The Journal of Neuroscience, September 14, 2011 • 31(37):13039–13045 • 13039

Behavioral/Systems/Cognitive

Behavioral and Neural Properties of Social Reinforcement Learning

Rebecca M. Jones,[1] Leah H. Somerville,[1] Jian Li,[2] Erika J. Ruberry,[1] Victoria Libby,[1] Gary Glover,[3] Henning U. Voss,[4] Douglas J. Ballon,[4] and B. J. Casey[1]

[1]Sackler Institute for Developmental Psychobiology, Weill Cornell Medical College, New York, New York 10065, [2]Department of Psychology, New York University, New York, New York 10003, [3]Lucas Center for Imaging, Department of Radiology, Stanford University, Stanford, California 94305, and [4]Citigroup Biomedical Imaging Center, Department of Radiology, Weill Cornell Medical College, New York, New York, 10065

Social learning is critical for engaging in complex interactions with other individuals. Learning from positive social exchanges, such as acceptance from peers, may be similar to basic reinforcement learning. We formally test this hypothesis by developing a novel paradigm that is based on work in nonhuman primates and human imaging studies of reinforcement learning. The probability of receiving positive social reinforcement from three distinct peers was parametrically manipulated while brain activity was recorded in healthy adults using event-related functional magnetic resonance imaging. Over the course of the experiment, participants responded more quickly to faces of peers who provided more frequent positive social reinforcement, and rated them as more likeable. Modeling trial-by-trial learning showed ventral striatum and orbital frontal cortex activity correlated positively with forming expectations about receiving social reinforcement. Rostral anterior cingulate cortex activity tracked positively with modulations of expected value of the cues (peers). Together, the findings across three levels of analysis—social preferences, response latencies, and modeling neural responses—are consistent with reinforcement learning theory and nonhuman primate electrophysiological studies of reward. This work highlights the fundamental influence of acceptance by one's peers in altering subsequent behavior.

Introduction

Successfully navigating our social environment depends on learning from positive and negative encounters with others and shaping future behavior toward those individuals. Psychologists have proposed that positive social exchanges are fundamentally rewarding for humans (Bandura and Walters, 1963; Baumeister and Leary, 1995; Steinberg, 2008), suggesting that learning from social interactions may draw on basic reinforcement learning mechanisms. The present study was designed to test this hypothesis by building on reinforcement learning studies in nonhuman primates and human imaging studies (Schultz et al., 1997; Fiorillo et al., 2003; McClure et al., 2003; D'Ardenne et al., 2008).

Reinforcement learning from primary (e.g., food) and secondary (e.g., money) reinforcers has been shown to engage specific neural circuitry. In its simplest form, it is explained by the classic Rescorla–Wagner model (Rescorla and Wagner, 1972). According to this model, learning to associate arbitrary cues with positive outcomes results in expectations of future positive outcomes in the presence of these cues. If there are discrepancies between the expected outcome to the cue and the actual outcome, a prediction error signal is generated. Nonhuman primate and human imaging studies have implicated the ventral striatum and orbital frontal cortex (OFC) in prediction error signaling (Schultz et al., 1997; Berns et al., 2001; Fiorillo et al., 2003; McClure et al., 2003). Studies have shown that as cues become reliably associated with receipt of a reward, manual responses to these cues quicken over time (O'Doherty et al., 2006; Spicer et al., 2007), while others demonstrate changes in choice behaviors based on reinforcement manipulations (Daw et al., 2006; Li and Daw, 2011). The learned association generates a neural signal to the cue that previously was associated with the reward itself (Schultz et al., 1997; O'Doherty et al., 2006). The current study examines whether similar changes in behavior (response latencies) and neural circuitry engaged during basic reinforcement learning are involved during learning within a social context.

This study tests the extent to which social reinforcement learning relies on similar learning mechanisms as those used in basic reinforcement learning. To do so, we created a task in which participants learned to differentiate three peers, each of whom was associated with a unique probability of social reinforcement (i.e., providing socially accepting feedback). Social reinforcement learning processes were evaluated at three levels of analysis—preference ratings, response latencies, and neural responses to expected cue values and prediction errors. We hypothesized that social preference ratings would become more favorable and response latencies would become faster toward the peer with the

Received June 13, 2011; revised July 12, 2011; accepted July 19, 2011.

Author contributions: R.M.J., L.H.S., and B.J.C. designed research; R.M.J., E.J.R., and V.L. performed research; J.L., G.G., H.U.V., D.J.B., and B.J.C. contributed unpublished reagents/analytic tools; R.M.J., L.H.S., and J.L. analyzed data; R.M.J., L.H.S., and B.J.C. wrote the paper.

This work was supported by NIDA R01 DA018879, NIDA T-32 Training Grant DA007274, the Mortimer D. Sackler family, and the Dewitt-Wallace fund. We gratefully acknowledge the assistance of the resources and staff at the Biomedical Imaging Core Facility of the Citigroup Biomedical Imaging Center at Weill Cornell Medical College.

The authors report no conflicts of interest.

Correspondence should be addressed to either Rebecca M. Jones or B. J. Casey, Sackler Institute for Developmental Psychobiology, Weill Cornell Medical College, 1300 York Avenue, Box 140, New York, NY 10065. E-mail: rej2004@med.cornell.edu or bjc2002@med.cornell.edu.

DOI:10.1523/JNEUROSCI.2972-11.2011

greatest probability of providing social acceptance to the participant. We applied a simple Rescorla–Wagner rule in behavioral and functional imaging analyses to target the neural bases of these behavioral changes, hypothesizing that the ventral striatum and OFC would code prediction error signals (Schultz et al., 1997; O'Doherty, 2007). Thus, the current study elucidates neurobiological mechanisms for key learning processes during social exchanges that shape behavior through positive interactions.

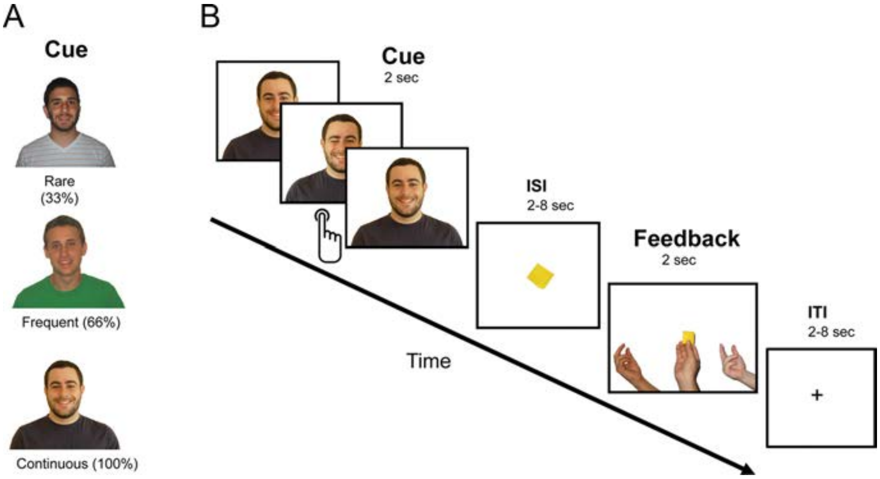

Figure 1. Task parameters. ***A***, Three peers chosen by the participant were associated with distinct probabilities of positive reinforcement. ***B***, Schematic of one trial within a run. The face of one peer (Cue) was displayed for 2 s, during which the face stimulus winked (500 ms) and participants pressed one of two buttons indicating in which eye the wink occurred, followed by a variable interstimulus interval (ISI), followed by the note outcome (Feedback). In this example, the participant received the note (positive social reinforcement) because it appeared in the middle hand. If the note appeared in one of the hands to the left or to the right of the middle hand, the participant did not receive the note (no positive social reinforcement). A variable intertrial interval (ITI) followed.

Materials and Methods

Participants. Forty-six adults (aged 18–28 years; 22 females) participated in the experiment. Thirty-six completed the task during functional magnetic resonance imaging (fMRI) (aged 18–28 years; all right-handed; 19 females). Three individuals in the fMRI group were excluded due to insufficient number of correct trials in any condition ($n = 2$; 1 male) or noncompliance with the task ($n = 1$, male). Participants reported no history of neurological and/or psychiatric disorders in a standard screening or on the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) (First et al., 2007) and imaging participants reported no contraindications for an MRI. Two participants did not complete the SCID due to time constraints. All participants provided written consent approved by the Institutional Review Board at Weill Cornell Medical College and were debriefed and compensated following their participation.

Experiment cover story. The experiment was conducted during two separate sessions. The first session introduced the cover story, leading participants to believe they would receive actual social feedback during a task that would be completed on the second visit. Participants were shown up to five photographs of gender- and ethnicity-matched peers. They then selected three with whom they would like to interact, and rated the three peers for how likeable and attractive they looked on a scale from 1 (not very) to 10 (very). Participants also completed a personal survey where they listed information about themselves (birthday; hometown; and favorite music, TV shows, books, quotes, and activities). Participants were told that each of the three selected peers would see their survey over the next few days as well as the surveys of two other supposed participants. These three peers would write notes indicating a positive interest in the participant's survey or in one of the other two surveys. Participants were told that each of these individuals could write a small number of notes, emphasizing their limited number and enhancing the positive value of receiving a note. Participants were then scheduled for a second session.

At the second session, participants were told that the experimenters had compiled the notes from the three selected peers. During the experiment, participants would be shown how often each of the peers decided to write notes to them (positive social reinforcement) or to one of the other supposed participants (no positive social reinforcement). Although it is possible that participants experienced the no positive social reinforcement trials as mildly rejecting, we have chosen not to adopt this interpretation because we do not have conclusive data supporting this possibility. Rather, these operational definitions were selected for consistency with studies of basic reward learning. At the beginning of the second session, participants were also reminded that receiving a note symbolized that the peer was interested in something written in their personal survey.

Unbeknownst to the participants, peer interaction (i.e., delivery of notes) was experimentally manipulated such that each of the three peers was associated with a distinct probability of social reinforcement (Fig. 1*A*) with Rare interaction defined by positive social reinforcement on 33% of the trials and no positive social reinforcement on 66% of the trials; Frequent interaction defined by positive social reinforcement on 66% of the trials and no positive social reinforcement on 33% of the trials; and Continuous interaction defined by positive social reinforcement on all trials (100%). The probability of reinforcement associated with each of the face stimuli was counterbalanced across participants to equate for low-level stimulus features across conditions.

Task parameters. At the start of each trial (Fig. 1*B*), a picture of one of the three peers was presented for two seconds (Cue). During the two seconds, the stimulus would wink for 500 ms in either the left or right eye, indicating that a note was ready to be passed. Participants signaled that they were ready to receive the note by pressing one of two buttons indicating whether the wink was in the left or the right eye. This behavioral component was included to ensure attention and to collect reaction time data as an index of learning about the reinforcement contingencies for each of the three peers across the experiment. After a jittered interstimulus interval of a picture of a folded note (2, 4, 6, or 8 s), three hands appeared at the bottom of the screen with one hand holding a note for 2 s (Feedback). Participants had been instructed that if the middle hand held the note, this signified that the participant had received a note from that peer (positive social reinforcement). If the note appeared in one of the hands to the left or right of the middle hand, this signified that the note was given to someone else (no positive social reinforcement). If the participant pressed incorrectly or did not respond during the cue, no feedback was given. A jittered intertrial interval (2, 4, 6, or 8 s) followed in which participants rested while viewing a fixation crosshair. Participants viewed 18 trials per run in a pseudorandomized order with six trials per condition (Rare, Frequent, Continuous) for six runs, for a total of 108 trials, 36 trials per condition. To enhance the believability of the cover story and keep participants engaged, one of the notes was shown between each run; these notes were generated by the experimenters and always indicated positive interest in the participant's personal survey (e.g., "I love playing soccer too, and I am part of a weekend league", "Where did you go when you visited Hawaii?", "I also have a golden retriever").

To further index learning with the reaction time data at the end of the experiment, after the six experimental runs, participants completed a reversal run (18 trials) during which reaction times were recorded. Contingencies were reversed for the Rare and Continuous conditions such that the Rare peer now provided 100% reinforcement to the participant and the Continuous peer now provided 33% reinforcement to the participant. The Frequent peer's probability (66%) did not change.

The task was presented using E-Prime software, and the participants who completed the task during fMRI viewed images on an overhead liquid crystal display panel with the Integrated Functional Imaging System-Stand Alone (IFIS-SA; fMRI Devices). E-Prime software, integrated with IFIS-SA, recorded button responses and reaction times using the Fiber Optic Button Response System (Psychology Software Tools).

At the end of the experiment, participants completed posttest ratings of attractiveness and likeability for each peer on the same scale used at the beginning of the experiment. To assess whether participants held explicit knowledge of the social reinforcement contingencies associated with each peer, they were asked whether any of the three peers provided positive reinforcement more often than any others. If the participant said yes, they were asked to describe what pattern they noticed, and descriptions were scored based on whether the participant accurately stated which peer provided the most, intermediate, and least positive social feedback. Three of the 43 participants correctly ranked the three peers in this way and were thus considered explicitly aware of the social reinforcement contingencies. Participants were then debriefed regarding the cover story and the rationale of the experiment.

Image acquisition. Participants were scanned with a Signa HDx 3.0T MRI scanner (General Electric Medical Systems) with a quadrature head coil. A high-resolution, 3D magnetization prepared rapid acquisition gradient echo anatomical scan (MPRAGE) was acquired (256 × 256 in-plane resolution, FOV = 240 mm; 124 1.5 mm sagittal slices). Functional scans were acquired with a spiral in and out sequence (TR = 2000 ms, TE = 30 ms, flip angle = 90°) (Glover and Thomason, 2004). Twenty-nine 5-mm-thick contiguous coronal slices were acquired per TR, for a total of 129 TRs per functional run with a resolution of 3.125 × 3.125 mm (64 × 64 matrix, FOV = 200 mm) covering the entire brain except for the posterior portion of the occipital lobe.

Behavioral analysis. Change in attractiveness and likeability of the peers before and after the task was tested with a 3 (probability: Rare, Frequent, Continuous) × 2 (time: before task, after task) repeated-measures ANOVA using PASW Statistics 18 software (SPSS). Attractiveness and likeability ratings for three of the 43 participants were lost due to technical error.

Reaction times were analyzed in response to the cue after the wink occurred. Reaction times were *z*-score transformed for each individual after removing outliers (defined as reaction times 3 SDs above or below probability and time (pre-interaction, post-interaction) on likeability ratings ($F_{(2,78)} = 5.48, p < 0.01$; Fig. 2*A*). *Post hoc* analyses indicated that post-task ratings decreased linearly with decreasing interaction probability, such that peers who interacted less with the participant were rated as less likeable (linear term: $F_{(1,39)} = 7.17, p < 0.02$). Whereas pre-task likeability ratings were equivalent for all three peers (ps > 0.48), after the task the Frequent ($t_{(39)} = -2.26, p < 0.03$) and Continuous ($t_{(39)} = -2.68$, $p < 0.02$) peers were rated as more likeable than the rarely reinforcing peer, though there was not a significant difference in likeability ratings after the task between the Frequent and Continuous peers ($t_{(39)} = -0.69$, $p > 0.49$). Attractiveness ratings were not significantly modulated by task conditions (main effects of reinforcement probability, time, and interactions: ps > 0.09).

Accuracy

Participants responded correctly to 95.63% of trials (SD = 3.54%). Response accuracy was not significantly modulated by the task conditions (main effects of reinforcement probability, time, and interactions: ps > 0.29).

Reaction time

Response latencies to the cue varied as participants learned the reinforcement contingency outcomes associated with each peer, F(1,39) as indicated by a significant inpj /T1 2 1 Tf 22t contis6(Fig.)-s6(Fig.0 Td 0.216 [(reinfor74uent)-274tion)--274ters (

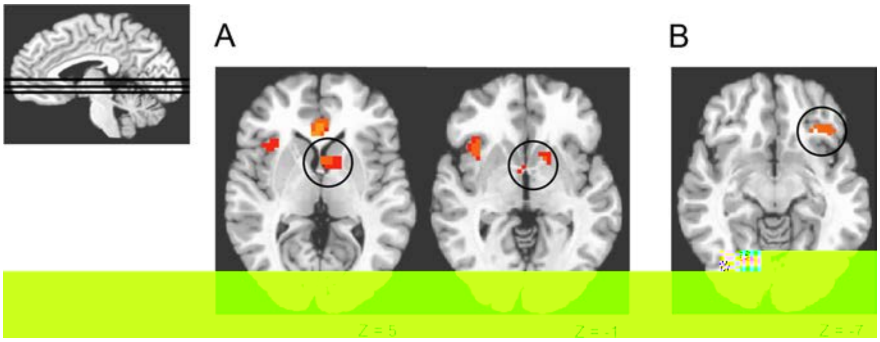

Figure 3. Brain regions reflecting positive correlations with prediction errors. ***A***, Circles denote activity in the ventral striatum. Image threshold $p < 0.05$, whole-brain corrected. ***B***, Circle denotes activity in the lateral orbital frontal cortex. Image threshold $p < 0.05$, small volume corrected (see Materials and Methods). All statistical activations are displayed on a representative high-resolution axial image. The left side of the image corresponds to the right side of the brain.

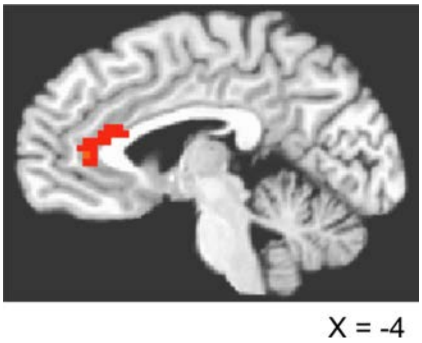

Figure 4. Neural activity with positive correlations with learned cue value. Activity in the rostral anterior cingulate cortex reflects a positive correlation with expected values for the cues. Image threshold $p < 0.05$, whole-brain corrected. Statistical activations are displayed on a representative high-resolution sagittal image.

ment outcomes. Specifically, one peer always provided positive social reinforcement, another one frequently provided positive social reinforcement, and the third rarely provided positive social reinforcement. Ratings of likeability changed from the beginning to the end of the experiment, with less reinforcing peers becoming less likeable, and more reinforcing peers yielding higher ratings of likeability by the end of the task. By asking participants to make a simple button response during the cue presentation, we tested whether speeding of response latencies (action tendencies) indexed learned associations between a given peer and their probability of providing positive social reinforcement. As expected due to the simplicity of the task, accuracy was at ceiling and there were no statistical differences in accuracy for the three peers.

In the current study, we observed faster responses to peers who provided positive social reinforcement more often, similar to studies where participants responded more quickly to cues that reliably predicted receiving a primary or secondary reward (O'Doherty et al., 2006; Spicer et al., 2007). Measuring differences in reaction times to cues to index learning differs from reinforcement studies that use modulated choice behavior as an indicator of learning (Tanaka et al., 2004; Daw and Doya, 2006; Schönberg et al., 2007). Choice tasks index changes in explicit preferences or a participant's strategy in maximizing reinforcement, while in the current study, changes in responses are thought to index differences in approach behaviors that are based on learning from a prior history of social feedback. In addition, participants showed faster reaction times after trials that did not provide positive social reinforcement. This finding is similar to studies that demonstrate improved performance on a trial that follows receiving punishment (Hester et al., 2010) or choosing to make a bet more often after losing money than after winning money (Liu et al., 2007), though the present study did not assess strategic behavior directly. Together, the behavioral findings demonstrate that participants learned the reinforcement contingencies and thus provide an objective index of social learning.

The changes in likeability ratings and response latencies did not appear to be conscious behavioral choices. The majority (93%) of participants were unable to articulate the reinforcement patterns, suggesting little if any explicit awareness of the reinforcement contingencies. These findings demonstrate that social preferences and actions can be influenced after only brief encounters with peers and without conscious awareness. Such rapid changes highlight the influence of positive social interactions on effectively altering subsequent behavior.

The neural correlates of these behavioral changes draw upon the same neural circuitry as that implicated in reinforcement learning (Alexander et al., 1986; Haber and Knutson, 2010). Prediction error (δ_t) learning engaged the ventral striatum and orbital frontal cortex, similar to previous studies using single-cell recordings (Schultz et al., 1997; Fiorillo et al., 2003; Sul et al., 2010) and human imaging studies with primary reinforcers such as juice (McClure et al., 2003; O'Doherty et al., 2003; D'Ardenne et al., 2008) and secondary reinforcers such as money or attractive or smiling faces (Bray and O'Doherty, 2007; Valentin and O'Doherty, 2009; Lin et al., 2011). Together, these findings support a role for the orbital frontostriatal circuit in generating learning signals from positive social reinforcement and provide a neural basis for how feedback during a social interaction is flexibly updated to inform subsequent social expectations.

The present study is distinct in its capacity to test whether registering violations in expectations of social acceptance draws on basic mechanisms that support prediction error learning. Using a simple Rescorla–Wagner learning model, we show that violations in expected social interaction are tightly coupled with changes in ventral striatal activity. No prior studies, to our knowledge, have applied a classic reinforcement learning model in the examination of learning from social reinforcers. In the social domain, studies have modeled trial-by-trial decisions about charitable donations (Hare et al., 2010) or intentions to trust a partner during economic exchanges (King-Casas et al., 2005). Furthermore, the current paradigm is distinct from previous studies that compare social acceptance to rejection (Eisenberger et al., 2003; Somerville et al., 2006; Guyer et al., 2009), as it targets the process of learning from the social feedback, rather than comparing acceptance to rejection. Therefore, the present study offers a unique explanation for how we learn from positive social interactions.

Our finding that the striatum is sensitive to expectations about receiving social feedback converges with other work targeting the neural mechanisms of social learning. Recently, Harris and Fiske (2010) showed sensitivity in this region to violations in expectations about personality trait information, and others have shown that the striatum is sensitive to violations of social group norms (Klucharev et al., 2009) as well as forming predictions about investors' decisions (King-Casas et al., 2005; Phan et al., 2010). Our results complement these studies by demonstrating a neural mechanism for how prior positive interactions with others shape our expectations for future interactions. Given the increased sensitivity in the ventral striatum to appetitive stimuli during adolescence (Galvan et al., 2006; Somerville et al., 2011), as well as the greater influence of peers during adolescence (Spear, 2000; Gardner and Steinberg, 2005), this work clearly raises the question of how peer interaction differentially impacts learning and behavior across development and how this may be differentially represented in the brain. Accordingly, it would be interesting to explore whether adolescents show increased sensitivity during social learning relative to children and adults.

The expected values (V_t) to the cues corresponded with greater activity in the rostral anterior cingulate cortex. Previous studies have shown the rostral anterior cingulate cortex/medial prefrontal cortex is sensitive to cues that predict reward receipt (Tanaka et al., 2004; Knutson et al., 2005; Palminteri et al., 2009) and may play a role in general learning about the value of information and using this information for future decisions (Rushworth and Behrens, 2008). Lesion studies in nonhuman primates have shown this region is important for establishing patterns of social interest in other individual male or female macaques

Li J, Daw ND (2011) Signals in human striatum are appropriate for policy update rather than value prediction. J Neurosci 31:5504–5511.

Li J, Delgado MR, Phelps EA (2011) How instructed knowledge modulates the neural systems of reward learning. Proc Natl Acad Sci U S A 108:55–60.

Lin A, Adolphs R, Rangel A (2011) Social and monetary reward learning engage overlapping neural substrates. Soc Cogn Affect Neurosci. Advance online publication. Retrieved March 22, 2011. doi:10.1093/scan/nsr006.

Liu X, Powell DK, Wang H, Gold BT, Corbly CR, Joseph JE (2007) Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. J Neurosci 27:4587–4597.

McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. Neuron 38:339–346.

O'Doherty JP (2007) Lights, camembert, action! The role of human orbitofrontal cortex in encoding stimuli, rewards, and choices. Ann N Y Acad Sci 1121:254–272.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. Neuron 38:329–337.

O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ (2006) Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. Neuron 49:157–166.

Palminteri S, Boraud T, Lafargue G, Dubois B, Pessiglione M (2009) Brain hemispheres selectively track the expected value of contralateral options. J Neurosci 29:13465–13472.

Phan KL, Sripada CS, Angstadt M, McCabe K (2010) Reputation for reciprocity engages the brain reward center. Proc Natl Acad Sci U S A 107:13099–13104.

Redcay E, Dodell-Feder D, Pearrow MJ, Mavros PL, Kleiner M, Gabrieli JD, Saxe R (2010) Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. Neuroimage 50:1639–1647.

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Classical conditioning II: current research and theory (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton Century Crofts.

Rudebeck PH, Buckley MJ, Walton ME, Rushworth MF (2006) A role for the macaque anterior cingulate gyrus in social valuation. Science 313:1310–1312.

Rushworth MF, Behrens TE (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci 11:389–397.

Schönberg T, Daw ND, Joel D, O'Doherty JP (2007) Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. J Neurosci 27:12860–12867.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599.

Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. Nature 429:664–667.

Somerville LH, Heatherton TF, Kelley WM (2006) Anterior cingulate cortex responds differentially to expectancy violation and social rejection. Nat Neurosci 9:1007–1008.

Somerville LH, Hare T, Casey BJ (2011) Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. J Cogn Neurosci 23:2123–2134.

Spear LP (2000) The adolescent brain and age-related behavioral manifestations. Neurosci Biobehav Rev 24:417–463.

Spicer J, Galvan A, Hare TA, Voss H, Glover G, Casey B (2007) Sensitivity of the nucleus accumbens to violations in expectation of reward. Neuroimage 34:455–461.

Steinberg L (2008) A social neuroscience perspective on adolescent risk-taking. Dev Rev 28:78–106.

Sul JH, Kim H, Huh N, Lee D, Jung MW (2010) Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron 66:449–460.

Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G (2009) The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. Neuron 62:269–280.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.

Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat Neurosci 7:887–893.

Valentin VV, O'Doherty JP (2009) Overlapping prediction errors in dorsal striatum during instrumental learning with juice and money reward in the human brain. J Neurophysiol 102:3384–3391.