

# Adolescents Adapt More Slowly than Adults to Varying Reward Contingencies

Amir Homayoun Javadi<sup>1,2\*</sup>, Dirk H. K. Schmidt<sup>1\*</sup>,  
and Michael N. Smolka<sup>1</sup>

## Abstract

■ It has been suggested that adolescents process rewards differently from adults, both cognitively and affectively. In an fMRI study we recorded brain BOLD activity of adolescents (age range = 14–15 years) and adults (age range = 20–39 years) to investigate the developmental changes in reward processing and decision-making. In a probabilistic reversal learning task, adolescents and adults adapted to changes in reward contingencies. We used a reinforcement learning model with an adaptive learning rate for each trial to model the adolescents' and adults' behavior. Results showed that adolescents possessed a shallower slope in the sigmoid curve governing the relation between expected value (the value of the expected feedback, +1 and –1 representing rewarding and punishing feedback, respectively) and probability of stay (selecting the same option as in the previous trial). Trial-by-trial change in expected values after being cor-

rect or wrong was significantly different between adolescents and adults. These values were closer to certainty for adults. Additionally, absolute value of model-derived prediction error for adolescents was significantly higher after a correct response but a punishing feedback. At the neural level, BOLD correlates of learning rate, expected value, and prediction error did not significantly differ between adolescents and adults. Nor did we see group differences in the prediction error-related BOLD signal for different trial types. Our results indicate that adults seem to behaviorally integrate punishing feedback better than adolescents in their estimation of the current state of the contingencies. On the basis of these results, we argue that adolescents made decisions with less certainty when compared with adults and speculate that adolescents acquired a less accurate knowledge of their current state, that is, of being correct or wrong. ■

## INTRODUCTION

A basic function of the brain is to evaluate the motivational and emotional importance of events and to adapt behavior accordingly (Jocham, Klein, & Ullsperger, 2011; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Schultz, 2006). On the basis of behavioral decision theories, decisions are guided by the value assigned to each potential option (Luce, 1959). Reward prediction error signals are used to reflect the difference between the expected value and the actual outcome of an action (O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Schultz, Dayan, & Montague, 1997). "Expected value" is defined as the value of the expected outcome. Positive values indicate expectation of a rewarding feedback and negative values expectation of punishment or loss. To behave adaptively in a changing world, these values must be continuously updated based on experience (Montague, 2006; Montague, Hyman, & Cohen, 2004).

Maturation of the human brain and reorganization of the neuronal structures related to emotional, motivational, and cognitive processes are essential for the establishment of behavioral control, cognitive flexibility, and

efficient brain function. Differences in the pattern of development of various brain areas and circuits have been proposed to lead to an "imbalance" in the adolescent brain (Casey, Jones, & Hare, 2008; Gogtay et al., 2004). Specifically, the subcortical brain circuitries and the frontal, cortical circuitries show a lead-lag gradient of maturation (Casey, Jones, et al., 2008; Steinberg, 2005), with subcortical processes developing earlier and reaching maturation already in adolescence, whereas the development of cortical frontal processes is much more protracted and reach maturation only in emerging adulthood.

One consequence of this is that adolescents engage in increased risky decision-making compared with other age groups, because they place greater value on the potential positive (as opposed to negative) consequences of risk-taking (Steinberg, 2010; Casey, Getz, & Galvan, 2008; Ernst, Pine, & Hardin, 2006). Brain imaging studies that focused on the developmental aspects of reward processing offered different explanations for risky adolescent behavior. On the one hand, it was hypothesized that lower activation (i.e., hyposensitivity) in the reward system of adolescents (compared with adults) may lead to more extensive reward seeking (Spear, 2000). On the other hand, higher activation (i.e., hypersensitivity) in the reward system has been hypothesized to lead to an increase in risk taking behavior (van Leijenhorst, Moor, et al., 2010; Galvan,

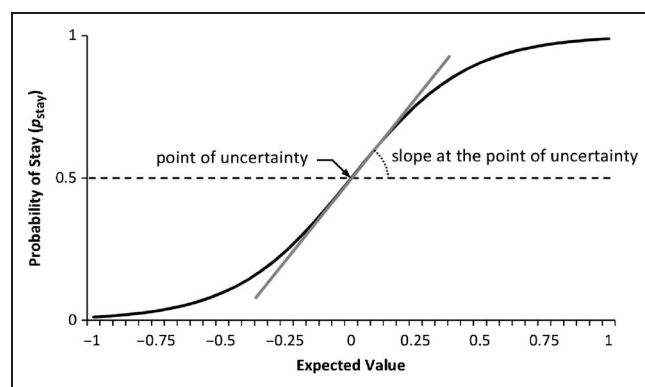
<sup>1</sup>Technische Universität Dresden, <sup>2</sup>University College London  
\*These authors contributed equally to the study.

Hare, Voss, Glover, & Casey, 2007). Bjork, Smith, Chen, and Hommer (2010) and Bjork et al. (2004) found the adolescents' reward system (especially the ventral striatum [VS]) to be hyposensitive compared with adults. Others found hypersensitivity of the VS (Galvan & McGlennen, 2013; Cohen et al., 2010; van Leijenhorst, Zanolie, et al., 2010; Galvan et al., 2006; Ernst et al., 2005). As for adults, it has been shown that they are not only adequately sensitive but also able to exert control over impulsive tendencies (Ripke et al., 2012; Cohen et al., 2010). Using a deterministic reversal learning task, van der Schaaf, Warmerdam, Crone, and Cools (2011) found that overall performance increases from age 10 to 25. Interestingly, punishment-based learning was best for the youngest age group, whereas reward-based learning was best in young adults.

The goal of this study was to investigate age-related differences in the behavioral effect and neural processing of rewarding and punishing feedback. Efficient processing of feedback is necessary for decision-making and, more importantly, for adaptive behavior in a changing environment.

We used a probabilistic reversal learning task to study how adolescents adapt to changes of reward contingencies, as well as how they deal with uncertainty in the system. We modeled adolescents' and adults' behavior, using a reinforcement learning method to compare their modeling parameters to achieve a better understanding of the underlying mechanisms of possible behavioral differences both groups.

In our model each decision is governed by a sigmoid curve, which relates reward expectation (expected value) and likelihood of behavioral stay ( $p_{\text{stay}}$ , selecting the same option in the subsequent trial). Figure 1 shows this curve with expected value spanning over  $[-1 \dots +1]$ , representing 100% punishment and 100% reward for the option chosen before in the two ends of the plot. Indifference or the uncertainty point is the point at which there is no difference between options, where  $p_{\text{stay}} = 0.5$ . The slope at this point indicates how one integrates expected values to make decisions with more certainty in subsequent trials, that is, making decisions with  $p_{\text{stay}}$  values smaller or greater than 0.5. In other words, the slope shows



**Figure 1.** Sigmoid curve that relates expected value and likelihood of behavioral stay, showing the point of uncertainty and slope at that point.

how fast one crosses the uncertainty point (toward either  $p_{\text{stay}} = 1$  or  $p_{\text{stay}} = 0$ ), that is, a higher slope corresponds to a faster passage of the uncertainty point and vice versa.

Regarding the neural correlates of parameters derived from such reinforcement learning algorithms, it has previously been shown that BOLD activity of the dorsal ACC (dACC) is correlated with learning rate (Krugel, Biele, Mohr, Li, & Heekeren, 2009; Behrens, Woolrich, Walton, & Rushworth, 2007; Klein et al., 2007), the VS with prediction error (Gläscher, Hampton, & O'Doherty, 2009; Hampton, Bossaerts, & O'Doherty, 2006), and the ventromedial pFC (vmPFC) with expected value (Gläscher et al., 2009; Hampton et al., 2006). Although it has to be acknowledged that other brain areas, such as the lateral orbital frontal cortex, the dorsolateral pFC, and the anterior insula are involved in reversal learning (Xue et al., 2013; Remijne, Nielen, Uylings, & Veltman, 2005), we focused on VS, dACC, and vmPFC, as combined signals from these three regions are reported to be predictive of behavior (Hampton & O'Doherty, 2007), which we expect to be different across age groups.

Given the work of van der Schaaf et al. (2011), we hypothesized that adolescents would show a lower performance during the task and a higher sensitivity to punishments, compared with adults. Regarding the applied reinforcement learning algorithm, we expected lower certainty and, consequently, a shallower slope in their decision curve. Further to this, we investigated the correlation of modeling parameters with BOLD brain activity and explored whether age related differences can be observed.

## METHODS

### Participants

The data set used in this study was part of the "Adolescent Brain" project, funded by the German Federal Ministry of Education and Research (BMBF). This project is a longitudinal study investigating the relationship between brain development and susceptibility to substance use disorders, involving two assessments over 4 years (Ripke et al., 2012).

Two hundred sixty adolescents were recruited from local secondary schools. We had to exclude 42 adolescents from the analysis because of excessive head movements (movements greater than 3 mm in any one direction), interruptions in scanning, faults in data transfer, or missing data. The remaining 218 adolescents (115 boys (52.75%), age range = 14–15 years, mean age = 14.61 years ( $SD = 0.32$ )) were included in the analysis. As a control group, we recruited 29 adult participants by board and Internet announcements (17 men (58.62%), age range = 20–39 years, mean age = 25.24 years ( $SD = 6.34$ )). Adolescents were screened with a structured, diagnostic interview "development and well-being assessment" (Goodman, Ford, Richards, Gatward, & Meltzer, 2000) according to the fourth edition of the *Diagnostic and Statistical Manual*



similar level of understanding. In the first phase of the training session, the rule for system change was implemented, but participants were provided with deterministic feedback. This means that they were always rewarded after correct responses and punished after wrong responses. The criterion to finish this phase was three system changes. In the second phase, participants were introduced to probabilistic feedback, without system changes. The criterion to finish this phase was to select the better option 10 times consecutively. The third phase combined probabilistic feedback with system changes. This phase was similar to the main task in the scanner. The criterion to finish this phase was to achieve three system changes. See Figure 2B for the procedure of the session.

Participants were instructed to maximize their gains. They were informed that, in addition to a fixed amount of €5, they would receive any extra money they accumulated at the end of the study. The duration of the task was 26 min.

### Computational Modeling

We used a similar model as described in Krugel et al. (2009) to model participants' behavioral choices. We considered a sigmoid curve (Equation 6), indicating the relation between difference of expected values for the two options,  $v_a(t)$  and  $v_b(t)$  for options  $a$  and  $b$ , respectively, to calculate the probability of the selection of each option,  $p_a(t + 1)$  and  $p_b(t + 1)$ . On the basis of these probabilities, we defined probability of behavioral stay ( $p_{\text{stay}}$ ), that is, selecting the same option in the current trial as the previous trial (Equation 8). We constructed the sigmoid curve based on the difference of expected values,  $v_a(t) - v_b(t)$ , and  $p_{\text{stay}}$ . We chose difference of expected values instead of expected value for each option,  $v_a$  and  $v_b$ , and  $p_{\text{stay}}$  instead of the probability of selection of that option ( $p_a$  and  $p_b$ ). Difference of expected values and  $p_{\text{stay}}$  combine  $v_a$  and  $v_b$  into a uniform parameter that is indifferent to the options per se.

$$v_{\text{selected}}(t) = \begin{cases} v_a(t) & \text{if option } a \text{ is selected} \\ v_b(t) & \text{if option } b \text{ is selected} \end{cases} \quad (1)$$

in which  $v_a(t)$  and  $v_b(t)$  show expected value on trial  $t$  for the two options  $a$  and  $b$ , namely circle and square.

$$\delta(t) = \text{reward}(t) - v_{\text{selected}}(t) \quad (2)$$

in which  $\delta(t)$  shows the prediction error and  $\text{reward}(t)$  shows reward, for trial  $t$ .

$$d_v(t) = \alpha(t) \cdot \delta(t) \quad (3)$$

$$v_{\text{selected}}(t + 1) = v_{\text{selected}}(t) + d_v(t) \quad (4)$$

in which  $\alpha(t)$  is the adaptive learning rate (see below).  $d_v(t)$  represents change of expectation. After each decision

the expected value for the two options were updated as follows:

$$\begin{cases} v_a(t + 1) = v_{\text{selected}}(t + 1) & \text{if option } a \text{ is selected} \\ v_b(t + 1) = v_b(t) & \\ v_b(t + 1) = v_{\text{selected}}(t + 1) & \text{if option } b \text{ is selected} \\ v_a(t + 1) = v_a(t) & \end{cases} \quad (5)$$

Subsequently the probability of selecting options  $a$  and  $b$  were calculated as follows:

$$p_a(t + 1) = \frac{1}{1 + \exp(-\gamma \cdot (v_a(t) - v_b(t)))} \quad (6)$$

$$p_b(t + 1) = 1 - p_a(t + 1) \quad (7)$$

where  $\gamma$  is the slope of the sigmoid curve, considered as the sensitivity parameter determining the influence of reward expectations on choice probabilities.

$p_{\text{stay}}(t + 1)$  and  $p_{\text{switch}}(t + 1)$  were calculated as follows:

$$\begin{cases} p_{\text{stay}}(t + 1) = p_a(t + 1) & \text{if option } a \text{ is selected} \\ p_{\text{stay}}(t + 1) = p_b(t + 1) & \text{if option } b \text{ is selected} \end{cases} \quad (8)$$

and

$$p_{\text{switch}}(t + 1) = 1 - p_{\text{stay}}(t + 1) \quad (9)$$

Because traditional approaches using constant learning rate do not allow for fast adaptation after the occurrence of a reversal, nor do they allow for stabilization of behavior once the best option is found, we used an adaptive learning rate (Krugel et al., 2009).  $\alpha(t)$  was updated as follows, where  $f(m)$  is a mapping function to ensure that  $\alpha(t)$  values are maintained in the range of ]0..1[ ,  $m(t)$  is the normalized value of first derivation of  $\delta(t)$  and  $\delta_{\text{abs}}(t)$  is the smoothed, unsigned value of  $\delta(t)$ .

$$\delta_{\text{abs}}(t) = \delta_{\text{abs}}(t - 1) \cdot (1 - \alpha(t)) + |\delta(t)| \cdot \alpha(t) \quad (10)$$

$$m(t) = \frac{\delta_{\text{abs}}(t) - \delta_{\text{abs}}(t - 1)}{(\delta_{\text{abs}}(t) + \delta_{\text{abs}}(t - 1))/2} \quad (11)$$

$$f(m(t)) = \text{sign}(m(t)) \cdot (1 - \exp(-(m(t)/\beta)^2)) \quad (12)$$

$$\alpha(t+1) = \begin{cases} \alpha(t) + f(m(t)) \cdot (1 - \alpha(t)) & \text{if } m(t) > 0 \\ \alpha(t) + f(m(t)) \cdot \alpha(t) & \text{if } m(t) < 0 \end{cases} \quad (13)$$

where  $\beta$  is a modulatory factor to which the derivation of  $\delta(t)$  affects  $\alpha(t + 1)$ .

Finally  $\gamma$ ,  $\alpha(1)$  and  $\beta$  were the three parameters that needed to be optimized using the logarithm of likelihood



of fit ( $\log L$ ).  $L$  represents how accurately the model can predict participants' behavior in a subsequent trial. We used the following formula to calculate  $L$ , where  $i$  represents trial number and  $n$  represents total number of trials ( $n = 120$ ).

$$L = \sum_{i=2}^n B_{i,\text{switch}} P_{i,\text{switch}} / \sum_{i=2}^n B_{i,\text{switch}} + \sum_{i=2}^n B_{i,\text{stay}} P_{i,\text{stay}} / \sum_{i=2}^n B_{i,\text{stay}}. \quad (14)$$

Figure 3 shows modeling of a sample session for choices, reward, and modeling parameters.

## Statistical Analysis

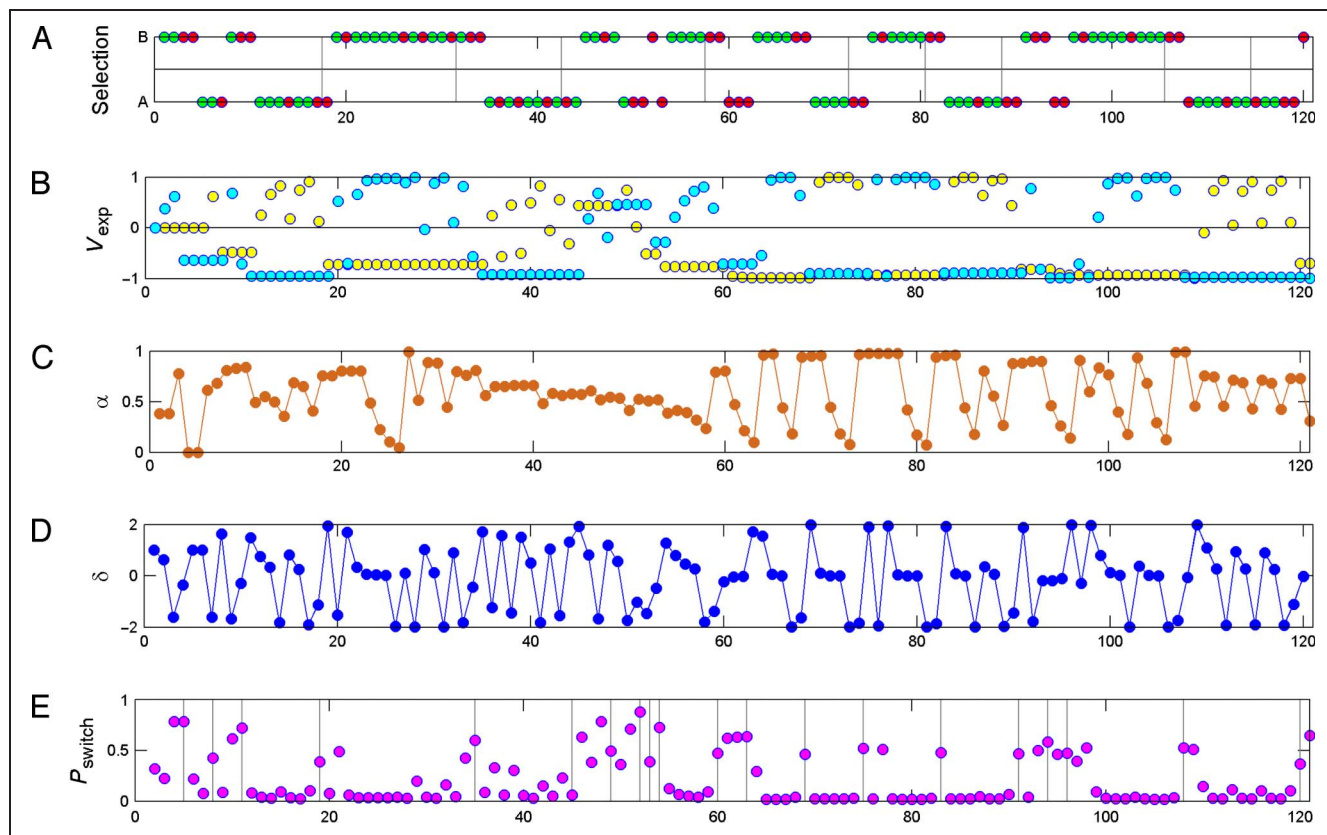
### Behavioral Measures

We compared the ratio of correct responses using an independent sample  $t$  test and the difference in the number of system changes between adolescents and adults using non-parametric Mann–Whitney  $U$  test. We

also analyzed effects on the switching rate, using a  $2 \times 2 \times 2$  mixed-factorial ANOVA with Response (correct/wrong) and Feedback (reward/punishment) as within-subject factors and Group (adults/adolescents) as between-subject factor. Subsequently, we compared switching rates of adolescents and adults in all four types of trials using independent sample  $t$  tests.

### Modeling Measures

Two sets of parameters were estimated in our models: the ones that model the behavior as a whole (learning rate for the first trial  $\alpha(1)$ , modulatory factor  $\beta$ , logarithm of the slope of the sigmoid curve  $\gamma$ , and logarithm of likelihood of fit  $L$ ) and the ones that model the behavior on each trial (learning rate  $\alpha$ , change of expected value  $d_v$ , and prediction error  $\delta$ ). The former set of parameters ( $\alpha(1)$ ,  $\beta$ ,  $\log \gamma$ , and  $\log L$ ) was subjected to independent sample  $t$  tests with group as the independent factor. The latter set of parameters ( $\alpha$ ,  $d_v$ , and  $\delta$ ) was subjected to three  $2 \times 2 \times 2$  mixed-factorial ANOVAs with Response (correct/wrong) and Feedback (reward/punishment) as within-subject factors and Group (adults/adolescents) as



**Figure 3.** (A) Selected task option. A and B represent the two options, square and circle, respectively. Red color indicates punishment, and green indicates reward. Vertical lines indicate trials in which a system change has occurred. As it is clear from the figure, each system change was preceded with at least four consecutive selections of the correct option, regardless of possible negative feedback. (B) Expected value for option A, yellow circles, and B, cyan circles. As shown, expected value for an option changes only when that option is selected. Its value increased with positive feedbacks. (C) Adaptive learning rate ( $\alpha$ ). (D) Prediction error is defined as the difference between reward and expected value,  $\delta(t) = \text{reward}(t) - v_{\text{selected}}(t)$ . (E) Probability of switch as calculated by the model. Vertical lines indicate trials in which a behavioral switch has occurred.

between-subject factor. Subsequently, Bonferroni-corrected independent sample *t* tests were used for post hoc comparisons. Data were checked for normality of distribution using the Kolmogorov–Smirnov test.

It should be mentioned that SPSS controls for highly imbalanced group sizes in independent two-sample *t* tests. The standard two-sample *t* test allows the sample sizes to be different (Press, Teukolsky, Vetterling, & Flannery, 2007). The sample variance is estimated by combining the sample variances from each group. Importantly, each is weighted by the number of samples in the group. So, in this sense, the standard *t* test already accommodates differences in sample size. A similar argument applies to ANOVAs. Variances were different between adolescents and adults; therefore, we report the result of tests with the assumption of inequality of variance. The distributions of *p* values for post hoc tests for each group of analyses were corrected for multiple comparisons according to the false discovery rate (FDR) procedure (Benjamini & Hochberg, 1995). We computed a *q* threshold for four comparisons per group that set the expected rate of false discoveries to 0.025 for  $q^* = 0.050$ .

### Image Acquisition

All MRI data were acquired at the Neuroimaging Centre at the Technische Universität Dresden, using a 3.0-T scanner (Magnetom Tim Trio, Siemens, Erlangen, Germany). Series of T<sub>2</sub>\*-weighted, EPIs with 42 transverse slices, tilted approximately 30° toward the coronal beyond the anterior and posterior commissure lines, with a 3-mm in-plane resolution and a slice thickness of 2 mm (1-mm gap resulting in a voxel size of 3 × 3 × 3 mm<sup>3</sup>), a field of view of 192 × 192 mm<sup>2</sup>, a flip angle of 80°, a repetition time of 2410 msec, a bandwidth of 2112 Hz/pixel, and an echo time of 25 msec, were acquired. The first 3 volumes were discarded to allow the magnetization to reach equilibrium. High-resolution three-dimensional anatomical images were acquired using a T<sub>1</sub>-weighted, magnetization-prepared, rapid acquisition gradient-echo sequence with a field of view of 256 × 224 mm<sup>2</sup>, 176 slices, a voxel size of 1 × 1 × 1 mm<sup>3</sup>, a repetition time of 1900 msec, an echo time of 2.26 mm, and a flip angle of 9°.

### Imaging Data Analysis

Imaging data analysis was done using SPM5 (Wellcome Trust, London, UK). Data were preprocessed to correct for slice timing and head motion, spatially normalized to a standard EPI template in MNI space and smoothed (8 mm FWHM isotropic Gaussian kernel). Templates were based on the MNI305 stereotaxic space (Cocosco, Kollokian, Remi, Pike, & Evans, 1997), an approximation of Talairach space (Talairach & Tournoux, 1988).

Following Gläscher et al. (2009) and Krugel et al. (2009), three binary and three parametric regressors of interest were specified. Binary regressors were convolved

with a canonical hemodynamic response function and modulated by respective parameters ( $\alpha$ ,  $\nu$ , and  $\delta$ ). Specifically we specified regressors for the response event (1 sec before the response until button press) modulated with the expected value ( $\nu$ ), the learning event (1 sec after onset of feedback for 1 sec) modulated with learning rate ( $\alpha$ ; Krugel et al., 2009), and the feedback event (from onset of feedback for 1 sec) modulated with prediction error ( $\delta$ ; Gläscher et al., 2009). Please note, however, that we did not split up the positive and negative prediction errors as in Krugel et al. (2009).

Additionally, we also conducted a similar first-level model with 12 regressors. These regressors were combinations of 3 parameters (learning rate/expected value/prediction error) × 2 response (correct/wrong) × 2 feedback (rewarded/punished). All these regressors were modulated by respective parameters ( $\alpha$ ,  $\nu$ , and  $\delta$ ) and convolved with a canonical hemodynamic response function. The parametric modulators were all corrected to achieve zero mean. This resulted into two sets of beta images, with slope representing correlation and the interception representing mean. In addition, the six scan-to-scan motion parameters produced during realignment were included to account for residual motion effects. These were fitted to each voxel individually using a standard general linear model (GLM).

To explore the neural correlates of changes in reinforcement learning parameters at the second level, we ran three 1-sample *t* tests using the respective first-level contrasts, condition against baseline, capturing the correlation of  $\alpha$ ,  $\nu$ , and  $\delta$  with brain activity. To compare adolescents' and adults' brain BOLD activity, we ran three independent sample *t* tests, using the same first-level contrasts and Group (adults/adolescents) as between-subject factors. Finally, we ran six 2 (Group: adolescents/adults) × 2 (Response: correct/wrong) × 2 (Feedback: rewarded/punished) mixed factorial ANOVAs, with the contrast reflecting the correlation (slope) and mean (intercept) of  $\alpha$ ,  $\nu$ , and  $\delta$  for the respective trial type. We report activations in the corresponding ROI when  $p < .05$  (small volume-corrected FDR) and with a minimum number of  $k = 10$  voxels in a cluster.

For small volume correction, three ROIs were specified based on probabilistic maps that are freely available online (Nielsen & Hansen, 2002). We made three binary images using a threshold value of 0.5 on the dorsal part of ACC (referred to as dACC), the VS, and the ventromedial part of the pFC (referred to as vmPFC).

## RESULTS

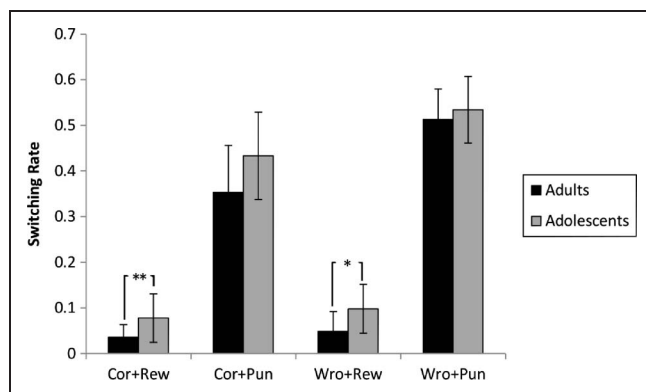
### Behavioral and Modeling

An independent sample *t* test showed no significant differences in task performance between groups, according to the ratio of correct responses (adolescents mean (*SD*) = 0.59 (0.07), adults 0.61 (0.06),  $t(42.653) = 1.292$ ,  $p = .203$ ).

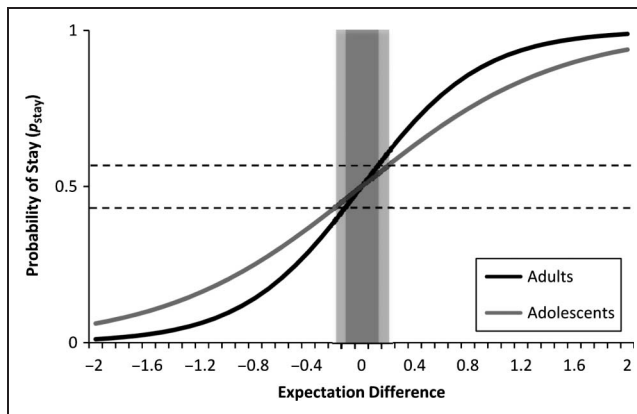
On the other hand, a nonparametric Mann–Whitney  $U$  test revealed that the number of system changes for adults was significantly higher compared with adolescents (median adolescents 6, adults 7,  $Z = -2.04$ ,  $p = .04$ ).

The  $2 \times 2 \times 2$  mixed-factor ANOVA revealed that adolescents switched choices from one trial to the next more frequently compared with adults (significant main effect of Group; adolescents 0.28 (0.10), adults 0.23 (0.10),  $F(1, 245) = 5.729$ ,  $p = .017$ ). This test showed significant three-way interaction of Group, Feedback, and Response,  $F(1, 245) = 4.169$ ,  $p = .042$ . Post hoc  $t$  tests comparing switching rates of adolescents and adults in all four conditions of Response  $\times$  Feedback showed a significant higher switching rates in the case of correct-rewarded,  $t(59.591) = 3.328$ ,  $p = .002$ , and wrong-rewarded trials in adolescents,  $t(40.592) = 2.569$ ,  $p = .014$ , and nonsignificant differences in the case of correct-punished,  $t(34.824) = 1.983$ ,  $p = .055$ , and wrong-punished,  $t(37.598) = 0.812$ ,  $p = .422$  (Figure 4).

Independent sample  $t$  tests showed no significant difference for  $\alpha(1)$  (adolescents 0.307 (0.251), adults 0.286 (0.179),  $t(44.228) = 0.578$ ,  $p = .567$ ) and no significant difference for  $\beta$  (adolescents 1.654 (1.177), adults 1.825 (1.337),  $t(34.026) = 0.654$ ,  $p = .518$ ). Similar  $t$  tests showed a highly significant difference in  $\log\gamma$  between the two groups, with adults achieving a higher value (adolescents 0.137 (0.311), adults 0.330 (0.342),  $t(34.456) = 2.847$ ,  $p = .007$ ). Figure 5 shows the decision curve for adolescents and adults. We should emphasize that, contrary to Figure 1, which shows reward expectation, Figure 5 shows expectation difference: the difference between the expected reward of the selected and unselected options. Expectation difference spans over  $[-2 \dots +2]$ , with 100% expectation of receiving reward for one option and 100% expectation of receiving punishment for the other option placed at either end of the curve. Logarithm of likelihood of fit ( $\log L$ ) was significantly different between adults and adolescents,  $t(33.667) = 3.031$ ,



**Figure 4.** Switching rates for adolescents and adults for different trial and response conditions. Switching rate reflects the ratio of behavioral switch to the total number of trials. Error bars reflect one standard deviation ( $SD$ ). Cor = Correct; Wro = Wrong; Rew = Rewarded; Pun = Punished.  $*p = .014$ ,  $**p = .002$ .



**Figure 5.** Decision curve used in the computational modeling showing shallower slope at  $p_{\text{stay}} = 0.5$  for adolescents when compared with adults. Shaded areas show uncertainty area for adolescents (lighter) and adults (darker). See Discussion for further explanation. Expectation difference shows the difference between the expected value of the selected and unselected options in any given trial. Upper and lower dashed lines show  $p_{\text{uncertainty, upper}}$  and  $p_{\text{uncertainty, lower}}$ , respectively.

$p = .005$ , with a better fit for adults ( $-0.481$  (0.085)) compared with adolescents ( $-0.531$  (0.071)).

A  $2 \times 2 \times 2$  mixed-factorial ANOVA with Response and Feedback as within-subject factors and Group as a between-subject factor on  $\alpha$  showed no significant difference for any of the comparisons ( $F < 1$ ). In contrast, two  $2 \times 2 \times 2$  mixed-factorial ANOVAs on  $d_v$  and  $\delta$  showed a significant effect of Response and Feedback, two-way interaction of Response and Group, and three-way interaction of Response, Feedback, and Group for both  $d_v$  and  $\delta$ , as well as a significant two-way interaction of Response and Feedback for  $d_v$ . The results of these ANOVAs are summarized in Table 1.

Independent sample  $t$  tests on the interaction of response, feedback, and group showed a significant difference between adolescents and adults for the wrong-punished condition, with adults having a smaller  $d_v$  ( $t(36.483) = 2.333$ ,  $p = .025$ ). No other comparison was significant ( $p > .145$ ). Figure 6A shows the change of expected values for all the post hoc comparisons.

Post hoc independent sample  $t$  tests on the interaction of response, feedback, and group showed a near-to-significant difference between adolescents and adults for the correct-punished condition, with adolescents having a smaller  $\delta$  ( $t(33.821) = 2.284$ ,  $p = .029$ ). No other comparison was significant ( $p > .225$ ). Figure 6B shows  $\delta$  values for all the post hoc comparisons.

## Brain Imaging

For the whole sample, we found that the trial-by-trial time course of  $\alpha$  was correlated with the BOLD response of the dACC,  $v$  was correlated with activity of the vmPFC, and activity of the VS reflected  $\delta$  (Figure 7; Krugel et al., 2009; Hampton et al., 2006). Independent sample  $t$  tests on the trial-wise correlation of  $\alpha$ ,  $v$ , and  $\delta$  with BOLD data

**Table 1.** Summary of  $2 \times 2 \times 2$  Mixed-factorial ANOVA with Response and Feedback as Within-subject Factors and Group as Between-subject Factor on Change of Expectation ( $d_v$ ) and Prediction Error ( $\delta$ )

Effect	$d_v$		$\delta$	
Main effect of Response	$F(1, 245) = 76.667$	$p < .001$	$F(1, 245) = 89.886$	$p < .001$
Main effect of Feedback	$F(1, 245) = 2330.9$	$p < .001$	$F(1, 245) = 18179$	$p < .001$
Main effect of Group	$F(1, 245) = 1.054$	$p = .306$	$F(1, 245) = 0.476$	$p = .491$
Interaction of Response and Feedback	$F(1, 245) = 8.512$	$p = .004$	$F(1, 245) = 2.338$	$p = .128$
Interaction of Feedback and Group	$F(1, 245) = 0.378$	$p = .539$	$F(1, 245) = 1.144$	$p = .286$
Interaction of Response and Group	$F(1, 245) = 3.508$	$p = .062$	$F(1, 245) = 3.135$	$p = .078$
Interaction of Response, Feedback, and Group	$F(1, 245) = 9.366$	$p = .002$	$F(1, 245) = 5.083$	$p = .025$

showed nonsignificant differences between adults and adolescents.

Three full-factorial GLM (with group as a between-subject factor and feedback and response as within-subject factors) on the correlation of  $\alpha$ ,  $v$ , and  $\delta$  with brain response did not show any significant main effect of Group or three-way interaction of Group  $\times$  Feedback  $\times$  Response. Three complimentary full-factorial GLM on the mean brain response (intercepts) of  $\alpha$ ,  $v$ , and  $\delta$  during the different trial types also showed no significant main effect of group or three-way interaction. Furthermore, a post hoc  $t$  test on the mean  $\delta$  in the VS showed nonsignificant differences

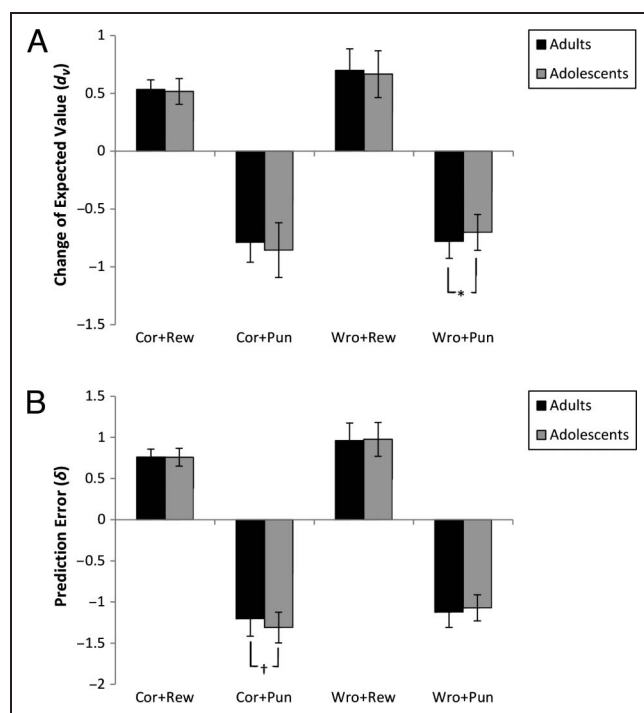
between both groups (adults/adolescents) in correct-punished trials.

## DISCUSSION

Reinforcement learning modeling has been used to investigate the underlying brain areas in decision-making (Krugel et al., 2009; Hampton et al., 2006). In contrast, we used it to achieve a better understanding of the contributing factors underlying behavioral differences in decision-making between adolescents and adults. On the basis of behavioral data that showed that adolescents switched more often than adults ( $p = .02$ ) and achieved a lower number of system changes (change of contingencies;  $p = .04$ ), we hypothesized that adolescents performed the task with lower certainty and consequently possessed a shallower slope in their decision-making curve.

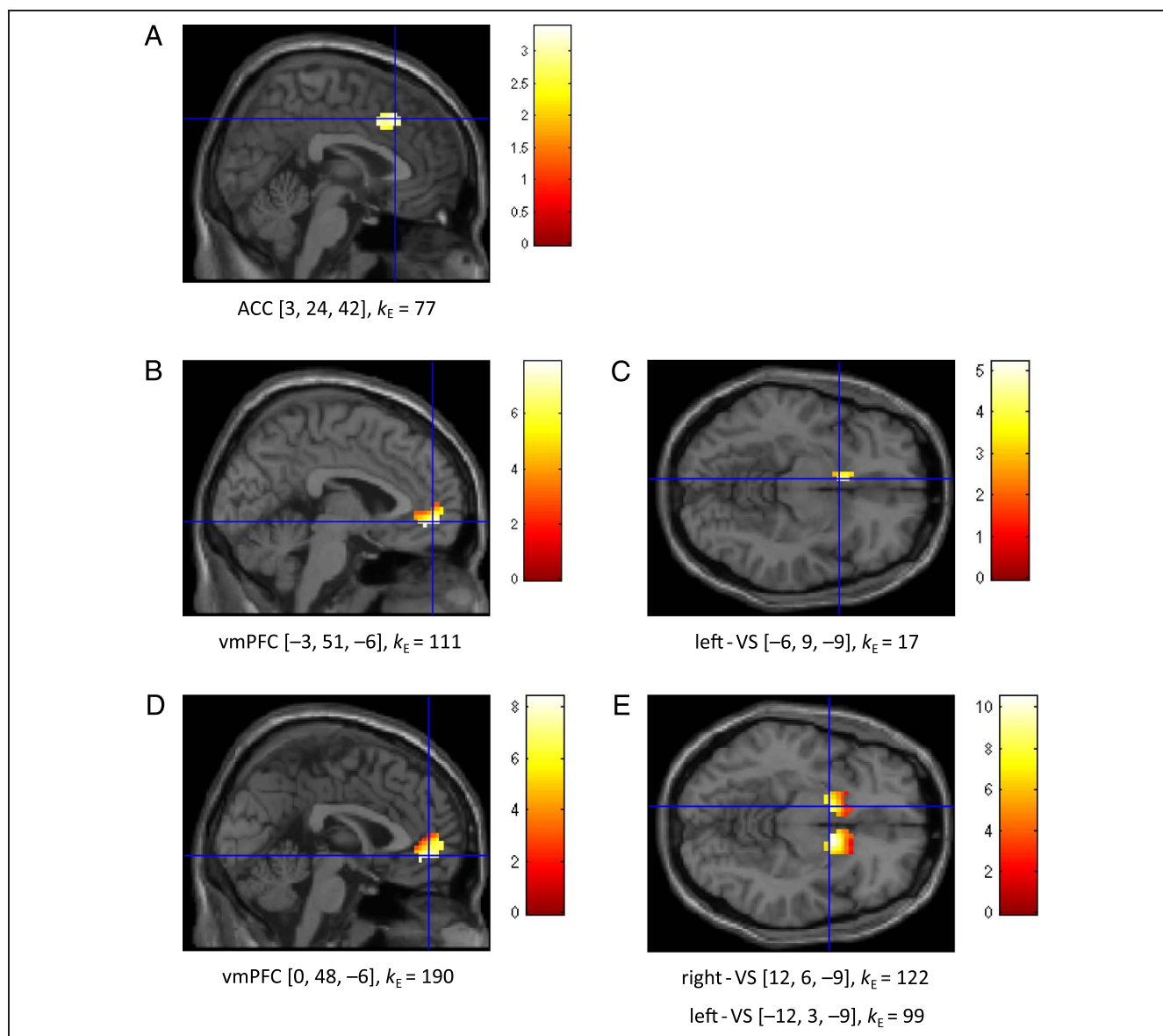
Our results are in line with our hypothesis. We defined  $p_{\text{stay}} = 0.5$  as the uncertainty point and considered slope at this point as the rate of transition from the uncertainty point toward a more certain area ( $p_{\text{stay}} = 1$  or  $p_{\text{stay}} = 0$ ). An alternative way is to define an uncertainty area. We can define the uncertainty area as the range of expectation difference values that correspond to  $p_{\text{stay}}$  values as  $p_{\text{uncertainty, lower}} < p_{\text{stay}} < p_{\text{uncertainty, upper}}$ . This range is shown as shaded bars in Figure 5. Because adolescents showed a shallower slope in their decision curve, they achieve a wider uncertainty range (lighter shading). This wider range of uncertainty can be interpreted as reduced decisiveness, that is, adolescents made decisions with lower certainty, compared with adults.

We investigated the correlation of BOLD activity with modeling parameters  $\alpha$ ,  $v$ , and  $\delta$ . In line with previous literature (Krugel et al., 2009; Hampton et al., 2006), our results showed that BOLD activity in the VS, dACC, and vmPFC is correlated with learning rate, expected value, and prediction error, respectively. Comparing the correlation of the three model parameters with BOLD signal between adolescents and adults showed no difference in the VS, dACC, and vmPFC. Moreover, no differences were found regarding the neural correlates of these parameters



**Figure 6.** (A) shows change of expected value ( $d_v$ ) and (B) shows prediction error ( $\delta$ ) for the three-way interaction of group, response, and punishment (rewarded/punished). Error bars reflect one standard deviation (SD). Cor = Correct; Wro = Wrong; Rew = Rewarded; Pun = Punished. \* $p = .025$ , † $p = .029$ .





**Figure 7.** Masked brain images showing the correlation of the BOLD activity of the adult and adolescent groups ( $p < .05$  small volume-corrected FDR with minimum number of  $k = 10$  voxels in a cluster) with (A) dynamic learning rate ( $\alpha$ ), (B and C) expected value ( $v$ ), and (D and E) prediction error ( $\delta$ ).  $k_E$  represents the number of voxels in a cluster. Coordinates refer to the peak voxel for each cluster.

during the four different trial types (correct-rewarded/correct-punished/wrong-rewarded/wrong-punished). Taken together, these results indicate that task-related brain activity does not or only slightly differs between adolescents and adults and that learning mechanisms in adolescents and adults are quite similar and therefore recruit similar brain regions.

In addition to our predictions, correlation of BOLD activity with prediction error was not limited to VS but was also found in the vmPFC. This is in line with the findings of Hampton et al. (2006). We also found a weak correlation in the VS with expected value. Correlation of BOLD activity with expected value is also reportedly not limited to the vmPFC. Gläscher (2009) and Hampton et al. (2006) showed that the amygdala's BOLD activity is correlated with expected value. We argue that finding prediction error

and expected value parameters to be correlated with BOLD activity in identical brain regions might either be because of an intercorrelation of dependent model parameters or because of correlations in regressors caused by the relatively rapid timing of events in our design.

The modeling fit, as measured by  $\log L$ , was significantly worse for adolescents than for adults. One might speculate that the differences in modeling parameters are merely the result of difference in model fit. We argue that although the degree of fit was different, the three modeling parameters were calculated with equal accuracy, as shown by the similarity of adolescents' and adults' correlation analysis of brain BOLD activity. Therefore, the difference in model fit can be interpreted as a result of the difference in predictability of adolescents' and adults' behavior, demonstrated by a higher rate of behavioral

switch in adolescents and a lower number of system changes, which we interpret as a higher level of uncertainty in adolescents. This behavioral difference is captured by the difference in slope of decision curves.

There is a strong agreement that dramatic behavioral changes during adolescence are driven by differences in reward processing and sensitivity (Somerville, Jones, & Casey, 2010; Steinberg, 2005; Dahl, 2004; for a review, see Blakemore & Robbins, 2012; Galvan, 2010). Although the interaction effect of feedback and group was not significant, the three-way interaction effect of response, feedback, and group was significant. Post hoc tests on this three-way interaction showed interesting results: first, adults achieved a smaller absolute value of prediction error for being punished after trials which they responded correctly to, and second, they achieved a higher absolute value of change in expectation for being punished after trials which they responded wrongly to. The former finding shows that adults were more capable of interpreting negative feedback as either leading or misleading and therefore had more accurate expectations. The latter finding, on the other hand, shows that they incorporated punishment when updating their state to a greater extent when they felt like they were mistaken. Has to be noted that the sample sizes were different, as was the variance of the two samples; hence, the adult group results are likely less stable than the adolescent group results.

Galvan et al. (2006) and Ernst et al. (2005) showed that adolescents are hypersensitive to reward, whereas Bjork et al. (2004) showed a hyposensitivity. Inconsistency in the findings might be because of task design and the developmental stage of the adolescents recruited. Cohen et al. (2010) argued that enhanced prediction error signal leads to adolescents' reward-seeking behavior. Our modeling results showed no difference between the two groups in response to rewarding feedback (no differences in post hoc comparisons on rewarding feedback on the interaction of feedback, response, and group). In contrast, we found significant differences in the response to punishing feedback after being wrong (difference in the change of expected value) and after being correct (difference in prediction error). Another reason for this inconsistency might be our choice of age range for adults. This range is not always consistent between studies (Blakemore & Robbins, 2012). For example, in some studies, the adult group is within our selected range (20–39 years old), and in other studies this range is higher. For instance, the adult age range for Chein, Albert, O'Brien, Uckert, and Steinberg (2011) was 24–29 years, for Jarcho et al. (2012) it was 23–40 years, and for Vaidya, Knutson, O'Leary, Block, and Magnotta (2013) it was 26–30 years old. To further investigate the effect of age in the adults group, we ran similar three full-factorial GLM (with Group as a between-subject factor and Feedback and Response as within-subject factors) on the correlation of  $\alpha$ ,  $\nu$ , and  $\delta$  with brain response in adults older than

24 years ( $n = 14$ ) and adolescents. These analyses showed no significant three-way interaction of the three factors of Group, Response, and Feedback, even with  $p < .01$  uncorrected and  $k = 5$ . These results, however, might be because of the small number of participants in the adults group.

Appropriate weighting and interpretation of both rewards and punishments are crucial for effective decision-making. Numerous studies have shown that rewards and punishments are processed and weighted differently in adults than in adolescents (Tversky & Kahneman, 1991; Kahneman & Tversky, 1979). Regardless of clear differences in the processing of reward and punishment, most of the attention in the developmental differences between adults and adolescents is focused on reward processing (Penolazzi, Gremigni, & Russo, 2012; Padmanabhan, Geier, Ordaz, Teslovich, & Luna, 2011; van Leijenhorst, Moor, et al., 2010; for review, see Blakemore & Robbins, 2012; Steinberg, 2005). Only recently has the developmental differences in the processing of punishment between adolescents and adults been studied (Galvan & McGlennen, 2013; Aite et al., 2012; Barkley-Levenson, van Leijenhorst, & Galvan, 2012; van der Schaaf et al., 2011). In a recent study, Galvan and McGlennen (2013) showed that adolescents are hypersensitive to punishments when compared with adults. In line with their findings, our results showed that adolescents possessed significantly higher absolute prediction error in response to punishments in correct trials.

Behavioral data showed that adolescents switched more often than adults in several conditions, even after receiving rewarding feedback. This fact is perfectly in line with this idea. Here, we argue that rewards possibly do not affect the change of expectation strongly enough to pass the uncertainty area, as seen by shallower slope, and thus, this leaves adolescents at a higher probability of switching because of a higher state of uncertainty.

In conclusion, from a developmental perspective, we showed that behavioral differences between groups are reflected in the slope, change of expected value, and prediction error parameters. We showed that (1) adults updated their expected value to a greater extent toward higher certainty and (2) they were adequately sensitive to negative feedback on correct and wrong trials. On the basis of these findings, we argued that adolescents performed the task with lower certainty, reflected by the shallower slope in their decision curves. Furthermore, we speculated about the possibility that adults acquired more accurate knowledge about their current status. Additionally, our approach shows that computational modeling can be effectively used to better understand the mechanisms of decision-making in developmental studies.

### Acknowledgments

We would like to thank Fraser Merchant and Ying Lee for proofreading the document. We would also like to thank the two anonymous reviewers for their constructive comments as

well as Thomas Hübner, Michael Marxen, Eva Mennigen, Kathrin U. Müller, Stephan Ripke, and Sarah Rodehacke for their help in the different stages of the project. This research was supported by the Deutsche Forschungsgemeinschaft (grants SM 80/7-1 and SFB 940) and the German Ministry of Education and Research (BMBF grant 01EV0711). A. H. J. was supported by Wellcome Trust.

Reprint requests should be sent to Amir Homayoun Javadi, Institute of Behavioral Neuroscience, University College London, 26 Bedford Way, WC1H 0AP, London, United Kingdom, or via e-mail: a.h.javadi@gmail.com or Michael N. Smolka, Section of Systems Neuroscience, Technische Universität Dresden, Würzburger Str. 35, 01187, Dresden, Germany, or via e-mail: michael.smolka@tu-dresden.de.

## REFERENCES

- Aïte, A., Cassotti, M., Rossi, S., Poirel, N., Lubin, A., Houdé, O., et al. (2012). Is human decision-making under ambiguity guided by loss frequency regardless of the costs? A developmental study using the Soochow Gambling Task. *Journal of Experimental Child Psychology*, *113*, 286–294.
- Barkley-Levenson, E. E., van Leijenhorst, L., & Galvan, A. (2012). Behavioral and neural correlates of loss aversion and risk avoidance in adolescents and adults. *Developmental Cognitive Neuroscience*, *3*, 72–83.
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214–1221.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, Methodological*, *57*, 289–300.
- Bjork, J. M., Knutson, B., Fong, G. W., Caggiano, D. M., Bennett, S. M., & Hommer, D. W. (2004). Incentive-elicited brain activation in adolescents: Similarities and differences from young adults. *The Journal of Neuroscience*, *24*, 1793–1802.
- Bjork, J. M., Smith, A. R., Chen, G., & Hommer, D. W. (2010). Adolescents, adults and rewards: Comparing motivational neurocircuitry recruitment using fMRI. *PLoS One*, *5*, e11440.
- Blakemore, S.-J., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience*, *15*, 1184–1191.
- Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review*, *28*, 62–77.
- Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*, 111–126.
- Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, *14*, F1–F10.
- Cocosco, C. A., Kollokian, V., Remi, K. S. K., Pike, G. B., & Evans, A. C. (1997). Brainweb: Online interface to a 3D MRI simulated brain database. *Neuroimage*, *5*, S425.
- Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., et al. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience*, *13*, 669–671.
- Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences*, *1021*, 1–22.
- Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., et al. (2005). Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage*, *25*, 1279–1291.
- Ernst, M., Pine, D. S., & Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. *Psychological Medicine*, *36*, 299–312.
- Galvan, A. (2010). Adolescent development of the reward system. *Frontiers in Human Neuroscience*, *4*, 6.
- Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., et al. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, *26*, 6885–6892.
- Galvan, A., Hare, T. A., Voss, H., Glover, G., & Casey, B. J. (2007). Risk taking and the adolescent brain: Who is at risk? *Developmental Science*, *10*, F8–F14.
- Galvan, A., & McClennen, K. M. (2013). Enhanced striatal sensitivity to aversive reinforcement in adolescents versus adults. *Journal of Cognitive Neuroscience*, *25*, 284–296.
- Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics*, *7*, 73–82.
- Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex*, *19*, 483–495.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 8174–8179.
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *41*, 645–655.
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *The Journal of Neuroscience*, *26*, 8360–8367.
- Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences, U.S.A.*, *104*, 1377–1382.
- Hornak, J., O'Doherty, J. P., Bramham, J., Rolls, E., Morris, R., Bullock, P., et al. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, *16*, 463–478.
- Jarcho, J. M., Benson, B. E., Plate, R. C., Guyer, A. E., Detloff, A. M., Pine, D. S., et al. (2012). Developmental effects of decision-making on sensitivity to reward: An fMRI study. *Developmental Cognitive Neuroscience*, *2*, 437–447.
- Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *The Journal of Neuroscience*, *31*, 1606–1613.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
- Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science*, *318*, 1642–1645.
- Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 17951–17956.
- Luce, R. D. (1959). Individual choice behavior: A theoretical analysis. *New York*, *115*, 191–243.
- Montague, P. R. (2006). *Why choose this book? How we make decisions*. New York: EP Dutton.

