Behavioral/Cognitive

Role of Prefrontal Cortex in Learning and Generalizing Hierarchical Rules in 8-Month-Old Infants

Denise M. Werchan,¹ Anne G.E. Collins,² Michael J. Frank,¹ and Dima Amso¹

¹Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, Rhode Island 02912, and ²Department of Psychology, University of California, Berkeley, California 94720

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 to record PFC activity in 8-month-old human infants during a simple audiovisual hierarchical-rulelearning task. Behavioral results confirmed that infants adopted hierarchical rule sets to learn and generalize spoken object–label mappings across different speaker contexts. 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 dorsolateral PFC 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.

Key words: development; eye blink rate; frontostriatal circuitry; generalization; prefrontal cortex; rule learning

Significance Statement

Hierarchical rule learning is a powerful learning mechanism that allows rules to be selected in a context-appropriate fashion and transferred or reused in novel contexts. Data from computational models and adults suggests that this learning mechanism is supported by dopamine-innervated interactions between prefrontal cortex (PFC) and striatum. Here, we provide evidence that PFC also supports hierarchical rule learning during infancy, challenging the current dogma that PFC is an underdeveloped brain system until adolescence. These results add new insights into the neurobiological mechanisms available to support learning and generalization in very early postnatal life, providing evidence that PFC and the frontostriatal circuitry are involved in organizing learning and behavior earlier in life than previously known.

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 (Frank and Badre, 2012; Collins and

Received April 24, 2016; revised Aug. 18, 2016; accepted Aug. 18, 2016.

The authors declare no competing financial interests.

Correspondence should be addressed to Dima Amso, Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer St., Providence, RI 02912. E-mail: Dima_Amso@Brown.edu. DOI:10.1523/JNEUROSCI.1351-16.2016

Copyright © 2016 the authors 0270-6474/16/3610314-09\$15.00/0

Frank, 2013; Collins et al., 2014; Donoso et al., 2014; Collins and Frank, 2016). 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; Fig. 1).

Prior behavioral work suggests that adults spontaneously and incidentally create hierarchical rules when learning simple stimulusaction 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 (Frank and Badre, 2012; Collins and Frank, 2013). Specifically, these models posit that frontostriatal circuits are hierarchically nested, such that a higher-order context cues a valid rule set to up-

Author contributions: D.M.W., A.G.E.C., M.J.F., and D.A. designed research; D.M.W. performed research; D.M.W., A.G.E.C., M.J.F., and D.A. analyzed data; D.M.W., M.J.F., and D.A. wrote the paper.

This work was supported by the National Science Foundation (Graduate Research Fellowship under Grant DGE-1058262 to D.M.W.) and the National Institutes of Health (Grant R01 MH099078 D.A. and D. Badre). We thank TechEn CW6 Systems for their generous and tireless support and the infants and families who made this research possible.



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 stayed the same 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.

date 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-monthold 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 (Karson, 1983), also increased when higherorder rules switched during learning. However, PFC has a protracted developmental course, not reaching maturity until adolescence (Huttenlocher, 1979). Therefore, 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, because the striatum shows relatively high functionality compared with PFC in early infancy (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 (Fig. 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 with 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 (Fig. 1). Importantly, we expected that 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).

Materials and Methods

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 before participation. All families were compensated for time and travel to our laboratory.

Procedure

Behavioral procedures

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 (Fig. 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 among 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 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 in which the higher-order stimulus (e.g., face–voice speaker con-



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 dIPFC. *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 dIPFC.

texts) 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 in which the higher-order rule stayed the same on each trial (Stay 1 and Stay 2 conditions; Fig. 1*B*). We recorded infants' frontal cortical activity across left and right dorsolateral PFC (dlPFC) and medial PFC (mPFC) during the learning task using NIRS (Fig. 2*A*,*B*).

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 objectlabel 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; Fig. 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 toyword pairing from the generalization task. The purpose of this inference test was to determine 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; Fig. 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.

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 among 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 eight trials during learning, for a total of 32 trials (eight 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 1500 ms while a recorded female voice said the artificial word associated with the pairing. There was a 1000 ms interval between trials. The four blocks consisted of two rule switch condi-

tions (Switch 1 and Switch 2) in which the higher-order face–voice context changed each trial and two rule stay conditions (Stay 1 and Stay 2) in which the higher-order face–voice context stayed the same each trial. Each block was preceded by a 10 s 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 s is sufficient for blood volume to return to baseline levels (Wilcox et al., 2008, 2009). The order of blocks was counterbalanced between participants.

Generalization task. Immediately after 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. In addition, one novel toy–word association was added to the rule set (Fig. 1A). Infants again received eight trials per pairing, as in the learning task, for a total of 24 trials.

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 (Fig. 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 >2 s, up to a maximum of 60 s. The dependent behavioral measure was the average looking time during the consistent trials compared with 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). Interrater reliability was high (r = 0.97).

NIRS recording procedures

Infants' frontal 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 12 optodes (4 sources and 8 detectors, resulting in 10 source-detector channels) with an interoptode separation of 3 cm was placed over infants' frontal brain regions on each hemisphere. The array was arranged in a lattice pattern (Fig. 2*A*), 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



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 indicate SEM.

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 (Fig. 2*A*; Jasper, 1958). The vertical midline of the optode array was centered across the nasion–inion line.

After recording, the data were preprocessed in HomER 2.0 using the default preprocessing pipeline (Huppert et al., 2009). First, the raw signals (acquired at 50 Hz) were digitally band-pass filtered at 0.01-0.1 Hz to remove systematic physiological and movement artifacts (White et al., 2009; Homae et al., 2010). Second, the change in optical density was calculated for each wavelength relative to the 10 s baseline before 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 $\geq \pm 5 \ \mu$ M over a 0.5 s range in each channel (Lloyd-Fox et al., 2009; Emberson et al., 2015). This indicated that no motion artifacts occurred during the four blocks of interest (Stay 1, Stay 2, Switch 1, and Switch 2), so all blocks were retained for subsequent data analysis. Finally, changes in oxygenated hemoglobin (relative to the 10 s baseline) in each of the 10 source-detector channels were exported for subsequent analysis by averaging across every 4 s of each 24 s block starting 4 s after block onset to account for the hemodynamic response lag (Miezin et al., 2000). This created a total of five 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 dlPFC, the three center-left channels corresponding to right mPFC, the three center-right channels corresponding to right dlPFC (Fig. 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 (Fig. 2*B*; Aasted et al., 2015).

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 (Karson, 1983; Shukla, 1985; Blin et al., 1990; Kleven and Koek, 1996; Taylor et al., 1999; e.g., Colzato et al., 2009). 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 nonhuman primates (Karson, 1983; Elsworth et al., 1991; Kleven and Koek, 1996; Jutkiewicz and Bergman, 2004). 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 (Deuschl and Goddemeier, 1988; Bodfish et al., 1995; Blandini et al., 2000) 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 with 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 (Tada, 1978; Fukuda et al., 2005; Oh et al., 2012). Blink rate is also predictive of the effects of dopamine agonists on reinforcement learning (Cavanagh et al., 2014). Although 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 (for review, see Bacher and Smotherman, 2004). 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 the involvement of dopamineinnervated frontostriatal circuitry during learning (note that we do not have the ability to image subcortical areas directly 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 dopamineinnervated 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.

Results

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 (Fig. 3*A*). 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 = 0.05), replicating our earlier finding that infants created hierarchical rule sets and



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 indicate SEM.

used these sets to generalize toy–word mappings across speaker contexts (Fig. 3*B*).

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 withinsubject factors. This test revealed a block by condition interaction ($F_{(1,36)} = 4.084$, p = 0.051). Subsequent within- and betweencondition 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 = 0.004, Bonferroni corrected α set to 0.008; Fig. 4).

NIRS results

We examined differences in infants' cortical activity, as measured by changes in oxygenated hemoglobin (HbO₂), 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; Fig. 2A, B) and assessed whether this neural activity varied as a function of whether infants successfully generalized learned rules. Changes in HbO₂ responses were calculated by averaging the HbO₂ response across every 4 s of each 24 s block (starting 4 s after block onset to account for the hemodynamic response lag; Miezin et al., 2000). This created a total of five time intervals for each block. Changes in HbO₂ responses were measured relative to a 10 s baseline period before 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 HbO2 response across the five time intervals for each condition), region of interest (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 among time interval, region, and condition $(F_{(36,1260)} = 1.686, p = 0.007, \eta_p^2 = 0.046)$. To determine which of the four regions of interest exhibited significant differences in activity in this three-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 (α set to a corrected 0.0125 value) revealed significant interactions between time interval and condition for right dlPFC only ($F_{(12,420)} = 2.219$, p = 0.01, $\eta_p^2 = 0.060$). This interaction was not significant for left dlPFC ($F_{(12,420)} = 0.798$, p = 0.653, $\eta_p^2 = 0.022$), left mPFC ($F_{(12,420)} = 0.601$, p = 0.842, $\eta_p^2 = 0.017$), or right mPFC ($F_{(12,420)} = 0.575$, p = 0.862, $\eta_p^2 = 0.016$).

We next examined right dlPFC activity separately for all combinations of conditions from the previous interactions (Switch 1 vs Stay 1, Switch 1 vs Switch 2, Stay 1 vs Stay 2, Stay 2 vs Switch 2, Switch 1 vs Stay 2, and Stay 1 vs 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 α value set to 0.008 (0.05/6). These analyses showed a significant 3-way interaction among interval, condition, and learning score for right dlPFC in Stay 2 versus Switch 2 only ($F_{(4,140)} = 5.328$, p = 0.001, $\eta_p^2 = 0.132$; 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 (Figs. 5, 6A). To verify this interpretation, we tested the precise relations among infants' learning scores and HbO2 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 toward predicting learning scores ($F_{(5,31)} = 2.081$, $p = 0.095, R^2 = 0.251$), but right dlPFC activity during Stay 2 was not a significant predictor of learning scores ($F_{(5,31)} = 1.324, p =$ $0.280, R^2 = 0.176$).

Interactions among 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 investigate this, we calculated the interaction between the change in right dIPFC activity and eye blink rate by multiplying the change in right dIPFC 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; Fig. 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 vs Stay 1, Switch 1 vs Switch 2, Stay 1 vs Stay 2, Stay 2 vs Switch 2, Stay 1 vs Switch 2, and Switch 1 vs Stay 2) using learning score as a continuous variable. We used a Bonferroni-corrected α value set to 0.008 (0.05/6). These analyses indicated that there was a significant main effect of condition in the interaction for Switch 2 relative to Switch 1 only ($F_{(1,35)} = 8.447, p = 0.006, \eta_p^2 = 0.194$), which reflected the fact that right dlPFC activity and eye blink rate

Table 1. Repeated-measures ANOVAs for all comparisons of conditions

		Switch 1 vs Switch 2			Switch 1 vs Stay 1			Switch 1 vs Stay 2			Stay 1 vs Stay 2			Stay 1 vs Switch 2			Stay 2 vs Switch 2		
	d.f.	F	р	η_p^2	F	р	η_p^2	F	р	η_p^2	F	р	η_p^2	F	р	η_p^2	F	р	η_p^2
Interval	4	1.09	0.362	0.030	2.21	0.071	0.059	1.06	0.381	0.029	9.41	< 0.001	0.212	10.87	<0.001	0.237	7.99	<0.001	0.186
Condition	1	0.30	0.589	0.008	0.57	0.455	0.016	0.16	0.694	0.004	0.39	0.537	0.011	2.50	0.123	0.067	2.23	0.144	0.060
Learning score	1	0.002	0.965	< 0.001	0.13	0.719	0.004	0.60	0.443	0.017	0.20	0.888	0.001	0.30	0.585	0.009	0.99	0.328	0.027
Interval*condition	4	3.56	0.008	0.092	3.62	0.008	0.094	2.88	0.025	0.076	0.30	0.879	0.008	0.93	0.447	0.026	0.60	0.667	0.017
Interval*learning score	4	1.99	0.100	0.054	0.16	0.959	0.005	1.43	0.228	0.039	1.96	0.103	0.053	0.38	0.826	0.011	1.61	0.174	0.044
Condition*learning score	1	0.01	0.906	< 0.001	0.28	0.602	0.008	0.41	0.525	0.012	2.40	0.130	0.064	0.33	0.567	0.009	1.52	0.226	0.042
Interval*condition*learning score	1	0.09	0.987	0.002	1.76	0.141	0.048	2.25	0.067	0.060	2.25	0.067	0.060	2.91	0.024	0.077	5.33	0.001	0.132

Bold indicates significant at p < 0.008. d.f., degrees of freedom.



Figure 5. Right dIPFC 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 illustrate this interaction visually, the baseline-corrected change in right dIPFC activity is shown in the 10 infants with the worst learning scores (**B**). Error bars indicate SEM.



Figure 6. *A*, Significant cortical activity was observed in channels over right dIPFC (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 among right dIPFC 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 dIPFC 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 dIPFC activity showed better transfer performance.

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 = 0.010, $\eta_p^2 = 0.173$), 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 = 0.520, $\eta_p^2 = 0.012$), Stay 2 versus Switch 2 ($F_{(1,35)} = 0.728$, p = 0.399, $\eta_p^2 = 0.020$), Switch 1 versus Stay 2 ($F_{(1,35)} = 5.227$, p = 0.028, $\eta_p^2 = 0.184$), or Stay 1 versus Switch 2 ($F_{(1,35)} = 0.001$, p = 0.981, $\eta_p^2 = 0.069$).

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 ($\beta = 0.885$, p = 0.001; Fig. 6*B*). 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 ($\beta = 0.512$, p = 0.097) and explained statistically more variance than either eye blink rate or PFC activation alone ($\beta = -0.242$, p = 0.371, and $\beta = -0.218$, p = 0.300, respectively). Together, these findings indicate that infants who had better learning and subsequent 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.

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 (Karson, 1983; Shukla, 1985; Blin et al., 1990; Kleven and Koek, 1996; Taylor et al., 1999; Colzato et al., 2009), 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 (Karson, 1983; Shukla, 1985; Blin et al., 1990; Kleven and Koek, 1996; Taylor et al., 1999; e.g., Colzato et al., 2009). However, computational models suggest that phasic dopamine firing relates to reward prediction errors during learning, which drives learning of hierarchical rule structures (Frank and Badre, 2012; Collins and Frank, 2013). 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; Niv et al., 2007); therefore, 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 and Frank, 2013). Before 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, so 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 are also consistent with prior work showing a more general role for dlPFC in the application and maintenance of abstract rules after initial learning (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. Because significant dIPFC 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 dopaminerelated 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 reward prediction errors are greatest early in learning and that relevant mappings between contexts and rule structures must be learned before PFC can reliably switch between the rule structures (Collins and Frank, 2013).

Our results are somewhat surprising, however, given that PFC is not thought to reach maturity until adolescence (Huttenlocher, 1979) and rule learning and working memory processes are thought to have a similarly protracted course of development (Crone et al., 2006; Bunge and Zelazo, 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 ecolog-

ical 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. Therefore, 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. Although 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 have to be learned simultaneously, whereas the infant studies do not require such action control. Therefore, infants might engage similar circuits and computations that 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.

References

- Aasted CM, Yücel MA, Cooper RJ, Dubb J, Tsuzuki D, Becerra L, Petkov MP, Borsook D, Dan I, Boas DA (2015) Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. Neurophotonics 2:020801. CrossRef Medline
- Bacher LF, Smotherman WP (2004) Spontaneous eye blinking in human infants: a review. Dev Psychobiol 44:95–102. CrossRef Medline
- Blandini F, Nappi G, Tassorelli C, Martignoni E (2000) Functional changes of the basal ganglia circuitry in Parkinson's disease. Prog Neurobiol 62: 63–88. CrossRef Medline
- Blin O, Masson G, Azulay JP, Fondarai J, Serratrice G (1990) Apomorphineinduced blinking and yawning in healthy volunteers. Br J Clin Pharmacol 30:769–773. Medline
- Bodfish JW, Powell SB, Golden RN, Lewis MH (1995) Blink rate as an index

of dopamine function in adults with mental retardation and repetitive behavior disorders. Am J Ment Retard 99:335–344. Medline

- Bunge SA (2004) How we use rules to select actions: a review of evidence from cognitive neuroscience. Cogn Affect Behav Neurosci 4:564–579. CrossRef Medline
- Bunge SA, Zelazo PD (2006) A brain-based account of the development of rule use in childhood. Curr Dir Psychol 15:118–121. CrossRef
- Cavanagh JF, Wiecki TV, Kochar A, Frank MJ (2014) Eye tracking and pupillometry are indicators of dissociable latent decision processes. J Exp Psychol-Gen 143:1476–1488. CrossRef Medline
- Chugani HT (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 187–195. San Diego: Academic.
- Collins AG, Frank MJ (2013) Cognitive control over learning: Creating, clustering, and generalizing task-set structure. Psychol Rev 120:190–229. CrossRef Medline
- Collins AG, Frank MJ (2016) Neural signature of latent structure learning predicts hierarchical clustering and transfer of rule sets. Cognition 152: 160–169. CrossRef Medline
- Collins AG, Cavanagh JF, Frank MJ (2014) Human EEG uncovers latent generalizable rule structure during learning. J Neurosci 34:4677–4685. CrossRef Medline
- Colzato LS, van den Wildenberg WP, Hommel B (2008) Reduced spontaneous eye blink rates in recreational cocaine users: evidence for dopaminergic hypoactivity. PLoS One 3:e3461. CrossRef Medline
- Colzato LS, Slagter HA, van den Wildenberg WP, Hommel B (2009) Closing one's eyes to reality: evidence for a dopaminergic basis of psychoticism from spontaneous eye blink rates. Pers Individ Diff 46:377–380. CrossRef
- Crone EA, Wendelken C, Donohue S, van Leijenhorst L, Bunge SA (2006) Neurocognitive development of the ability to manipulate information in working memory. Proc Natl Acad Sci U S A 103:9315–9320. CrossRef Medline
- Deuschl G, Goddemeier C (1998) Spontaneous and reflex activity of facial muscles in dystonia, Parkinson's disease and in normal subjects. J Neurol Neurosurg Psychiatry 64:320–324. CrossRef Medline
- Donoso M, Collins AG, Koechlin E (2014) Foundations of human reasoning in the prefrontal cortex. Science 344:1481–1486. CrossRef Medline
- Dreisbach G, Müller J, Goschke T, Strobel A, Schulze K, Lesch KP, Brocke B (2005) Dopamine and cognitive control: the influence of spontaneous eyeblink rate and dopamine gene polymorphisms on perseveration and distractibility. Behav Neurosci 119:483–490. CrossRef Medline
- Elsworth JD, Lawrence MS, Roth RH, Taylor JR, Mailman RB, Nichols DE, Redmond DE Jr (1991) D1 and D2 dopamine receptors independently regulate spontaneous blink rate in the vervet monkey. J Pharmacol Exp Ther 259:595–600. Medline
- Emberson LL, Richards JE, Aslin RN (2015) Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. Proc Natl Acad Sci U S A 112:9585–9590. CrossRef Medline
- Frank MJ, Badre D (2012) Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb Cortex 22: 509–526. CrossRef Medline
- Freed WJ, Kleinman JE, Karson CN, Potkin SG, Murphy DL, Wyatt RJ (1980) Eye-blink rates and platelet monoamine oxidase activity in chronic schizophrenic patients. Biol Psychiatry 15:329–332. Medline
- Fukuda K, Stern JA, Brown TB, Russo MB (2005) Cognition, blinks, eyemovements, and pupillary movements during performance of a running memory task. Aviat Space Environ Med 76:C75–C85. Medline
- Goel V, Dolan RJ (2000) Anatomical segregation of component processes in an inductive inference task. J Cogn Neurosci 12:110–119. CrossRef Medline
- 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: 4877–4882. CrossRef Medline
- Huppert TJ, Diamond SG, Franceschini MA, Boas DA (2009) HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. Appl Opt 48:D280–D298. CrossRef Medline
- Huttenlocher PR (1979) Synaptic density in human frontal cortex: developmental changes and effects of aging. Brain Res 163:195–205. CrossRef Medline
- Jasper HH (1958) The ten-twenty electrode system of the International Federation. Electroenceph Clin Neurophysiol 10:371–375.

- Jutkiewicz EM, Bergman J (2004) Effects of dopamine D1 ligands on eye blinking in monkeys: efficacy, antagonism, and D1/D2 interactions. J Pharmacol Exp Ther 311:1008–1015. CrossRef Medline
- Karson CN (1983) Spontaneous eye-blink rates and dopaminergic systems. Brain 106:643–653. CrossRef Medline
- Kleven MS, Koek W (1996) Differential effects of direct and indirect dopamine agonists on eye blink rate in cynomolgus monkeys. J Pharmacol Exp Ther 279:1211–1219. Medline
- Lloyd-Fox S, Blasi A, Volein A, Everdell N, Elwell CE, Johnson MH (2009) Social perception in infancy: a near infrared spectroscopy study. Child Dev 80:986–999. CrossRef Medline
- Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (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. CrossRef Medline
- Mohr C, Sándor PS, Landis T, Fathi M, Brugger P (2005) Blinking and schizotypal thinking. J Psychopharmacol 19:513–520. CrossRef Medline
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology 191:507–520. CrossRef Medline
- Oh J, Jeong SY, Jeong J (2012) The timing and temporal patterns of eye blinking are dynamically modulated by attention. Hum Mov Sci 31:1353–1365. CrossRef Medline
- Rovee-Collier C, Cuevas K (2009) Multiple memory systems are unnecessary to account for infant memory development: an ecological model. Dev Psychol 45:160–174. CrossRef Medline
- Seger CA, Poldrack RA, Prabhakaran V, Zhao M, Glover GH, Gabrieli JD (2000) Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. Neuropsychologia 38: 1316–1324. CrossRef Medline

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

Slagter HA, Georgopoulou K, Frank MJ (2015) Spontaneous eye blink rate

predicts learning from negative, but not positive, outcomes. Neuropsychologia 71:126–132. CrossRef Medline

- Strange BA, Henson RN, Friston KJ, Dolan RJ (2001) Anterior prefrontal cortex mediates rule learning in humans. Cereb Cortex 11:1040–1046. CrossRef Medline
- Tada H (1978) Spontaneous blinking during a visual tracking performance. Fukushima J Med Sci 25:91–100. Medline
- Taylor JR, Elsworth JD, Lawrence MS, Sladek JR Jr, Roth RH, Redmond DE Jr (1999) Spontaneous blink rates correlate with dopamine levels in the caudate nucleus of MPTP-treated monkeys. Exp Neurol 158:214–220. CrossRef Medline
- Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A, Casey BJ, Nelson C (2009) The NimStim set of facial expressions: Judgments from untrained research participants. Psychiatry Res 168:242–249. CrossRef Medline
- Werchan DM, Collins AG, Frank MJ, Amso D (2015) Spontaneous hierarchical rule learning in 8-month-old infants. Psychol Sci 26:805–815. CrossRef Medline
- White BR, Snyder AZ, Cohen AL, Petersen SE, Raichle ME, Schlaggar BL, Culver JP (2009) Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. Neuroimage 47:148–156. CrossRef Medline
- White IM, Wise SP (1999) Rule-dependent neuronal activity in the prefrontal cortex. Exp Brain Res 126:315–335. CrossRef Medline
- Wilcox T, Bortfeld H, Woods R, Wruck E, Boas DA (2008) Hemodynamic response to featural changes in the occipital and inferior temporal cortex in infants: a preliminary methodological exploration. Dev Sci 11:361–370. CrossRef Medline
- 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. CrossRef Medline