A Neural Correlate of Strategic Exploration at the Onset of Adolescence

Andrew S. Kayser[1,2], Zdeňa Op de Macks[3], Ronald E. Dahl[3], and Michael J. Frank[4]

Abstract

■ The onset of adolescence is associated with an increase in the behavioral tendency to explore and seek novel experiences. However, this exploration has rarely been quantified, and its neural correlates during this period remain unclear. Previously, activity within specific regions of the rostrolateral PFC (rlPFC) in adults has been shown to correlate with the tendency for exploration. Here we investigate a recently developed task to assess individual differences in strategic exploration, defined as the degree to which the relative uncertainty of rewards directs responding toward less well-evaluated choices, in 62 girls aged 11–13 years from whom resting state fMRI data were obtained in a separate session. Behaviorally, this task divided our participants into groups of explorers ($n = 41$) and nonexplorers ($n = 21$). When seed ROIs within the rlPFC were used to interrogate resting state fMRI data, we identified a lateralized connection between the rlPFC and posterior putamen/insula whose strength differentiated explorers from nonexplorers. On the basis of Granger causality analyses, the preponderant direction of influence may proceed from posterior to anterior. Together, these data provide initial evidence concerning the neural basis of exploratory tendencies at the onset of adolescence. ■

INTRODUCTION

The decision to continue to exploit a known source of reward or to explore the environment for a potentially greater one depends upon a balance of factors whose weighting is subjective. For example, when confronted with the choice between an activity whose reward is known (such as eating ice cream) and a new one that might—or might not—be even better (such as trying a new athletic activity or smoking a cigarette for the first time), some individuals may choose what they know, whereas others may elect to try the unknown option. More generally, the brain must weigh the advantages of exploiting the action associated with a more certain outcome against exploring an action whose payoff is more unspecified. Notably, this type of exploration is strategic rather than random: it is driven by the relative uncertainty of options within the reward space, so that outcomes maximize information that has most potential to improve the status quo (Frank, Doll, Oas-Terpstra, & Moreno, 2009).

Adolescence is thought to be a time of exploration (Forbes & Dahl, 2010; Kelley, Schochet, & Landry, 2004). In conjunction with changes in sensation seeking, risk tolerance, and other traits, such strategic exploration may serve an evolutionary purpose by encouraging adolescents to develop behaviors adaptive to function in new social and behavioral contexts (Kelley et al., 2004). Importantly, the outcomes of exploratory behaviors are likely to differ across adolescents: In some individuals, for example, exploration may increase vulnerability to risky behaviors by exposing them to previously undiscovered detrimental activities perceived to be rewarding (e.g., cigarette smoking), whereas in others, exploration may increase resilience by ensuring that other more constructive rewards (e.g., athletic participation) are continually assessed. Moreover, the degree to which adolescents strategically navigate the exploration–exploitation tradeoff is likely variable across individuals. To our knowledge, this latter question has yet to be rigorously evaluated.

Because of their potential importance in periadolescence, differences in exploratory behaviors between individuals are likely to have a neural correlate. Rostrolateral PFC (rlPFC) is thought to be important for evaluating the efficacy of behavioral strategies and deciding whether alternative strategies need to be pursued (Donoso, Collins, & Koechlin, 2014). Recent studies have directly linked exploratory behavior to activity within the rlPFC, as measured using fMRI (Badre, Doll, Long, & Frank, 2012; Boorman, Behrens, Woolrich, & Rushworth, 2009; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006), and our previous work identified a specific region within the right rlPFC for which activity varied with outcome uncertainty in participants who demonstrated strategic exploratory behaviors, compared with those without (Badre et al., 2012).

This finding holds particular interest for adolescents, in whom the rlPFC has not yet fully developed (reviewed in

[1]University of California, San Francisco, [2]VA Northern California Health Care System, Martinez, CA, [3]University of California, Berkeley, [4]Brown University

Dumontheil, 2014). Gray matter density peaks in late childhood and early adolescence before declining, and the fractional anisotropy of white matter tracts, depending on brain region, does not reach its adult levels until late adolescence or early adulthood (Lebel, Walker, Leemans, Phillips, & Beaulieu, 2008). These structural changes are accompanied by functional changes: in paradigms assessing the development of higher-order reasoning, for example, adolescents have been shown to activate the same regions as adults, though with relative differences in activity (both increases and decreases; Dumontheil, 2014). In contrast, the relative maturity of subcortical systems has led to theories that motivated behaviors in adolescence may reflect reward sensitivity within corticostriatal loops (Alexander, DeLong, & Strick, 1986), but in the context of limited cognitive control (Gladwin, Figner, Crone, & Wiers, 2011; Galvan, 2010; Somerville & Casey, 2010; Steinberg, 2008). Such theories suggest that exploratory behaviors at the onset of adolescence may engage the rlPFC, but that individual differences in exploration in this age range may reflect disparities in the connectivity of this area with subcortical and posterior brain regions.

In particular, hypothesized changes in connectivity are likely to be found with regions important for strategic exploration: those involved in the assessment of uncertainty and the association of action plans with reward. The insula, and the salience network more generally, are known to be involved in the evaluation of uncertainty, whether in the context of rewards or selected actions, in both adults and adolescents (Smith, Steinberg, & Chein, 2014; White, Engen, Sorensen, Overgaard, & Shergill, 2014; Preuschoff, Quartz, & Bossaerts, 2008; Huettel, 2006). Likewise, the striatum, including the ventral striatum, has a well-established role in associating rewards with motor plans, possibly via an anatomical structure that progressively links the nucleus accumbens with the dorsolateral striatum (Haber & Knutson, 2010; Haber, Fudge, & McFarland, 2000). These previous results suggest that, in adolescents who engage in exploratory behaviors, the rlPFC may demonstrate behaviorally relevant connectivity with uncertainty and reward-related areas that include the insula and striatum. To our knowledge, this possibility has yet to be addressed.

Given these open questions, here we use a previously validated task to assess individual differences in exploration within a group of early adolescent girls, ages 11–13 years old, from whom resting state fMRI (rs-fMRI) images were also obtained. We hypothesized that differences in strategic exploration should correlate with the degree to which neural representations of uncertainty and action are incorporated by the rlPFC and, specifically, that participants with greater exploration should show greater connectivity between the rlPFC and relevant subcortical/posterior brain regions. Moreover, we predicted that, because a tendency toward exploration may remain consistent across transient behavioral states,

changes in connectivity should be reflected in the resting state. Finally, we hypothesized that this distinct neural representation would argue for a behavioral dissociation of strategic exploration from other factors relevant to early adolescence, including risk aversion.

METHODS

Study Population

At a single time point within a larger study designed to investigate longitudinal changes in adolescent girls, we evaluated 76 healthy periadolescent girls who were without a history of neurological or psychiatric disorders and between 11 and 13 years old at the time of behavioral testing. Sixty-six of these girls participated in MRI scanning, of whom 62 completed the exploration–exploitation task and formed the subject group. Twenty-eight participants were 11 years old, 20 were 12 years old, and 14 were 13 years old. Scores on the pubertal development scale (Petersen, Crockett, Richards, & Boxer, 1988) ranged from 1.2 to 3.8, with a mean of 2.5 ± 0.7. We limited our study to girls because pubertal status could be more sharply defined and to avoid confounds resulting from potentially differential effects of gender on exploration. A parent or guardian gave written informed consent for each participant in accordance with the Committee for the Protection of Human Subjects at the University of California, Berkeley. All participants also provided written assent and were paid approximately $75 via gift card for their participation.

Experimental Paradigm

Participants performed the exploration–exploitation task outside the MRI scanner, and their behavior was correlated with rs-fMRI data. As in our previous work (Kayser, Mitchell, Weinstein, & Frank, 2015; Badre et al., 2012; Frank et al., 2009), on each trial participants observed a clock that completed a revolution over 5 sec. Following instructions that sometimes they would do better by responding faster and sometimes by responding slower, they stopped the clock with a key press during the 5 sec in an attempt to win points. Rewards were delivered with a probability and magnitude that varied as a function of RT; together, these factors defined the reward space for each condition. Importantly, participants were not cued to the nature of the reward space beforehand, requiring them to learn how reward probability and magnitude varied with duration from trial onset. Over the course of 50 trials, participants explored each of four conditions, named in accordance with the change in expected value (Probability × Magnitude) with increasing RT: increasing expected value (IEV), decreasing expected value (DEV), constant expected value (CEV), and constant expected value-reversed (CEVR; see Figure 1). CEV and CEVR are distinguished by contrasting reward frequency

Figure 1. (A) Participants viewed a clock face while deciding when to stop a marker that made one clockwise rotation over 5 sec. (B) The probability of reward for each of the four conditions across the five second trial time. (C) The magnitude of reward for each of the four conditions across the five second trial time. (D) The expected value of the reward (Probability × Magnitude) for each of the four conditions. Note that DEV and IEV strongly differentiate expected value but share monotonic declines in reward probability.
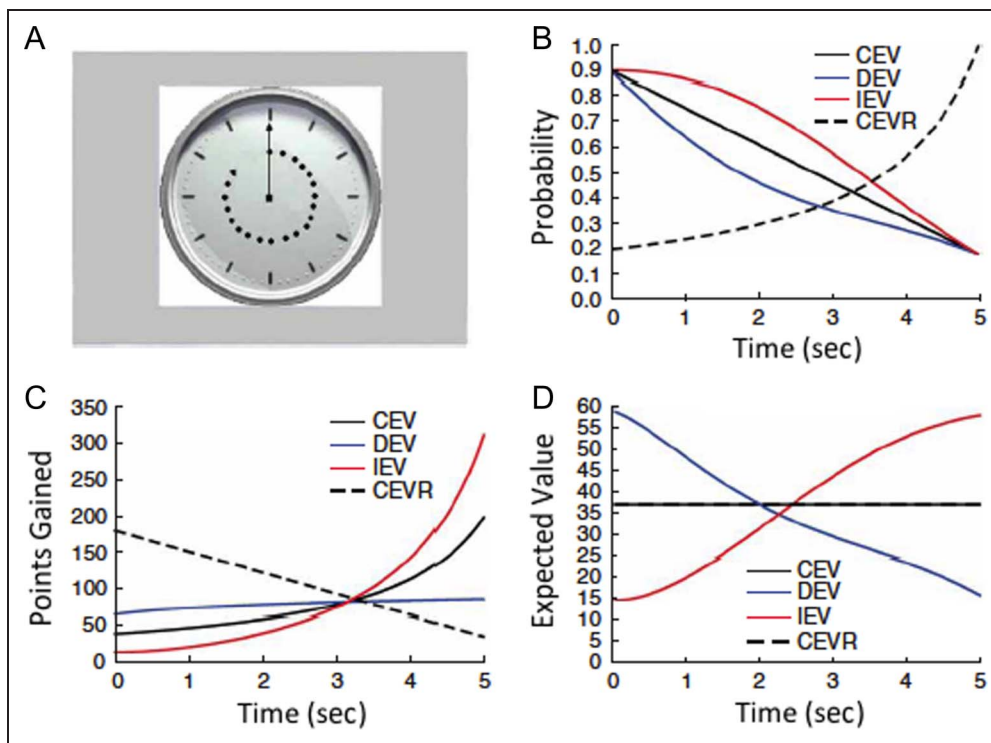
and reward magnitude curves whose product gives rise to overlapping expected value curves. Each participant completed a total of 200 trials (4 conditions × 50 trials per condition), with the order of the conditions counterbalanced across participants. Although participants were not directly cued to the nature of the reward space, the clock face changed color between runs to indicate that a different task condition was present.

A computational model was fit to each participant's behavioral data to estimate the magnitude of the exploration parameter ε (see below for additional details on the full model). Briefly, ε captures the degree to which uncertainty in the reward space drives individual participants to explore. At the beginning of learning, participants have little knowledge as to whether faster or slower responses will produce greater expected reward (Figure 2). (Without limiting the generality of the results, our previous studies have shown that RTs in the exploration–exploitation task can be well predicted by assuming that participants track the outcome statistics associated with two general classes of responses, "fast" versus "slow"; Badre et al., 2012; Frank et al., 2009.) After participants experience more trials, the feedback about rewards allows them to reduce the uncertainty related to their prior beliefs, manifest as a reduction in the variance of the relevant belief distribution. The explore parameter ε indexes the degree to which participants guide exploration toward the more uncertain of these (fast/slow) distributions to increase their understanding of the reward space. In equation form,

$$\text{Exploration }(t) = \varepsilon * [\sigma_{\text{slow}}(t) - \sigma_{\text{fast}}(t)]$$
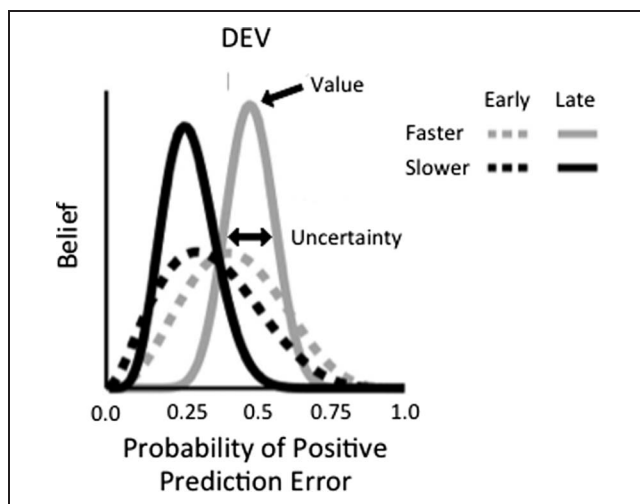
Figure 2. Intuition for the explore parameter. The strength of a participant's belief (y axis) about the probability of a better than expected outcome (x axis) depends upon both the stage of learning and the task condition (shown for DEV). In early stages (dashed lines) of the DEV condition the participant has not yet determined whether a faster (in gray) or slower (in black) response is more likely to be rewarded, a state reflected in both a weaker belief strength (value on the y axis) and greater uncertainty (a broader belief distribution). Later in the task (solid lines), belief strength increases and the uncertainty about reward likelihood decreases—that is, the participant may learn that faster responses are more likely to yield reward (i.e., a positive prediction error). The explore parameter indexes the degree to which participants use the relative uncertainty between the faster and slower distributions to explore the reward space (i.e., to reduce the variance in the more uncertain distribution).

where the magnitude of Exploration in milliseconds at time t is proportional to the difference between the standard deviations σ of the fast and slow distributions at that time, weighted by a scale factor ε derived from each participant's data. ε therefore represents how strongly each participant uses the differences between the standard deviations to drive RT. In keeping with other conceptualizations of exploration (e.g., Daw et al., 2006), this component of the model predicts that, in a given trial, participants will increase RTs if the outcome statistics are more uncertain for slow than fast responses, and vice versa.

Unlike the explore parameter ε, which scales the difference in the uncertainties of the two (fast/slow) distributions, the exploitation parameter ρ scales the degree to which participants adjust their responses as a function of the relative difference in the mean expected values:

$$\text{Exploitation}\,(t) = \rho * \left[\mu_{\text{slow}}(t) - \mu_{\text{fast}}(t) \right]$$

where the magnitude of Exploitation at time t is proportional to the difference between the mean rewards μ of the fast and slow distributions at that time, weighted by a scale factor ρ fit to each participant's data. This component of the model predicts that participants will increase RTs if the average reward is greater for slow than fast responses, and vice versa—that is, it reflects exploitation of prior learning. The exploitation factor therefore complements factors associated with exploration, in that reward depends on knowledge of the mean but exploration depends on the variance.

More specifically, both exploitation of the RTs producing the highest rewards and exploration for even better rewards are driven by errors of prediction in tracking expected reward value V. On trial t, the expected reward for each clock face is calculated as

$$V(t) = V(t-1) + a\delta(t-1)$$

where a determines how rapidly V is updated and δ is the reward prediction error. This V value represents the average expected value of rewards for each clock face; the participant's goal is to attempt to maximize her rewards by selecting RTs that produce the largest number of positive deviations from this average. Because it is unreasonable to assume that learners track a separate value for each RT from 0 to 5000 msec, a simplifying assumption is that the learners can track the values of responses that are faster or slower than average, as noted above.

For consistency with prior reports, the full RT model included additional contributions to responding that were not a focus of the present experiment. The full model estimates RT ($\hat{R}T$) on trial t as follows:

$$\begin{aligned} \hat{R}T(t) = {} & K + \lambda RT(t-1) - \text{go}(t) + \text{no go}(t) \\ & + \rho[\mu_{\text{slow}}(t) - \mu_{\text{fast}}(t)] + \nu\left[RT_{\text{best}} - RT_{\text{avg}}\right] \\ & + \text{Explore}(t) \end{aligned}$$

where K is a free parameter capturing baseline response speed, λ reflects autocorrelation between the current and previous RT, go and no-go learning reflect a striatal bias to speed (slow) responding as a function of positive (negative) reward prediction errors, and ν captures the tendency to adapt RTs toward the single largest reward experienced thus far. Their inclusion allows the model to better fit overall RTs, but we have found that they do not change the pattern of findings associated with exploration or exploitation. For other details regarding the primary continuous RT model, including alternative models that provide poorer behavioral fits, please see Frank et al. (2009). For each participant, the Simplex method was used to estimate free parameters that minimized the sum of the squared error between predicted and observed RTs. Goodness of fit was assessed by a sum of squared error term that reflected the difference between participant behavior and model predictions.

MRI Image Acquisition

MRI scanning was conducted on a Siemens MAGNETOM Trio 3T MR Scanner (Berlin, Germany) at the Henry H. Wheeler, Jr., Brain Imaging Center at the University of California, Berkeley. Anatomical images consisted of 160 slices acquired using a T1-weighted MPRAGE protocol (repetition time = 2300 msec, echo time = 2.98 msec, field of view = 256 mm, matrix size = 256 × 256, voxel size = 1 mm³). During two 5-min rs-fMRI runs, functional images consisting of 24 slices were acquired in interleaved fashion with a gradient echo-planar imaging protocol (repetition time = 1370 msec, echo time = 27 msec, field of view = 225 mm, matrix size = 96 × 96, voxel size = 2.3 × 2.3 × 3.5 mm).

fMRI Preprocessing

fMRI preprocessing was performed using both the AFNI (afni.nimh.nih.gov) and FSL (www.fmrib.ox.ac.uk/fsl/) software packages. Functional images were converted to 4-D NIfTI format and corrected for slice-timing offsets. Motion correction was carried out using the AFNI program *3dvolreg*, with the reference volume set to the mean image of the first run in the series. Coregistration with the anatomical scan was performed using the AFNI program *3dAllineate*, and anatomical images were normalized to a standard volume (MNI_N27) using the FSL program *fnirt*. The same normalization parameters were later applied to native-space statistical maps to generate group statistical maps.

Multivariate Analysis

Resting state data were smoothed by a 5-mm FWHM Gaussian kernel before temporal bandpass filtering between 0.009 and 0.08 Hz to reduce the influence of cardiac and respiratory artifact (Fox et al., 2005). Movement parameters and the white matter and ventricular time series, but not the global mean signal, were included

as regressors of no interest. Because motion can severely impact resting state data, data were then scrubbed (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012). On the basis of our acquisition parameters, we removed frames in which the summed variance of the temporal derivative of the BOLD signal (DVARS) was greater than 0.03 and the maximal motion displacement was greater than 2.5 mm. Various ROIs within the rostral PFC (see Results and Table 1 later in the article) were then selected, based on previous work with this and related tasks (Badre et al., 2012; Boorman et al., 2009; Daw et al., 2006). Each ROI was defined by a set of MNI coordinates that formed the center for a sphere with 5-mm radius. A time course defined by averaging across voxels in this region was then correlated either with all other voxels in the brain (whole-brain analyses) or with specific ROIs (ROI–ROI analyses), and correlation coefficients were Fisher-transformed to allow for the application of parametric statistical tests. (As for other rs-fMRI analyses, the so-called "univariate" contrasts were not possible because of the lack of a contrasting baseline condition and the absence of discrete task epochs.) For whole-brain analyses, images were normalized to the MNI template before the application of group level statistics. Map-wise significance ($p < .05$, corrected for multiple comparisons) was determined by applying a cluster-size correction derived from the AFNI programs *3dFWHMx* and *3dClustSim* to data initially thresholded at a value of $p < .005$, uncorrected. Because of our hypotheses about changes in frontostriatal connectivity, the volume of a frontal mask (AAL regions 3–32 and 71–76; Tzourio-Mazoyer et al., 2002) was used to calculate the appropriate cluster size correction (equal to 16 contiguous voxels).

To evaluate the temporal influence of these regions in ROI–ROI analyses, we used bivariate Granger causality. This technique determines whether the time series in one voxel or region helps to predict upcoming time points in a second time series; if so, that voxel or region is said to be Granger causal for the second. Using custom Matlab-based analysis scripts (www.mathworks.com)

developed in our previous work, we restricted our analysis to linear autoregressive models (see Kayser, Sun, & D'Esposito, 2009, for full details).

RESULTS

Sixty-two early adolescent girls completed both the exploration–exploitation task and resting state MRI scans. Before examining exploration explicitly, we first ensured that participants performed the task well. Out of 12,400 total trials (50 trials per condition × 4 conditions × 62 participants), there were 51 no-responses (0.41%), indicating excellent task engagement. More importantly, participants' performance in the different task conditions could be readily distinguished via their mean RTs over the latter half of trials (Figure 3A). In a two-way ANOVA including the factor of Task condition with Participants as a random effect, a strongly significant effect of Task condition could be seen, $F(3, 61) = 4.59$, $p = .004$. In post hoc t tests, this finding was driven by RTs within the CEVR condition, which were significantly longer than in the CEV ($t(61) = 2.97$, $p = .004$), DEV ($t(61) = 3.08$, $p = .003$), and IEV ($t(61) = 2.79$, $p = .007$) conditions. Furthermore, only in the CEVR condition did participants demonstrate a significant increase in RT from the first half of trials to the second ($t(61) = 2.02$, $p = .048$). Importantly, neither age nor scores on the pubertal development scale (Petersen et al., 1988) influenced performance on any of the individual task conditions (age: all p values > .22; PDS: all p values > .08).

Notably, this pattern of performance demonstrates a form of risk aversion or probability-magnitude bias: Because both the CEV and CEVR conditions have constant expected value across the entire duration of the clock, the significant difference between them indicates a differential sensitivity to reward frequency (Figure 1)—that is, participants were more averse to low-frequency rewards despite their larger magnitudes. Results for the IEV condition (Figure 3A) were in keeping with this idea.
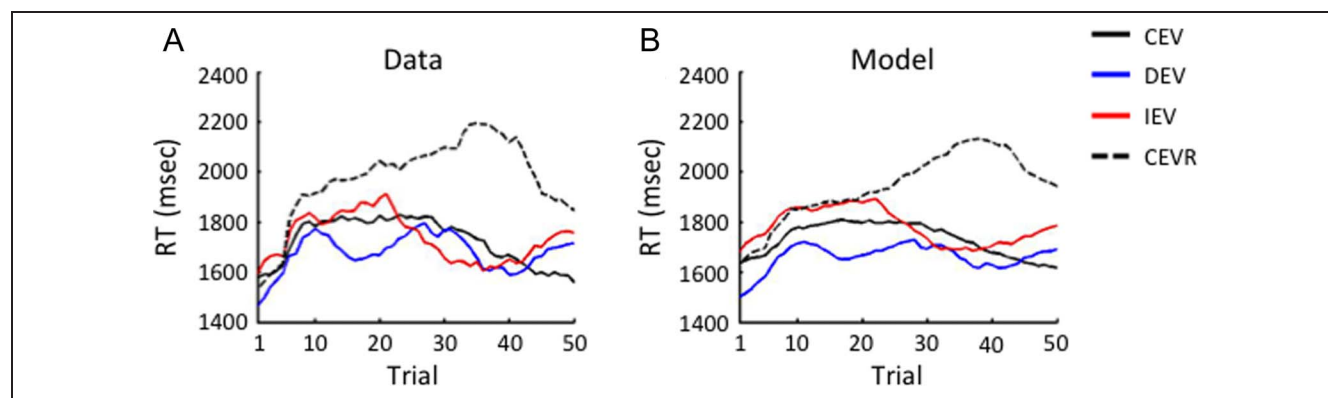
Figure 3. (A) Shown are the average RTs across all 62 participants for each of the four conditions, smoothed by a 10-trial weighted average. As evident in the plots, participant behavior strongly differentiated the conditions, $F(3, 61) = 4.59$, $p = .004$. (B) Model predictions for the average RT across trials for all 62 participants, demonstrating good agreement with the behavioral data.

Specifically, participants who maximized expected value would be anticipated to slow responding as they learned about the reward space for IEV, whereas participants who preferred more frequent but smaller rewards over proportionally larger, less frequent rewards would be anticipated to respond rapidly. Consistent with the latter possibility, no increase in RT was seen between the first and second halves of trials for the IEV condition ($t(61) = -0.96$, $p = .34$), and no difference in RT was evident between the two conditions, IEV and DEV, that most strongly differentiated expected value ($t(61) = -0.10$, $p = .9$. Rather, as noted above, participants responded later only when reward frequency, not expected value, increased with time (the CEVR condition). This result stands in contrast with our previous work in adults (Kayser et al., 2015; Badre et al., 2012; Frank et al., 2009), who strongly tracked expected value.

To ensure that this pattern of responding reflected strategic, rather than random, evaluation of the reward space, we compared trial-by-trial changes in RT across participants with model-derived estimates of participants' relative uncertainty about the outcomes of faster versus slower responses (Figure 4). Within our group of 62 early adolescent participants, 41 evinced a positive explore parameter ("explorers"), and 21 did not ("nonexplorers"). If explorers used uncertainty about the reward space to
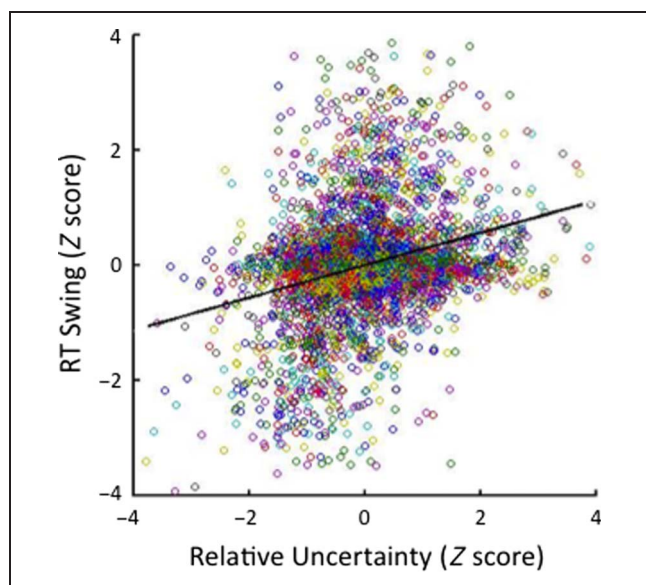

Figure 4. The correlation between the change in RT from one trial to the next ("RT swing") and the difference in the standard deviations of the fast and slow belief distributions ("Relative uncertainty") for all sessions with a positive explore parameter across participants. To facilitate comparison, both RT swing and relative uncertainty values have been converted to Z scores. As predicted, greater relative uncertainty on a given trial correlated with greater changes in RT from that trial to the next (mean regression coefficient = 0.28, $p < .001$ for the t test evaluating whether these coefficients are different from zero across participants). This positive correlation provides important support for the notion that experience-derived relative uncertainty, rather than undirected responding, drives exploration of the reward space.

drive responding, then greater relative uncertainty about slower responses (Figure 4: positive values, x axis), for example, should be correlated with slowing of RT on the next trial (Figure 4: positive swing in RT, y axis). In contrast, random evaluation of the reward space would lead to no relationship. Consistent with participants' strategic exploration of the reward space, a strongly significant positive correlation was seen between relative uncertainty and RT swing in the explorers (mean regression coefficient across the group = 0.28, significantly different from zero; $p < .001$).

Importantly, differences in exploration between participants could not be explained by other behavioral and demographic variables. RT data for the CEV, CEVR, DEV, and IEV conditions were not significantly different between explorers and nonexplorers (all p values > .062). Risk aversion, defined here as the difference between RTs in the CEVR and CEV conditions to minimize RT-related learning effects, was also no different between explorers and nonexplorers ($t(24) = -1.63$, $p = .11$, corrected for unequal variances). Additionally, the number of explorers did not vary by age: 21, 12, and 8 participants for ages 11, 12, and 13 years old, respectively (of 28, 20, and 14 total) had nonzero explore parameters ($X(2) = 0.62$, $p = .73$); and explorers showed no difference in scores on the pubertal development scale when compared with nonexplorers ($t(37) = -0.58$, $p = .57$, corrected for unequal variances). Finally, there were no significant differences in model fit between explorers and nonexplorers ($t(57) = 1.32$, $p = .19$, allowing for unequal variances) or across age groups ($F(2, 59) = 0.89$, $p = .42$), and no correlation was found between model fit and score on the pubertal development scale ($r = 0.13$, $p = .33$).

To assess the neural correlates of exploration in resting state data, we started with the single region in right rlPFC (MNI coordinates [24 46 20]) that differentiated explorers from nonexplorers in our previous work in adults (Badre et al., 2012). We hypothesized that connectivity with this region would likewise distinguish explorers from nonexplorers in this early adolescent sample. Because other studies have implicated not only right, but also left, rlPFC in exploratory behaviors, we also calculated resting state connectivity between a corresponding (mirror image) region in left rlPFC (MNI coordinates [-24 46 20]). As shown in Figure 5, the seed region in right rlPFC was more strongly connected to the right posterior putamen/insula ([32 -10 8]; cluster size 19 voxels, peak t value = 3.20) in explorers compared with nonexplorers, whereas the seed region in left rlPFC was more strongly connected to the left posterior putamen/insula ([-38 -14 -6]; cluster size 39 voxels, peak t value = 3.63) in explorers compared with nonexplorers (both results $p < .05$, corrected). In subsequent ROI–ROI analyses, the strengths of this rlPFC–putamen/insula connectivity across individuals did not correlate with either age ($ps > .21$ or score on the pubertal development scale ($ps > .45$, and the statistical difference in connectivity between explorers and

Figure 5. Adolescents who displayed exploratory behavior demonstrated greater lateralized resting state connectivity between seed regions in rlPFC (top, in green) and posterior striatum/insula (lower, in red) than nonexplorers. Two clusters were identified across the group comparison: for the seed centered at [24 46 20], a 19-voxel cluster at MNI coordinates [32 −10 8] with peak *t* score 3.20 (left), and for the seed centered at [−24 46 20], a 39-voxel cluster at MNI coordinates [−38 −14 −6] with peak *t* score 3.63 (both results *p* < .05, corrected). The adjacent dot plots demonstrate that these differences were not driven by outliers. As determined by Granger causality, the direction of the influence (blue arrows) between these regions proceeded from posterior striatum/insula to rlPFC.
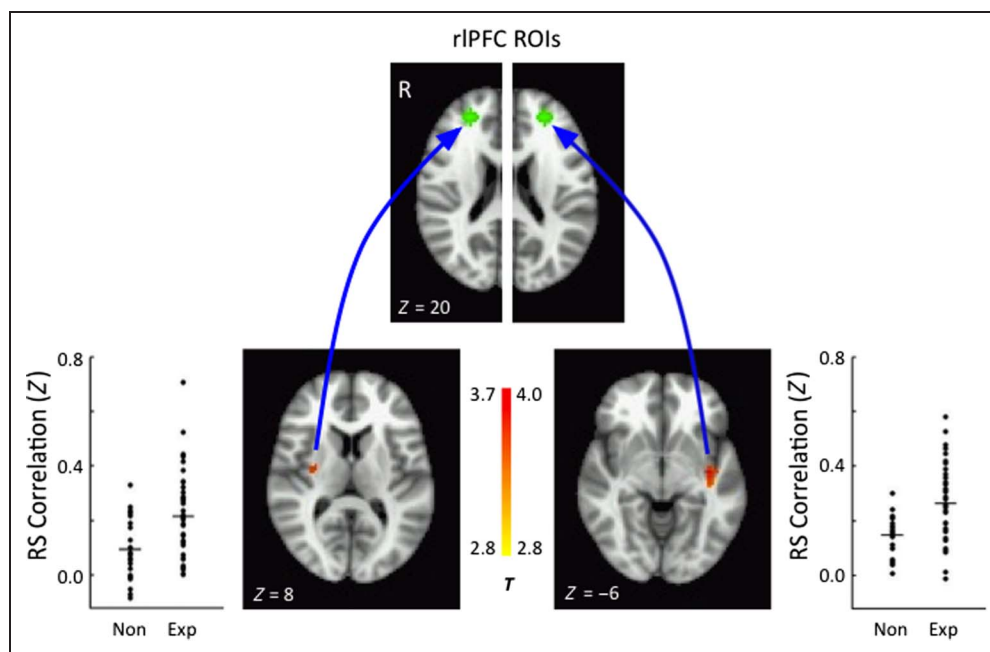
nonexplorers remained strongly significant when the variance explained by age and pubertal development scale was first removed by linear regression before the differences were calculated (ROI–ROI analyses for right-sided regions: $t(60) = 3.2, p = .002$; left-sided regions: $t(60) = 3.5, p = .0009$). Moreover, when participants were divided into exploiters ($n = 45$) and nonexploiters ($n = 17$) and whole-brain connectivity maps for the same rlPFC seed regions were evaluated, a significant difference between exploiters and nonexploiters was only identified for connectivity between one contralateral region and the right rlPFC seed. However, this contralateral region ([MNI coordinates [−29 55 28]) was closely adjacent to the area of the left rlPFC seed (MNI coordinates [−24 46 20]), encompassed voxels that were potentially outside the MNI template brain (data not shown), and was not confirmed by related ROI–ROI analyses (see below).

To ensure that these findings were not tied to a particular rlPFC seed region, we replicated these results for strategic exploration by evaluating connectivity differences between explorers and nonexplorers in an additional ROI–ROI analysis (Table 1) using the identified posterior putamen/insula regions and other functionally defined rlPFC seeds derived from previous work that investigated exploratory behaviors (Badre et al., 2012; Boorman et al., 2009; Daw et al., 2006). Similarly, to ensure that this finding was specific to exploration, we repeated these ROI–ROI analyses, but instead used the exploitation parameter to divide individuals into exploiters and nonexploiters. The number of true exploration-related positives (Table 1, shaded areas) was significantly greater

than that expected by chance ($p = .00002$, binomial theorem), whereas neither exploration- nor exploitation-related false positives occurred more than expected by chance ($p = .19$ and $p = .88$, respectively). Additionally, we found no significant, exploitation-related connectivity between the seed regions themselves for any of the seeds (data not shown). Finally, to evaluate whether this connectivity was directional, we applied Granger causality to the original finding. For both right- and left-sided connections (Figure 5, blue arrows), a significant lateralized Granger causal influence was found from the posterior striatum/insula to the rlPFC (right putamen/insula to right rlPFC: $p = .025$, left putamen/insula to left rlPFC: $p = .023$).

DISCUSSION

Here we demonstrate that the tendency for uncertainty-guided exploration shows significant individual variation around the onset of adolescence and that explorers show greater connectivity between the rlPFC and the putamen/insula than do nonexplorers. Importantly, these differences do not correlate with a measure of risk aversion, pubertal status, or age, suggesting that exploration itself represents an independent and differentiable component of cognitive function. Moreover, the direction of this connectivity proceeds from posterior to anterior ("bottom–up" rather than "top–down") as assessed by Granger causality, arguing that activity within these regions may reflect input to, rather than output from, rlPFC—a finding potentially consistent with previous theories that responses in subcortical and posterior structures mature

before those in PFC (Gladwin et al., 2011; Galvan, 2010; Somerville & Casey, 2010; Steinberg, 2008).

Despite the importance that individual differences in strategic exploratory behaviors are thought to play in adolescence, they have not previously been studied quantitatively. In contrast with the current task, a previous study evaluated noncontingent exploration, in which a model of individual choice variability in a multiple one-armed bandit task was dependent upon previous selections, but not upon previous rewards (Christakou et al., 2013). Interestingly, this particular form of choice sensitivity was both explained by age effects and correlated with activity within premotor cortex—perhaps consistent with individual variability in motor planning. Other work has focused on potentially related behaviors such as risk tolerance and self-regulation, also conceptualized as the presence of increased appetitive drive in the context of weak or immature control processes during adolescence. The current results demonstrate that exploration represents an additional complexity; here it was independent of the degree to which participants

Table 1. Additional ROI–ROI Analyses

	Exploration (ε)		Exploitation (ρ)	
ROI	R Str/Ins	L Str/Ins	R Str/Ins	L Str/Ins
24 46 20 (a)	*		~	
−24 46 20 (a)		*		
22 54 28 (a)	*	~		
−22 54 28 (a)		*		
36 56 −8 (a)				
−36 56 −8 (a)		*		
27 57 6 (b)	*			
−27 48 4 (b)	*			
36 54 0 (c)	*			
−34 56 −8 (c)		~		

To ensure that differences in the magnitude of resting state correlations were not tied to an idiosyncratic rlPFC ROI, additional rlPFC ROIs derived from previous work—(a) Badre et al. (2012), (b) Daw et al. (2006), and (c) Boorman et al. (2009)—were tested to determine whether connectivity with the regions shown in Figure 3 distinguished explorers from nonexplorers. Moreover, to ensure that results were specific to exploration (ε), we tested whether these same regions could distinguish exploiters from nonexploiters (ρ). Where bilateral coordinates were not available, the x coordinate was reflected about the midline to generate a contralateral ROI. Shaded cells within the table indicate where significant results would be expected if the findings displayed in Figure 5 generalized to other regions in lateralized fashion but remained specific to exploration. The number of true exploration-related positives (shaded areas) was significantly greater than that expected by chance ($p = .00002$, binomial theorem), whereas neither exploration-nor exploitation-related false positives occurred more than expected by chance ($p = .19$ and $p = .88$, respectively). For the first two ROIs (coordinates [24 46 20] and [−24 46 20]), results recapitulate the findings of Figure 5 with respect to the exploration parameter. Asterisks indicate $p < .05$; tildes indicate $p < .10$; blank cells specify nonsignificant values.

pursued large rewards without consideration for their frequency. Moreover, in post hoc analyses, it was independent of sensation seeking, as defined by scores on the sensation seeking scale for children (SSSC; Russo et al., 1991). Specifically, there were no differences between explorers and nonexplorers in the total SSSC score (11.3 vs. 12.3, $t(45) = -0.93$, $p = .36$), nor for any of the SSSC subscales (all $ps > .21$). Finally, in exploratory individuals, this exploration was strategic, in that it was driven by relative uncertainty in the reward space, rather than by chance responding. Such exploration may therefore reflect the participation of rlPFC-based corticostriatal circuits necessary for adaptive responding when participants are confronted by uncertain rewards.

Notably, these periadolescent participants as a whole also displayed a probability-magnitude bias in which they chose highly probable rewards over proportionally greater rewards of lower probability. This finding is consistent with some reports that adolescents at this time point are relatively risk averse (Tymula, Rosenberg Belmaker, Ruderman, Glimcher, & Levy, 2013), perhaps more so for losses than for the gains studied here (Wolf, Wright, Kilford, Dolan, & Blakemore, 2013). Alternatively, this bias may partly result from a potential inability of these participants to effectively integrate probability and magnitude to generate an estimate of expected value. This cognitive explanation may align with other evidence that the capacity for abstraction may develop at older ages than the ones studied here (Dumontheil, 2014), but it alone would not explain why participants emphasized probability over magnitude in the choices that they did make. Moreover, this bias is unlikely to reflect a problem with either learning in general or exploration in particular, as Figures 3 and 4 demonstrate an understanding of the task structure and relative uncertainty, respectively.

Because these strategic exploratory behaviors are not yet well evaluated in adolescence, their behavioral and neurophysiological understanding leans heavily on the extant adult work. Across different tasks in adults, including both one-armed bandit tasks and the current one, regions within rlPFC have consistently been shown to respond to exploration of the reward space (Badre et al., 2012; Boorman et al., 2009; Daw et al., 2006). As we have discussed elsewhere (Badre et al., 2012), these similarities are present though the tasks themselves can have considerable differences, including in the quality of exploration itself. In a task in which participants selected between multiple slot machines on each trial (Daw et al., 2006), for example, they behaved as though only the last trial of the task was informative (i.e., the event history was limited), indicating that all unchosen options from the previous trial were equally uncertain on the next one. In contrast, here participants showed clear effects of learning across 50 trials, as reflected in the contribution of previous understanding of the reward space (incorporated into estimates of uncertainty) to trial-by-trial explo-

ration (Figure 4). Despite differences in the computation of the uncertainty that guided exploration in previous tasks, rlPFC activity nonetheless correlated with exploratory behaviors in these data.

A related question concerns the engagement of rlPFC in periadolescent participants at all. Importantly, although exploration has not been well studied in this age group, other tasks that engage rlPFC in adults have been shown to likewise engage this region in adolescence (Dumontheil, 2014). For example, studies of relational integration, the capacity to evaluate relationships between multiple cognitive representations, demonstrate that rlPFC activity in such adolescents correlates with the processing of abstract representations, though this activity becomes more specific for higher-order representations in later adolescence (Wendelken, O'Hare, Whitaker, Ferrer, & Bunge, 2011; Dumontheil, Houlton, Christoff, & Blakemore, 2010). Moreover, our robustness checks (Table 1) indicate that this finding is not limited to one specific rlPFC region. Thus, although rlPFC activity near the onset of adolescence may not demonstrate the specificity of adult activity—likely consistent with studies that show development of both cortical thickness (O'Donnell, Noseworthy, Levine, & Dennis, 2005) and myelination (Miller et al., 2012) within rlPFC into adulthood—such activity has nonetheless been found consistently in this age group.

Given that our current connectivity results generalize to multiple exploration-associated rlPFC regions but remain specific to exploration (as opposed to exploitation), the meaning of a behaviorally defined, lateralized difference in connectivity with the insula and putamen is intriguing. The insula is thought to be important for evaluating uncertainty (Smith et al., 2014; White et al., 2014; Preuschoff et al., 2008; Huettel, 2006) and for integrating interoception with cognition (Craig, 2009), suggesting the hypothesis that exploration may be more strongly coupled to interoceptive awareness in explorers. The location of the insula activation in the current study is likely to be critical; our current activation is more ventral and posterior, where social–emotional and sensorimotor representations may be present (Kurth, Zilles, Fox, Laird, & Eickhoff, 2010). Such representations could conceivably provide important context for exploratory decisions by allowing for moment-by-moment analysis of uncertainty, motivation, and motor state (Craig, 2009). Of course, the insula may also participate in other, distinct cognitive processes, and other brain regions may likewise participate in the processing of uncertainty at different times; more broadly, the problems of such reverse inference are well known (e.g., Poldrack, 2006; but see Hutzler, 2014). In addition to more directly evaluating these possibilities, future work might therefore test how aversion to uncertainty affects exploratory behavior both in this task and more generally, especially when losses are also possible (Payzan-LeNestour & Bossaerts, 2011).

Similarly, connections with the more posterior putamen may be related to the activation of reward-relevant motor plans. Importantly, however, this connectivity would not be directly tied to actual motor performance in this study, given that our imaging data were obtained in the resting state. Because exploration in more ethological settings might be expected to depend on multiple factors, including the current physiological state, the stronger connection between these regions in explorers could instead reflect a history of greater use or greater sensitivity of this pathway in those individuals.

Along those lines, Granger causality analyses suggest that activity within the insula and putamen may more strongly influence rlPFC than the reverse. This finding is interesting in the context of theories arguing that reward-related behaviors in adolescence reflect a mismatch between early maturation of brain regions important for incentive salience and relative immaturity of those important for cognitive control (Gladwin et al., 2011; Galvan, 2010; Somerville & Casey, 2010; Steinberg, 2008). On the basis of our results, activity within these posterior regions may be more likely to provide an input to rlPFC than to indicate the results of top–down influence from this region. Thus, it is conceivable that rlPFC-based exploration in early adolescence is more strongly influenced by immediate consideration of reward rather than prospective evaluation of different paths of action (Donoso et al., 2014).

On a methodological note, the use of resting state data to investigate these questions was based on the hypothesis that exploration represents a somewhat stable cognitive phenotype and should therefore be reflected even in the absence of task. However, the use of resting state data has previously raised concerns in adolescent studies, including the recent widely discussed possibility that many resting state studies in adolescence may be influenced by subject motion (Power et al., 2012). Fortunately, participants in this study were quite still overall—the largest single linear displacement was less than one voxel size (2.3 mm) in 56 of 62 participants—and resting state data were scrubbed (Power et al., 2012) before analysis to reduce the influence of motion. Importantly, neither the mean ($p = .44$) nor the maximum ($p = .31$) motion displacement distinguished explorers from nonexplorers, and there was no significant correlation between either the mean or maximal motion displacement and the strength of connectivity between the rlPFC and posterior putamen/insula across participants (all p values $> .16$). Thus, a movement confound is unlikely to explain our results.

In summary, our data demonstrate that individual differences in strategic exploration have a behavioral and neural correlate at the onset of adolescence. Given these results, our understanding of so-called risky behaviors and adolescent vulnerability to psychiatric disorders may potentially be understood in light of exploration. Of course, additional factors are likely at play: for example,

future work might consider the influence of prospection (the ability to envision the future). More broadly, we hypothesize that puberty marks the beginning of an important inflection point in the developmental trajectory in which brain development may proceed in either a more resilient or more vulnerable direction (Crone & Dahl, 2012). Because this study is limited by the lack of longitudinal developmental data, future work to evaluate changes in strategic exploration over time would be critical, especially given the fact that other task outcomes, such as the probability-magnitude bias, appear to differ between our participants and adults. Furthermore, task-related fMRI results would provide important additional constraints about whether and how the rlPFC contributes to task performance. Nonetheless, this baseline work will hopefully provide predictors when longitudinal data points are obtained, thereby permitting within-subject evaluation of developmental changes in exploration across pubertal maturation, as well as correlation with real-world behaviors. Ultimately, understanding the neural systems and hormonal influences that underlie individual differences in decision-making at the onset of adolescence may have great relevance to understanding potentially pathological states in older adolescents and young adults, as well as advancing the possibility of periadolescence as a window of opportunity for early intervention.

Acknowledgments

REFERENCES

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9,* 357–381.

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron, 73,* 595–607.

Boorman, E. D., Behrens, T. E., Woolrich, M. W., & Rushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron, 62,* 733–743.

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and psychological maturation of decision-making in adolescence and young adulthood. *Journal of Cognitive Neuroscience, 25,* 1807–1823.

Craig, A. D. (2009). How do you feel—Now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10,* 59–70.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience, 13,* 636–650.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441,* 876–879.

Donoso, M., Collins, A. G., & Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science, 344,* 1481–1486.

Dumontheil, I. (2014). Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex. *Development of Cognitive Neuroscience, 10C,* 57–76.

Dumontheil, I., Houlton, R., Christoff, K., & Blakemore, S. J. (2010). Development of relational reasoning during adolescence. *Developmental Science, 13,* F15–F24.

Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition, 72,* 66–72.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 9673–9678.

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience, 12,* 1062–1068.

Galvan, A. (2010). Adolescent development of the reward system. *Frontiers in Human Neuroscience, 4,* 6.

Gladwin, T. E., Figner, B., Crone, E. A., & Wiers, R. W. (2011). Addiction, adolescence, and the integration of control and motivation. *Developmental Cognitive Neuroscience, 1,* 364–376.

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *Journal of Neuroscience, 20,* 2369–2382.

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology, 35,* 4–26.

Huettel, S. A. (2006). Behavioral, but not reward, risk modulates activation of prefrontal, parietal, and insular cortices. *Cognitive, Affective & Behavioral Neuroscience, 6,* 141–151.

Hutzler, F. (2014). Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *Neuroimage, 84,* 1061–1069.

Kayser, A. S., Mitchell, J. M., Weinstein, D., & Frank, M. J. (2015). Dopamine, locus of control, and the exploration–exploitation tradeoff. *Neuropsychopharmacology, 40,* 454–462.

Kayser, A. S., Sun, F. T., & D'Esposito, M. (2009). A comparison of Granger causality and coherency in fMRI-based analysis of the motor system. *Human Brain Mapping, 30,* 3475–3494.

Kelley, A. E., Schochet, T., & Landry, C. F. (2004). Risk taking and novelty seeking in adolescence: Introduction to part I. *Annals of the New York Academy of Sciences, 1021,* 27–32.

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure & Function, 214,* 519–534.

Lebel, C., Walker, L., Leemans, A., Phillips, L., & Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *Neuroimage, 40,* 1044–1055.

Miller, D. J., Duka, T., Stimpson, C. D., Schapiro, S. J., Baze, W. B., McArthur, M. J., et al. (2012). Prolonged myelination in human neocortical evolution. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 16480–16485.

O'Donnell, S., Noseworthy, M. D., Levine, B., & Dennis, M. (2005). Cortical thickness of the frontopolar area in typically developing children and adolescents. *Neuroimage, 24,* 948–954.

Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Computational Biology, 7,* e1001048.

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence, 17,* 117–133.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59–63.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage, 59,* 2142–2154.

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience, 28,* 2745–2752.

Russo, M. F., Lahey, B. B., Christ, M. A., Frick, P. J., McBurnett, K., Walker, J. L., et al. (1991). Preliminary development of a sensation seeking scale for children. *Personality and Individual Differences, 12,* 399–405.

Smith, A. R., Steinberg, L., & Chein, J. (2014). The role of the anterior insula in adolescent decision making. *Developmental Neuroscience, 36,* 196–209.

Somerville, L. H., & Casey, B. J. (2010). Developmental neurobiology of cognitive control and motivational systems. *Current Opinion in Neurobiology, 20,* 236–241.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28,* 78–106.

Tymula, A., Rosenberg Belmaker, L. A., Ruderman, L., Glimcher, P. W., & Levy, I. (2013). Like cognitive function, decision making across the life span shows profound age-related changes. *Proceedings of the National Academy of Sciences, U.S.A., 110,* 17143–17148.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15,* 273–289.

Wendelken, C., O'Hare, E. D., Whitaker, K. J., Ferrer, E., & Bunge, S. A. (2011). Increased functional selectivity over development in rostrolateral prefrontal cortex. *Journal of Neuroscience, 31,* 17260–17268.

White, T. P., Engen, N. H., Sorensen, S., Overgaard, M., & Shergill, S. S. (2014). Uncertainty and confidence from the triple-network perspective: Voxel-based meta-analyses. *Brain and Cognition, 85,* 191–200.

Wolf, L. K., Wright, N. D., Kilford, E. J., Dolan, R. J., & Blakemore, S. J. (2013). Developmental changes in effects of risk and valence on adolescent decision-making. *Cognitive Development, 28,* 290–299.