excluded from analysis, which resulted in removal of 16 trials (1.67 %). Additionally, our paradigm was not gaze-contingent and infants were allowed to explore the screen freely throughout the trial. Thus, we did not remove trials if infants looked first to the incorrect side before looking to the correct side, but we only recorded reaction times for entry to the correct location for analysis. In this way, relatively slower reaction times indicate that infants are making exploratory eye movements before the target appears and had not yet learned the rule.

Our measures of interest were (1) learning of the stimulus-response associations during the initial learning phase of the task, and (2) generalization of learning to novel higher-order contexts during the generalization phase of the task. Learning of the stimulus-response associations was operationalized as a decrease in infants' eye movement reaction times with trial exposure (Amso and Johnson, 2006; Frank et al., 2009a; Kirkham et al., 2007; Marcus et al., 2007, 1999; Markant and Amso, 2013a; Tummelshammer and Amso, 2018;

Tummeltshammer and Kirkham, 2013; Werchan et al., 2015). We calculated changes in infants' reaction times by binning every four consecutive trials in the learning phase to create four learning trial bins. Generalization was operationalized as faster reaction times during the generalization phase *relative* to the learning phase within each infant. This strategy corrects for speed of processing differences across individual infants that may otherwise shape the data. We calculated infants' generalization scores by subtracting the average reaction times during the generalization phase from the average reaction times during the learning phase of the abstract rule learning task. Thus, positive values indicate better generalization of abstract rules, and values near or less than zero indicate no generalization. We verified that these measures were normally distributed through visual inspection and statistically using the Kolmogorov-Smirnov test for normality, all $ps > 0.092$.

4. Experiment 1 results

4.1. Baseline attention biases and rule learning and generalization performance

We first examined infants' baseline biases to color/shape, measured during the attention bias priming task, as well as infants' initial learning of the stimulus-response associations during the learning phase of the rule learning task. At baseline, we found that all infants had a significantly greater attention bias to color over shape (calculated as a difference score of looking to *color-match* relative to *shape-match* items), $t(39) = 3.416, p = 0.001$, and that there were no significant differences in this bias between infants in the Color Contexts condition ($M = 108.35$ -ms, $SD = 277.42$ -ms) and the Shape Contexts condition ($M = 179.83$ -ms, $SD = 257.83$ -ms), $t(38) = 0.844, p = 0.404$.

We then examined infants' eye movement reaction times during the learning phase to test whether infants were learning the initial stimulus-response associations with trial exposure. We conducted a repeated-measures ANOVA with Learning Trial Bin (1, 2, 3, 4) as a within-subjects variable, Condition (Shape Contexts, Color Contexts) as a between-subjects variable, and infants' averaged eye movement reaction times as the dependent variable. This analysis revealed a main effect of Learning Trial Bin, $F(3,114) = 8.453, p < .001$, and a main effect of Condition, $F(1,38) = 12.849, p = .001$. These findings indicate that infants learned the stimulus-response associations with trial exposure. They also show that infants' who were assigned to the Color Contexts condition had overall faster reaction times ($M = 2508.54$ -ms, $SD = 44.87$ -ms) than those assigned to the Shape Contexts condition ($M = 2753.64$ -ms, $SD = 42.23$ -ms). This main effect of Condition did not interact further with Learning Trial Bin, $F(3,114) = 0.968, p = .411$, indicating that the *rate* of learning was equivalent between the two conditions.

We next examined whether infant learning (change in reaction time with trial exposure) reflects low-level stimulus-response associative learning, or whether infants formed abstract rules based on shape/color higher-order contexts. Infants were presented with a generalization task where novel contexts governed previously learned stimulus-response associations (Fig. 2). We conducted an ANOVA with Condition (Shape Contexts, Color Contexts) as a between-subjects variable and infants' generalization scores (calculated as a difference score of average reaction times during Learning – Generalization) as the dependent variable. Results indicated that all infants, as a group, had faster reaction times in generalization relative to learning, as evidenced by generalization scores significantly greater than zero, $F(1,38) = 15.599, p < .001$. We also found a main effect of Condition, $F(1,38) = 4.394, p = .043$, indicating that generalization performance was better in the Color Contexts condition ($M = 245.85$ -ms, $SD = 44.06$ -ms) than in the Shape Contexts condition ($M = 75.21$ -ms, $SD = 68.44$ -ms). We then conducted two one-sample t tests comparing generalization scores to chance (0) separately for each condition. These analyses revealed that generalization scores were significantly greater than chance in the Color Contexts condition, t

(19) = 5.579, $p < 0.001$, but performance was overall more variable in the Shape Contexts condition, $t(19) = 1.099, p = 0.285$ (Fig. 3).

To further ensure that generalization was specific to the higher-order shape or color context, we verified that reaction times during the fourth bin of the shape-contexts learning task were not correlated with reaction times during the first bin of the color-contexts generalization task, $r(20) = -0.197, p = .433$. Reaction times during the fourth bin of the color-contexts learning task were also not correlated with reaction times during the first bin of the shape-contexts generalization task, $r(20) = -0.203, p = .391$.

4.2. Influence of abstract rule learning on change in attention biases to color/shape from baseline

We next examined the critical prediction of this work: whether the efficacy of learning abstract rules cued by color or shape influenced what visual features infants subsequently attended to after the rule learning task, relative to baseline. We conducted correlational analyses to test whether the efficacy of infants' generalization, here the only measure of abstract rule learning, related to the degree of change in infants' attention biases from baseline to post-learning. Our results revealed that infants' generalization scores were correlated with the change in infants' attention bias to shape (calculated as a difference score of infants' shape biases after learning relative to baseline) in the Shape Contexts condition, $r(20) = 0.535, p = .015$ (Fig. 4A), such that better generalization of abstract rules cued by shape was associated with a greater change in infants' attention biases to shape. These data indicate that there was some degree of abstract rule learning for infants assigned to the Shape Contexts condition, despite the observed variability in infants' generalization scores. Mirroring these findings, infants' generalization scores in the Color Contexts condition was also correlated with a greater change in infants' attention biases to color (calculated as a difference score of infants' color biases after learning relative to baseline), $r(20) = 0.730, p < .001$ (Fig. 4B). Taken together, these results provide behavioral evidence that individual differences in the efficacy of abstract rule learning influences infants' attention to the visual features that organize inputs into rules for learning.

5. Experiment 2

In Experiment 1, we found that individual differences in abstract rule learning were related to the change in infants' attention biases to higher-order features, providing evidence that rapidly acquired top-down knowledge can influence subsequent downstream visual selection in infancy. We also found that infants assigned to the Color Contexts condition garnered more robust generalization results than infants assigned to the Shape Contexts condition. Nonetheless, infants in both conditions showed similar learning rates, as well as changes in their attention bias to color or shape, respectively. This suggests that the increased variability in the generalization data observed in the Shape Contexts condition may have been influenced by external factors or characteristics of the sample. As such, we repeated the Shape Contexts condition in Experiment 2 to attempt to replicate and extend our findings. This was the first goal of Experiment 2.

A second goal was to examine whether functional connectivity between the PFC, which we have shown to be involved in hierarchical rule learning in infancy (Werchan et al., 2016), and visual cortex is associated with task performance. In adults, top-down visual attention is thought to be mediated by functional interactions between the PFC and visual cortex (Gilbert and Li, 2013; Baluch and Itti, 2011; Paneri and Gregoriou, 2017). Prior studies also indicate that individual differences in PFC activation during rule learning is related to the efficacy of learning and generalization of hierarchical rules in infants (Werchan et al., 2016). Therefore, we measured whether infants' cortical activity when learning an abstract rule relates to the efficacy of generalization as well as the subsequent change in infants' attention biases. We tested a

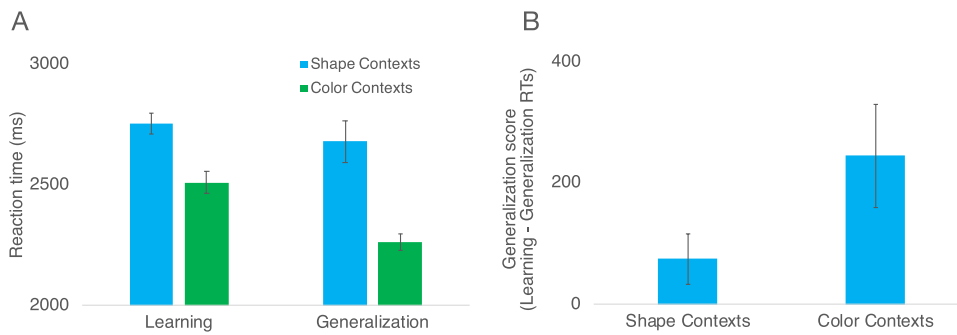


Fig. 3. (A) Infants' averaged eye movement reaction times to the cartoon reward during the learning and generalization phases of the abstract rule learning task in the Shape Contexts and Color Contexts conditions. Eye movement reaction times were calculated relative to trial onset (i.e., when the central cue first appeared on the screen). (B) Infants' generalization scores, which were calculated by subtracting their average eye movement reaction times to the cartoon reward during the generalization phase from their average reaction times during the learning phase of the abstract rule learning task.

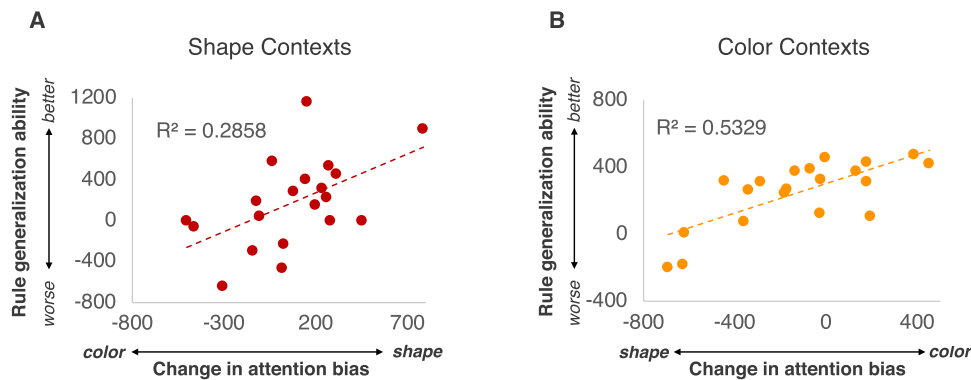


Fig. 4. The relation between infants' generalization performance and changes in attention biases to shape information in the Shape Contexts condition (A) and to color information in the Color Contexts Condition (B).

sample of infants using the same experimental paradigm as in Experiment 1, and we used functional near-infrared spectroscopy (fNIRS) to measure cortical activity and connectivity between infants' PFC and visual cortex 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. Examining individual differences in cortical activation and connectivity, and the relation to behavioral performance, provides an opportunity to glean insight into the *mechanisms* of developmental change in these emergent systems.

6. Experiment 2 method

6.1. Participants

The final sample consisted of 30 nine-month-old infants ($M = 9.5$ months, $SD = 0.48$ months, 16 females, 12 males, 23 white non-Hispanic, 4 Hispanic, 1 Asian, 1 black, and 1 other). An additional 6 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 ($n = 4$).

6.2. 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.

6.3. Procedure

The exact same experimental methods described in Experiment 1 for the Shape Contexts condition were used in Experiment 2, with the following exceptions: each of the two 32-s blocks of the rule learning task and the generalization task block were 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. We verified that infants' attention bias scores and average reaction times during the learning and generalization phase of the rule learning task were normally distributed through visual inspection and statistically using the Kolmogorov-Smirnov test for normality, all $ps > 0.090$.

6.4. fNIRS recording

Infants' frontal and visual cortical activity was recorded during the learning phase of the abstract rule 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 right/left 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 et al., 2007; Emberson et al., 2015; Werchan et al., 2018, 2016).

6.5. fNIRS data preprocessing

After recording, the fNIRS data were preprocessed prior to analyses using Homer 2.0 software (Huppert et al., 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 5 M 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 in the residual errors 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 and Asakawa, 2007; Werchan et al., 2016). This created a total of 7 time intervals for each of the two 32-s blocks during the learning phase of the abstract rule 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 Atlas-Viewer NIRS image reconstruction tools (Aasted et al., 2015). We verified that mean levels of cortical activation for each region were normally distributed through visual inspection and statistically using the Kolmogorov-Smirnov test for normality, all $ps > 0.130$.

6.6. fNIRS functional connectivity

To examine functional interactions between PFC and visual cortex, we first calculated an undirected measure of functional connectivity based on temporal correlations between regions of interest (Friston, 2011). Following the methods in prior fNIRS work exploring task-based functional connectivity in infants (Homaie et al., 2011; Keehn et al., 2013; Werchan et al., 2018), we computed a Pearson's r value for each infant by temporally correlating the PFC activations with the visual cortex activations across the seven averaged time intervals during the rule learning task. We then converted the r values to z scores using Fischer's z transformation to make their statistical distributions close to normal.

In addition to this undirected measure of functional connectivity, we also examined directed functional connectivity by applying Granger causality to the raw time courses of PFC and visual cortex activations. Granger causality is based on the concept that signal 1 affects or influences signal 2 if the predictions of signal 2 based on the past values of signal 1 are better than predictions of signal 2 based on past values of signal 2 alone (Barnett and Seth, 2014). For example, in the context of brain connectivity, if Granger causality for Region 1 \rightarrow Region 2 is greater than Region 2 \rightarrow Region 1, then this indicates a stronger functional influence from region 1 on region 2. Granger causality has been shown to be a suitable method for studying directional functional connectivity on cerebral blood oxygen response time courses (e.g., Goebel et al., 2003; Roebroeck et al., 2005; Seth et al., 2015; Wen et al., 2013).

We analyzed directed functional connectivity analyses by referring to prior fNIRS studies (e.g., Arizono et al., 2016; Medvedev, 2014; Sun and Wang, 2019; Zhou et al., 2016). Briefly, Granger causality values between the PFC and visual cortex were calculated for each infant using multivariate autoregressive (AR) modeling on the raw hemodynamic time courses, which was implemented using the MVGC MATLAB toolbox (Barnett and Seth, 2014). To meet assumptions of AR models, the raw time courses for the hemodynamic responses were differenced prior to analysis to remove non-stationarity in the signals. The optimal model

order was chosen using the Bayesian information criterion, model stationarity was checked by Augmented Dickey Fuller test, and model validity and consistency were verified using the Durbin-Watson test for autocorrelated residuals as implemented in the MVGC MATLAB toolbox (Barnett and Seth, 2014).

7. Experiment 2 results

7.1. Behavioral results

We first examined the distribution of infants' baseline attention biases, which again indicated that infants had a greater initial attention bias to color over shape, $t(29) = 3.703$, $p = 0.001$, $M = 183.12$, $SD = 270.83$. We then examined infants' eye movement reaction times across learning trial bins, which revealed a main effect of Learning Trial Bin, $F(3,87) = 2.698$, $p = 0.051$, indicating that infants' were learning the stimulus-response associations with trial exposure.

We next examined whether infants generalized the stimulus-response associations to novel shape contexts, as indexed by faster reaction times during learning relative to generalization. We conducted a one-sample t test comparing infants' generalization scores (calculated as a difference score of average reactions times during Learning – Generalization) to chance (0), which revealed that infants' eye movement reaction times were significantly faster during generalization relative to learning, $t(29) = 4.749$, $p < 0.001$ (Fig. 5). This provides evidence that infants learned an abstract rule and generalized this rule to a novel shape context. Critically, we also found that individual differences in infants' generalization scores were correlated with the change in infants' attention bias to shape from baseline to post-test, $r(30) = 0.522$, $p = 0.003$ (Fig. 6). These results are consistent with the findings from Experiment 1 in a larger independent sample of infants. Moreover, infants' generalization performance was more robust than infants assigned to the Shape Contexts condition in Experiment 1.

7.2. fNIRS results

We next examined how infants' cortical activity during the two 32-s blocks of the learning phase of the abstract rule learning task related to individual differences in infants' behavioral generalization scores and the change in infants' attention bias to shape, measured by relative looking time. We first conducted an omnibus repeated-measures ANCOVA on infants' cortical activations using Region (Visual Cortex, Left PFC, Right PFC), Learning Block (Block 1, Block 2), and Time Interval (seven 4-s intervals) as within-subjects factors and infants' Generalization scores and Attention Bias Change scores as continuous

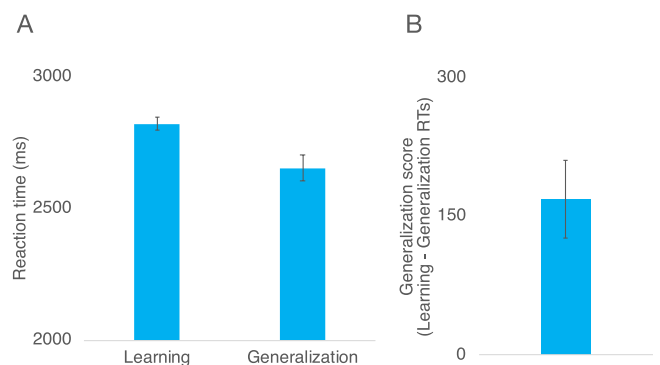


Fig. 5. (A) Infants' average eye movement reaction times to the correct location during the learning and generalization phases of the abstract rule learning task in Experiment 2. (B) Infants' generalization scores, which were calculated by subtracting their average eye movement reaction times to the cartoon reward during the generalization phase from their average reaction times during the learning phase of the abstract rule learning task.

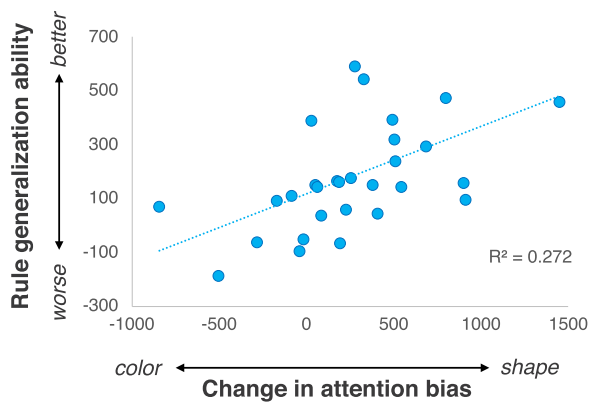


Fig. 6. Relation between infants' generalization performance during the rule learning task and change in attention bias to shape in Experiment 2.

variables. This analysis revealed a significant Region \times Learning Block \times Time Interval \times Generalization \times Attention Bias Change interaction, Wilks' lambda = 0.296, $F(12,15) = 2.975$, $p = 0.025$. There were no other significant main effects or interactions, all $ps > 0.143$, all $Fs < 1.787$.

We followed up on this significant interaction by examining each of the two 32-s learning blocks separately (Bonferroni corrected alpha = .025). A repeated-measures ANCOVA using Region (Visual Cortex, Left PFC, Right PFC) and Time Interval (seven 4-s intervals) as within-subjects factors and infants' Generalization scores and Attention Bias Change scores as continuous variables showed no significant main effects or interactions in the second 32-s block of the learning task, all $ps > 0.092$, all $Fs < 2.074$. This same test on the first 32-s block of the learning task, however, revealed a significant Region \times Attention Bias Change interaction, Wilks' lambda = 0.582, $F(2,25) = 8.992$, $p = 0.001$, and a significant Region \times Time \times Generalization \times Attention Bias Change interaction, Wilks' lambda = 0.294, $F(12,15) = 3.008$, $p = 0.024$. Planned Helmert contrasts indicated that the visual cortex differed from the right and left PFC in the Region \times Attention Bias Change interaction, $F(1,26) = 12.848$, $p = 0.001$, and in the Region \times Generalization \times Attention Bias Change interaction, $F(1,26) = 4.695$, $p = 0.040$. Additionally, the right PFC differed from the left PFC in the Region \times Generalization interaction, $F(1,26) = 4.945$, $p = 0.035$. No other significant effects or interactions were found, all $ps > 0.141$, all $Fs < 2.301$. We followed up on these interactions by examining activations across Time Interval separately by region.

7.2.1. Visual cortex activation

A repeated-measures ANCOVA on infants' visual cortex activations, including infants' Attention Bias Change scores and Generalization scores as continuous variables, revealed a significant Time \times Attention Bias Change interaction, Wilks' lambda = 0.385, $F(6,21) = 5.587$, $p = 0.001$. This interaction indicates that lower levels of mean visual cortex activation during the first 32-s block of learning were related to a greater change in infants' attention biases to shape after the rule learning task relative to baseline, $r(30) = -0.398$, $p = 0.029$. However, visual cortex activation was not correlated with infants' generalization scores during the rule learning task, $r(30) = 0.095$, $p = 0.618$.

7.2.2. PFC activation

A repeated-measures ANCOVA on infants' Left PFC activations, including infants' Generalization scores as continuous variables, did not reveal a significant interaction between Left PFC activation and Generalization, Wilks' lambda = 0.880, $F(6,21) = 0.476$, $p = 0.819$. However, this same analysis on infants' Right PFC revealed a significant interaction between Right PFC activation and Generalization, Wilks' lambda = 0.557, $F(6,21) = 2.786$, $p = 0.037$. This interaction indicates that higher levels of mean right PFC activation during the first 32-s block of learning

was related to better subsequent generalization performance, $r(30) = 0.480$, $p = 0.007$. However, mean levels of PFC activation did not relate to the change in infants' attention biases after learning relative to baseline, $r(30) = 0.191$, $p = 0.312$.

7.2.3. Functional connectivity

Finally, we examined how functional connectivity between the right PFC and visual cortex during the first 32-s block of learning related to individual differences in subsequent generalization performance, as well as the change in infants' attention biases to shape. A one-sample t test indicated that infants' functional connectivity correlational values were significantly greater than 0, $t(29) = 2.115$, $p = 0.043$. We then analyzed how individual differences in infants' functional connectivity correlational values related to subsequent behavioral Generalization scores and Attention Bias Change scores. Results indicated that higher functional connectivity was related to better generalization performance, $r(30) = 0.456$, $p = 0.011$, but did not correlate with the change in infants' attention bias to shape, $r(30) = 0.200$, $p = 0.290$.

We followed up on these analyses by examining the directionality of functional connectivity between the right PFC and visual cortex, indexed by Granger causality (GC) values for right PFC \rightarrow Visual Cortex and for Visual Cortex \rightarrow right PFC. A paired-samples two-tailed t test indicated that, at a group level, GC values for right PFC \rightarrow Visual Cortex ($M = 0.018$, $SD = 0.010$) were not different than GC values for Visual Cortex \rightarrow right PFC ($M = 0.021$, $SD = 0.011$), $t(29) = 1.360$, $p = 0.184$.

We next examined how individual differences in right PFC \rightarrow Visual Cortex and Visual Cortex \rightarrow right PFC GC values related to individual differences in subsequent behavioral generalization scores. We found that GC values for Visual Cortex \rightarrow right PFC did not correlate with infants' generalization performance, $r(30) = -0.239$, $p = 0.203$. Critically, however, our results revealed that higher GC values for right PFC \rightarrow Visual Cortex were correlated with better subsequent generalization performance, $r(30) = 0.415$, $p = .022$ (Fig. 7; Bonferroni corrected alpha = .025). These results suggest that the extent of the right PFC's functional influence over visual cortex during abstract rule learning is related to subsequent behavioral generalization performance. Taken together, these findings add 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 abstract rule learning.

8. General discussion

Visual attention is a fundamental capacity that supports the flexible selection of information based on relevant rules and goals that guide 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 and Li, 2013; Baluch and Itti, 2011; Paneri and Gregoriou, 2017). Yet, it is unclear whether similar mechanisms operate

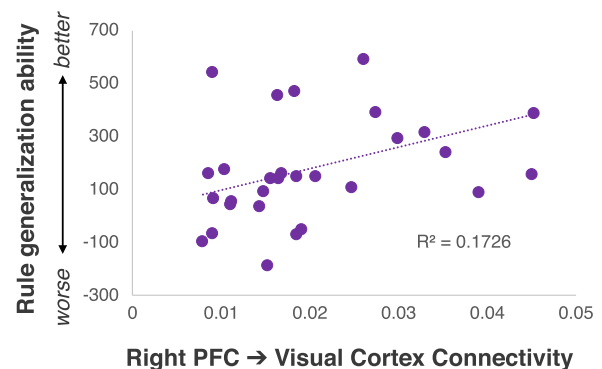


Fig. 7. Relation between infants' directed functional connectivity from the right PFC to visual cortex during learning and subsequent generalization performance.

in infants. We recently showed that 8-month-old infants can use the PFC to structure visual inputs into abstract rules (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 and Amso, 2018). Across two experiments, our behavioral data here provide support that 9-month-old infants can use top-down knowledge rapidly acquired through abstract rule learning mechanisms to modulate visual attention of learned behaviorally-relevant visual features.

We presented infants with an abstract rule learning task where they could use a simple visual feature (either color or shape) as a higher-order context to organize visual inputs into rules for learning. Our results indicated that infants who showed better generalization performance, as measured by faster reaction times during the generalization phase relative to the learning phase of the abstract rule learning task, also showed a greater change in their attention biases, from baseline to post-learning, to the visual feature that cued these abstract rule structures. Additionally, we found that infants showed similar learning rates and changes in attention biases to learned behaviorally-relevant visual features irrespective of condition assignment. However, it was unclear from this sample whether infants assigned to the Shape Contexts condition generalized rule learning to a novel shape context. We thus repeated the Shape Contexts condition with a larger sample of infants in Experiment 2, where we observed similar generalization performance as infants assigned to the Color Contexts condition. Thus, these results provide behavioral evidence that infants learned and generalized an abstract rule, which subsequently influenced infants' visual selection and attentional biases towards information relevant for learning.

We also examined the neural underpinnings of these processes by using fNIRS to record infants' frontal and visual cortex activity during learning. We found that infants' who had lower mean levels of visual cortex activation during the first 32-second block of the rule learning task also had greater subsequent changes in their attention biases to shape. This is consistent with prior work implicating reduced visual cortex activation over the course of perceptual learning, likely due to sharpened tuning of neuronal representations (e.g., Mukai et al., 2007). Moreover, we saw that infants with higher mean levels of right PFC activation during the first 32-second block of the rule learning task demonstrated better subsequent generalization performance. These findings mirror prior work implicating increased right dorsolateral PFC activation due to working memory processes in similar rule learning tasks in both infants (Werchan et al., 2016) and adults (Collins et al., 2014). We also found that these effects were specific to the first 32-second block of the rule learning task, which is consistent with prior work indicating that PFC involvement is more pronounced in early stages of learning, which can consist of as few as 8 trials, relative to late stages of learning (see Kelly and Garavan, 2005, for a review). It is also consistent with prior work indicating that learning and transfer of abstract rules stabilizes during the early trials of a task (e.g., Bhandari and Badre, 2018; Bhandari and Duncan, 2014; Cole et al., 2011).

We interpret these findings as preliminary evidence that greater PFC influence over visual cortex during initial rule learning might support better learning and subsequent generalization of abstract rules organizing visual inputs into predictable sequences. In turn, better learning and generalization of abstract rules may then lead to a greater attention bias to the relevant feature that acts as a higher-order context cueing these abstract rule structures. In support of this interpretation, our results also revealed that individual differences in infants' functional connectivity between the right PFC and visual cortex during the first 32-s block of learning was correlated with the efficacy of subsequent generalization performance, such that infants with higher functional connectivity showed better generalization. Moreover, Granger causality analyses assessing the directionality of these functional interactions indicated that greater PFC influence on visual cortex was related to better generalization performance. These results are consistent with prior work implicating functional interactions between the PFC and

visual cortex in modulating visual processing in adults (Corbetta and Shulman, 2002; Gazzaley et al., 2007; Kastner and Ungerleider, 2000; Morishima et al., 2009; Rossi et al., 2009; Taylor et al., 2007).

Prior work shows that resting state functional connectivity in human infants is dominated by short-range intra-cortical connections relative to long-range intercortical connections (Fransson et al., 2011; Gao et al., 2011), which makes our fNIRS results somewhat surprising. 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 and Amso, 2018). Thus, our behavioral and neuroimaging findings provide preliminary evidence that corticocortical connections between the PFC and visual cortex might be involved in top-down guidance of visual attention in infants. These findings add to growing behavioral evidence showing that infants can use top-down knowledge to guide visual attention, as well as add new insights into the functional circuitry that support 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 attention in infants, it is a correlational measure that cannot be used to make statements of causality.

In sum, our findings show that infants can rapidly acquire top-down knowledge using abstract rule learning mechanisms, and that this top-down knowledge subsequently influences infants' visual selection and attention biases to visual features relevant for learning. This initial demonstration of top-down knowledge influencing visual attention in infants may help infants learn to 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.

Author contributions

D. M. Werchan and D. Amso developed the study concept and design. D. M. Werchan collected the data. D. M. Werchan and D. Amso analyzed the data and drafted the manuscript.

Declaration of Competing Interest

The authors report no conflicts of interest.

Acknowledgements

We thank members of the Developmental Cognitive Neuroscience Lab at Brown University for help with recruitment, data collection, and analysis, and all of the infants and families who made this research possible. This work was funded by the James S. McDonnell Scholar Award in Understanding Human Cognition (to DA), and by the Robin Chemers Neustein Graduate Award in Brain Science (to DMW).

References

- Aasted, C.M., Yücel, M.A., Cooper, R.J., Dubb, J., Tsuzuki, D., Becerra, L., et al., 2015. Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. *Neurophotonics* 2 (2). <https://doi.org/10.1117/1.NPh.2.2.020801>.
- Althaus, N., Mareschal, D., 2012. Using saliency maps to separate competing processes in infant visual cognition. *Child Dev.* 83 (4), 1122–1128.
- Amso, D., Johnson, S.P., 2005. Selection and inhibition in infancy: evidence from the spatial negative priming paradigm. *Cognition* 95, B27–B36. <https://doi.org/10.1016/j.cognition.2004.08.006>.

- Amso, D., Johnson, S.P., 2006. Learning by selection: visual search and object perception in young infants. *Dev. Psychol.* 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* 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>.
- Arizono, N., Ohmura, Y., Yano, S., Kondo, T., 2016. Functional connectivity analysis of NIRS data under rubber hand illusion to find a biomarker of sense of ownership. *Neural Plast.* 2016 <https://doi.org/10.1155/2016/6726238>.
- Badre, D., 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci. (Regul. Ed.)* 12 (5), 193–200. <https://doi.org/10.1016/j.tics.2008.02.004>.
- Baluch, F., Itti, L., 2011. Mechanisms of top-down attention. *Trends Neurosci.* 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 Res. Bull.* 52 (5), 319–330.
- Barceló, F., Suwazono, S., Knight, R.T., 2000. Prefrontal modulation of visual processing in humans. *Nat. Neurosci.* 3 (4), 399–403. <https://doi.org/10.1038/73975>.
- Barnett, L., Seth, A.K., 2014. The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *J. Neurosci. Methods* 223, 50–68. <https://doi.org/10.1016/j.jneumeth.2013.10.018>.
- Bhandari, A., Badre, D., 2018. Learning and transfer of working memory gating policies. *Cognition* 172 (December), 89–100. <https://doi.org/10.1016/j.cognition.2017.12.001>.
- Bhandari, A., Duncan, J., 2014. Goal neglect and knowledge chunking in the construction of novel behaviour. *Cognition* 130 (1), 11–30. <https://doi.org/10.1016/j.cognition.2013.08.013>.
- 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>.
- 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. *J. Neurosci.* 29 (18), 5863–5872. <https://doi.org/10.1523/JNEUROSCI.0539-09.2009>.
- Carrasco, M., 2011. Visual attention: the past 25 years. *Vision Res.* 51 (13), 1484–1525. <https://doi.org/10.1016/j.visres.2011.04.012>.
- Cheng, C., Kaldy, Z., Blaser, E., 2019. Focused attention predicts visual working memory performance in 13- month-old infants: a pupillometric study. *Dev. Cogn. Neurosci.* 36 <https://doi.org/10.1016/j.dcn.2019.100616>.
- Cohen, J.D., Braver, T.S., Brown, J.W., 2002. Computational perspectives on dopamine function in prefrontal cortex. *Curr. Opin. Neurobiol.* 12 (2), 223–229. Retrieved from: <http://www.ncbi.nlm.nih.gov/pubmed/12015241>.
- Cole, M.W., Etzel, J.A., Zacks, J.M., Schneider, W., Braver, T.S., 2011. Rapid transfer of abstract rules to novel contexts in human lateral prefrontal cortex. *Front. Hum. Neurosci.* 5, 1–13. <https://doi.org/10.3389/fnhum.2011.00142>.
- Collins, A.G.E., Cavanagh, J.F., Frank, M.J., 2014. Human EEG uncovers latent generalizable rule structure during learning. *J. Neurosci.* 34 (13), 4677–4685. <https://doi.org/10.1523/JNEUROSCI.3900-13.2014>.
- Collins, A.G., Frank, M.J., 2013. Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychol. Review* 120, 190–229.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. <https://doi.org/10.1038/nrn755>.
- Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
- Donoso, M., Collins, A.G., Koechlin, E., 2014. Foundations of human reasoning in the prefrontal cortex. *Science* 344, 1481–1486.
- 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. *Proc. Natl. Acad. Sci.* 112 (31), 9585–9590. <https://doi.org/10.1073/pnas.1510343112>.
- Frank, M.J., Badre, D., 2012. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits I: computational analysis. *Cereb. Cortex* 22, 509–526. <https://doi.org/10.1111/j.1467-7687.2008.00794.x>.
- Frank, M.C., Slemmer, J.A., Marcus, G.F., Johnson, S.P., 2009a. Information from multiple modalities helps 5-month-olds learn abstract rules. *Dev. Sci.* 4 (12), 504–509. <https://doi.org/10.1111/j.1467-7687.2008.00794.x>.
- Frank, M.C., Vul, E., Johnson, S.P., 2009b. 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>.
- Fransson, P., Skjold, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., Aden, U., 2007. Resting-state networks in the infant brain. *Proc. Natl. Acad. Sci. U.S.A.* 104 (39), 15531–15536. <https://doi.org/10.1073/pnas.0704380104>.
- Fransson, P., Aden, U., Blennow, M., Lagercrantz, H., 2011. The functional architecture of the infant brain as revealed by resting-state fMRI. *Cereb. Cortex* 21 (1), 145–154. <https://doi.org/10.1093/cercor/bhq071>.
- Friston, K.J., 2011. Functional and effective connectivity: a review. *Brain Connect.* 1 (1), 13–88. <https://doi.org/10.1089/brain.2011.0008>.
- 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. *Cereb. Cortex* 17, 125–135. <https://doi.org/10.1093/cercor/bhm113>.
- Gilbert, C.D., Li, W., 2013. Top-down influences on visual processing. *Nat. Rev. Neurosci.* 14 (5), 350–363. <https://doi.org/10.1038/nrn3476>.
- Goebel, R., Roebroeck, A., Kim, D., Formisano, E., 2003. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn. Reson. Imaging* 21, 1251–1261. <https://doi.org/10.1016/j.mri.2003.08.026>.
- Goldman-Rakic, P.S., 1987. Development of cortical circuitry and cognitive function. *Child Dev.* 58 (3), 601–622. <https://doi.org/10.1111/j.0963-7214.2006.00419.x>. Retrieved from.
- Homae, F., Watanabe, H., Otobe, T., Nakano, T., Go, T., Konishi, Y., Taga, G., 2010. Development of global cortical networks in early infancy. *J. Neurosci.* 30 (14), 4877–4882. <https://doi.org/10.1523/JNEUROSCI.5618-09.2010>.
- Homae, F., Watanabe, H., Nakano, T., Taga, G., 2011. Large-scale brain networks underlying language acquisition in early infancy. *Front. Psychol.* 2, 1–14. <https://doi.org/10.3389/fpsyg.2011.00093>.
- 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. *Appl. Opt.* 48 (10), D280–D298.
- Johnson, M.H., Vecera, S.P., 1996. Cortical differentiation and neurocognitive development: the parcellation conjecture. *Behav. Processes* 36 (2), 195–212. [https://doi.org/10.1016/0376-6357\(95\)00028-3](https://doi.org/10.1016/0376-6357(95)00028-3).
- Johnson, M.H., Posner, M.I., Rothbart, M.K., 1991. Components of visual orienting in early infancy: contingency learning, anticipatory looking, and disengaging. *J. Cogn. Neurosci.* 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. *Psychol. Sci.* 5 (2), 90–93.
- Johnson, S.P., Amso, D., Slemmer, J.A., 2003. Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proc. Natl. Acad. Sci. U.S.A.* 100 (18).
- Kastner, S., Ungerleider, L.G., 2000. Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* 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. *Front. Hum. Neurosci.* 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. *Cereb. Cortex* 15, 1089–1102. <https://doi.org/10.1093/cercor/bhi005>.
- Kirkham, N.Z., Slemmer, J.A., Richardson, D.C., Johnson, S.P., 2007. Location, location, location: development of spatiotemporal sequence learning in infancy. *Child Dev.* 78 (5), 1559–1571.
- Kolb, B., Mychasiuk, R., Muhammad, A., Li, Y., Frost, D.O., Gibb, R., 2012. Experience and the developing prefrontal cortex. *Proc. Natl. Acad. Sci.* 109 (Supplement_2), 17186–17193. <https://doi.org/10.1073/pnas.1121251109>.
- 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 Dev.* 80 (4), 986–999.
- Marcus, G.F., Vijayan, S., Bandi Rao, S., Vishton, P.M., 1999. Rule learning by seven-month-old infants. *Science* 283 (January), 77–81.
- Marcus, G.F., Fernandes, K.J., Johnson, S.P., 2007. Infant rule learning facilitated by speech. *Psychol. Sci.* 18 (5), 387–391.
- Markant, J., Amso, D., 2013a. Selective memories: infants' encoding is enhanced in selection via suppression. *Dev. Sci.* 16 (6), 926–940. <https://doi.org/10.1111/desc.12084>.
- Markant, J., Amso, D., 2013b. Selective memories: infants' encoding is enhanced in selection via suppression. *Dev. Sci.* 16 (6), 926–940. <https://doi.org/10.1111/desc.12084>.
- 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. *Dev. Psychobiol.* 355–365. <https://doi.org/10.1002/dev.21375>.
- Medvedev, A.V., 2014. Does the resting state connectivity have hemispheric asymmetry? A near-infrared spectroscopy study. *NeuroImage* 85 (0 1), 1–16. <https://doi.org/10.1016/j.neuroimage.2013.05.092>.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24 (1), 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>.
- Morishima, Y., Akaiishi, R., Yamada, Y., Okuda, J., Toma, K., Sakai, K., 2009. Task-specific signal transmission from prefrontal cortex in visual selective attention. *Nat. Neurosci.* 12 (1), 85–91. <https://doi.org/10.1038/nn.2237>.
- Mukai, I., Kim, D., Fukunaga, M., Japee, S., Marrett, S., Ungerleider, L.G., 2007. Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *J. Neurosci.* 27, 11401–11411.
- Noudouost, B., Chang, M.H., Steinmetz, N.A., Moore, T., 2010. Top-down control of visual attention. *Curr. Opin. Neurobiol.* 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: Wixted, J.T. (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, fourth ed. John Wiley & Sons, pp. 1–3. <https://doi.org/10.1002/9781119170174.epcn401>.

- Paneri, S., Gregoriou, G.G., 2017. Top-down control of visual attention by the prefrontal cortex. Functional specialization and long-range interactions. *Front. Neurosci.* 11, 1–16. <https://doi.org/10.3389/fnins.2017.00545>.
- 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. *Eur. J. Neurosci.* 16, 291–310. <https://doi.org/10.1046/j.1460-9568.2002.02090.x>.
- Posner, M.I., 1980. Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.
- Richards, J.E., 2000. Localizing the development of covert attention in infants with scalp event-related potentials. *Dev. Psychol.* 36 (1), 91–108. <https://doi.org/10.1037/0012-1649.36.1.91>.
- Roebroeck, A., Formisano, E., Goebel, R., 2005. Mapping directed influence over the brain using Granger causality and fMRI. *NeuroImage* 25, 230–242. <https://doi.org/10.1016/j.neuroimage.2004.11.017>.
- Rossi, A.F., Pessoa, L., Desimone, R., Ungerleider, L.G., 2009. The prefrontal cortex and the executive control of attention. *Exp. Brain Res.* 192, 489–497. <https://doi.org/10.1007/s00221-008-1642-z>.
- Ross-Sheehy, S., Oakes, L.M., Luck, S.J., 2011. Exogenous attention influences visual short-term memory in infants. *Dev. Sci.* 14 (3), 490–501. <https://doi.org/10.1111/j.1467-7687.2010.00992.x>.
- 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. *Proc. Natl. Acad. Sci. U. S. A.* 102 (20), 7338–7343. <https://doi.org/10.1073/pnas.0502455102>.
- Ruff, C.C., Bestmann, S., Blankenburg, F., Bjoertomt, O., Josephs, O., Weiskopf, N., et al., 2008. Distinct causal influences of parietal versus frontal areas on human visual cortex: evidence from concurrent TMS-fMRI. *Cereb. Cortex* 18, 817–827. <https://doi.org/10.1093/cercor/bhm128>.
- Seth, A.K., Barrett, A.B., Barnett, L., 2015. Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* 35 (8), 3293–3297. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>.
- 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>.
- Sun, J., Wang, H., 2019. Connectivity properties in the prefrontal cortex during working memory: a near-infrared spectroscopy study cortex during working memory: a near-infrared. *J. Biomed. Opt.* 24 (5) <https://doi.org/10.1117/1.JBO.24.5.051410>.
- 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, P.C.J., Nobre, A.C., Rushworth, M.F.S., Hospital, J.R., Way, H., Ox, O., 2007. FEF TMS affects visual cortical activity. *Cereb. Cortex* 17, 391–399. <https://doi.org/10.1093/cercor/bhj156>.
- Tipper, S.P., 1985. The negative priming effect: inhibitory priming by ignored objects. *Q. J. Exp. Psychol.* 37A, 571–590.
- Tummeltshammer, K.S., Amso, D., 2018. Top-down contextual knowledge guides visual attention in infancy. *Dev. Sci.* 21, 1–9. <https://doi.org/10.1111/desc.12599>.
- Tummeltshammer, K.S., Kirkham, N.Z., 2013. Learning to look: probabilistic variation and noise guide infants' eye movements. *Dev. Sci.* 16 (5), 760–771. <https://doi.org/10.1111/desc.12064>.
- Tummeltshammer, K.S., Mareschal, D., Kirkham, N.Z., 2014. Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Dev.* 85 (5), 1981–1994. <https://doi.org/10.1111/cdev.12239>.
- Ungerleider, L.G., Gaffan, D., Pelak, V.S., 1989. Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys. *Exp. Brain Res.* 76, 473–484.
- Webster, M.J., Bachevalier, J., Ungerleider, L.G., 1994. Connections of inferior temporal areas TE0 and TE with parietal and frontal cortex in macaque monkeys. *Cereb. Cortex* 4 (5), 470–483.
- Wen, X., Rangarajan, G., Ding, M., 2013. Is Granger Causality a Viable Technique for Analyzing fMRI Data? *PLoS One* 8 (7). <https://doi.org/10.1371/journal.pone.0067428>.
- Werchan, D.M., Collins, A.G.E., Frank, M.J., Amso, D., 2015. 8-month-Old infants spontaneously learn and generalize hierarchical rules. *Psychol. Sci.* 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. *J. Neurosci.* 36 (40), 10314–10322. <https://doi.org/10.1523/JNEUROSCI.1351-16.2016>.
- Werchan, D.M., Baumgartner, H.A., Lewkowicz, D.J., Amso, D., 2018. The origins of cortical multisensory dynamics: evidence from human infants. *Dev. Cogn. Neurosci.* 34, 75–81. <https://doi.org/10.1016/j.dcn.2018.07.002>.
- Werchan, D.M., Lynn, A., 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* 1–16. <https://doi.org/10.1111/inf.12309>.
- 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>.
- Wu, R., Kirkham, N.Z., 2010. No two cues are alike: depth of learning during infancy is dependent on what orients attention. *J. Exp. Child Psychol.* 107, 118–136. <https://doi.org/10.1016/j.jecp.2010.04.014>.
- 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. *Nat. Neurosci.* 14 (5), 656–661. <https://doi.org/10.1038/nn.2773>.
- Zhou, G., Liu, J., Ding, X.P., Fu, G., Lee, K., Lee, K., 2016. Development of Effective Connectivity during Own- and Other-Race Face Processing: A Granger Causality Analysis. *Front. Hum. Neurosci.* 10, 1–14. <https://doi.org/10.3389/fnhum.2016.00474>.