Article

Neuron

Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes

Highlights

- The VLPFC is necessary for updating stimulus-outcome probability estimates
- The OFC is necessary for updating stimulus-outcome values based on current needs
- The VLPFC is necessary for both contingent and noncontingent learning

Authors

Peter H. Rudebeck,
Richard C. Saunders,
Dawn A. Lundgren, Elisabeth A. Murray

Correspondence

peter.rudebeck@mssm.edu (P.H.R.),
murraye@mail.nih.gov (E.A.M.)

In Brief

Rudebeck et al. show that the orbitofrontal cortex (OFC) and ventrolateral prefrontal cortex (VLPFC) make dissociable contributions to decision-making. The VLPFC, but not the OFC, is critical for updating stimulus-outcome availability, whereas the OFC, but not the VLPFC, is vital for updating stimulus-outcome desirability.

Rudebeck et al., 2017, Neuron *95*, 1208–1220
August 30, 2017
http://dx.doi.org/10.1016/j.neuron.2017.07.042

CellPress

CellPress

Neuron
Article

Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes

Peter H. Rudebeck,[1,3,*] Richard C. Saunders,[2] Dawn A. Lundgren,[2] and Elisabeth A. Murray[2,*]
[1]Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY 10014, USA
[2]Section on the Neurobiology of Learning and Memory, Laboratory of Neuropsychology, National Institute of Mental Health, Building 49, Suite 1B80, 49 Convent Drive, Bethesda, MD 20892, USA
[3]Lead Contact
*Correspondence: peter.rudebeck@mssm.edu (P.H.R.), murraye@mail.nih.gov (E.A.M.)
http://dx.doi.org/10.1016/j.neuron.2017.07.042

SUMMARY

Advantageous foraging choices benefit from an estimation of two aspects of a resource's value: its current desirability and availability. Both orbitofrontal and ventrolateral prefrontal areas contribute to updating these valuations, but their precise roles remain unclear. To explore their specializations, we trained macaque monkeys on two tasks: one required updating representations of a predicted outcome's desirability, as adjusted by selective satiation, and the other required updating representations of an outcome's availability, as indexed by its probability. We evaluated performance on both tasks in three groups of monkeys: unoperated controls and those with selective, fiber-sparing lesions of either the OFC or VLPFC. Representations that depend on the VLPFC but not the OFC play a necessary role in choices based on outcome availability; in contrast, representations that depend on the OFC but not the VLPFC play a necessary role in choices based on outcome desirability.

INTRODUCTION

To choose the most advantageous course of action, humans and other animals need to combine information about the desirability of an option with a graded estimate of its potential availability, and economists have long appreciated these two aspects of valuation. By combining the probability of a particular outcome with its subjective value, the overall value of a particular course of action can be estimated. Although economic behavior of this sort is reasonably well understood at the behavioral level, the brain areas necessary for processing these two aspects of valuation remain uncertain.

The orbitofrontal cortex (OFC; Walker's areas 11, 13, and 14) is widely held to be important for learning about both reward value and reward contingency (Mishkin, 1964; Padoa-Schioppa, 2011; Rolls, 2000; Wallis, 2007). "Reward value" in the present context refers to subjective value based on preference or desirability of a particular food outcome as opposed to value as commonly computed in economic theory (probability × magnitude). Lesions of the granular OFC of primates disrupt the ability to use information about the desirability and probability of rewarding outcomes to guide decision-making (Camille et al., 2011; Hornak et al., 2004; Izquierdo et al., 2004; Walton et al., 2010), and similar observations have followed lesions of the agranular OFC of rodents (Burke et al., 2008; Mobini et al., 2002).

Recently, a role for the OFC in signaling reward probability has been questioned; monkeys with selective excitotoxic lesions of the OFC, unlike monkeys with aspiration lesions of the OFC, are unimpaired in learning and reversing object choices based on reward feedback in deterministic settings (Rudebeck et al., 2013). This finding raises a question about the learning of stimulus-outcome probabilities: is the OFC involved, and, if not the OFC, then which area is necessary for updating these representations in the primate brain? The work of Walton et al. (2010), combined with our previous results (Rudebeck et al., 2013), suggests that some area near the OFC might be the crucial area, rather than the OFC per se. The adjacent inferior convexity has been implicated in similar types of learning (Iversen and Mishkin, 1970; Rygula et al., 2010), but only with a deterministic experimental design similar to that used by Rudebeck et al. (2013). Accordingly, we tested the contributions of both regions to choices based on reward desirability and reward probability.

Here we report the effects of excitotoxic lesions of either the granular OFC or ventrolateral prefrontal cortex (VLPFC; Walker's areas 12, 45, and ventral 46)—a part of the granular prefrontal cortex adjacent to the OFC—on two tasks. One task is designed to assess the ability to use the updated probability of a predicted outcome to guide a choice among visual stimuli and the other to measure the ability to use the current desirability of a predicted outcome to make similar choices. In both tasks, monkeys chose between options depending on their expected value. In the first task (experiment 1), we manipulated the probability of receiving a single reward for a particular choice while holding the desirability and magnitude of reward constant. In the second task (experiment 2), we manipulated the subjective value of different food rewards with a selective satiation procedure while holding the probability and magnitude of reward stable.

Table 1. Quantification of OFC and VLPFC Lesions

Case No.	Left	Right	Mean	Left	Right	Mean
Monkeys that Received OFC Lesions						
	OFC (Intended)			VLPFC (Inadvertent)		
OFC 1	82.0	78.2	80.1	0.64	0.15	0.4
OFC 2	92.2	89.7	91.0	4.26	2.66	3.46
OFC 3	81.2	60.7	71.0	2.49	0.28	1.39
OFC 4	96.1	96.6	96.3	11.1	7.9	9.5
Mean	87.9	81.3	84.6	4.6	2.7	3.7
Monkeys that Received VLPFC Lesions						
	OFC (Inadvertent)			VLPFC (Intended)		
VLPFC 1	0.75	3.56	2.16	91.6	95.9	93.8
VLPFC 2	3.66	3.65	3.66	97.8	100.0	98.9
VLPFC 3	3.95	0	1.98	92.5	83.6	88.1
VLPFC 4	0	2.12	1.06	83.1	86.8	85.0
Mean	2.1	2.3	2.2	91.3	91.6	91.4

Percent of intended and unintended damage to either the OFC or VLPFC in monkeys that received OFC lesions (top) and monkeys that received VLPFC lesions (bottom).

RESULTS

Experiment 1: Updating Likelihood Estimates for Predicted Outcomes

We trained a group of unoperated control monkeys (n = 8) and a group of monkeys with excitotoxic OFC lesions (n = 4) to perform a three-choice probabilistic learning task (Figure 1; Walton et al., 2010). Four of the unoperated control monkeys subsequently completed additional preoperative testing, received excitotoxic lesions of the VLPFC (n = 4; Figure 1C; Figure S1), and then were retested on the three-choice probabilistic task. This difference in testing history between the OFC and VLPFC lesion groups meant that monkeys with OFC lesions were compared with concurrently run controls, whereas monkeys with VLPFC lesions were compared with their own preoperative performance (see Figure S7 for full details of the testing order).

Based on an MRI assessment, we estimated that the lesions destroyed a mean of 84.6% of the OFC (range, 71.0%–96.3%) and, in the other group of operated monkeys, 91.4% of the VLPFC (range, 85.0%–98.9%; Supplemental Information; Figure S1; Table 1). Importantly, there was minimal overlap between lesions with, on average, less than 5% of the nontarget structure affected (Table 1). Inadvertent damage was typically unilateral and inconsistent across subjects.

At the start of each 300-trial session, monkeys were presented with three novel stimuli on a touch screen monitor (Figure 1A). By sampling different stimuli over trials, monkeys could learn which of the three stimuli was the best option; i.e., the one associated with the highest probability of receiving a single banana-flavored pellet. Because the reward probabilities assigned to each option changed over the course of the session, to maximize the reward, the monkeys needed to continually update their representation of the best option. Reward delivery for selecting a particular stimulus was predetermined based on one of four different schedules, as described in Figure 1B, Figure S3, and Walton et al. (2010), and each trial was followed by a 5-s intertrial interval (ITI).

Unoperated monkeys, both the control group and monkeys before VLPFC lesions (a pre-operative group), quickly learned which image was associated with the highest probability of reward and were able to track the best option as it changed over the course of each test session (Figure 2A, gray line/shaded area). Contrary to reports of deficits in updating probabilistic outcomes after aspiration lesions of the OFC (Camille et al., 2011; Hornak et al., 2004; Mobini et al., 2002; Walton et al., 2010), monkeys with selective, excitotoxic lesions of the OFC have no impairment on this task. For instance, on schedule 2, the choices of monkeys with OFC lesions clearly overlapped with those of the unoperated controls (Figure 2A). To probe this null result, we used a reinforcement learning model to estimate, on a trial-by-trial basis, whether monkeys were choosing the image associated with the highest probability of reward based on their history of previous choices and outcomes on schedule 2 (Figure 2B). Estimating the best choice on each trial in this way confirmed that monkeys with OFC lesions did not differ from controls ($F(1,10) < 0.1$, $p > 0.9$). In addition, monkeys with OFC lesions also chose the option associated with the highest probability of reward at greater than chance levels (one-sample t test, $t(3) = 4.9$, $p < 0.01$). This null effect was consistent over all schedules on which the monkeys with OFC lesions were tested (Figure 3A; Figures S2 and S3; effect of group, $F(1,10) = 0.15$, $p > 0.7$, group by schedule interaction, $F(3,30) = 0.28$, $p > 0.8$) and was not dependent on the phase of the test session (first versus second 150 trials, effect of phase or phase by group interaction, F values < 2, $p > 0.15$).

In contrast, after excitotoxic lesions of the VLPFC, monkeys exhibited a profound deficit in the ability to learn probabilistic stimulus-outcome associations. The deficit was most prominent when the image associated with the highest probability of reward switched at the midpoint of the session (Figure 2A). Determining the best choice on each trial on schedule 2 using a reinforcement learning model further revealed that, after VLPFC lesions, monkeys were less likely to choose the option associated with the highest probability of reward in both the first and second 150 trials (Figure 2B; effect of surgery, $F(1,3) = 12.42$, $p < 0.05$; phase by surgery interaction, $F(1,3) = 1.95$, $p > 0.25$). This effect of VLPFC lesions on learning was observed across all of the schedules that the monkeys completed (Figure 3A; Figures S2 and S3; effect of surgery, $F(1,10) = 10.14$, $p = 0.05$; surgery by phase interaction, $F(1,3) = 0.22$, $p > 0.6$), with one exception: schedule 1 (surgery by schedule by phase interaction, $F(1,3) = 7.77$, $p < 0.05$). In the first 150 trials of this schedule, one option has a very high probability of reward compared with the other options (Figure 1B), and, in this situation, VLPFC lesions did not affect the ability of monkeys to learn the option associated with the highest probability of reward (effect of surgery first 150 trials, $F(1,3) = 2.64$, $p > 0.2$). Overall, this analysis indicates that VLPFC lesions affect learning of probabilistic associations, especially when the difference between options is small, and have less influence when there is one good option.

Previous reports have interpreted the effects of OFC lesions in terms of a perseverative impairment related to the loss of inhibitory control (Rolls et al., 1994; cf. Walton et al., 2010), and the

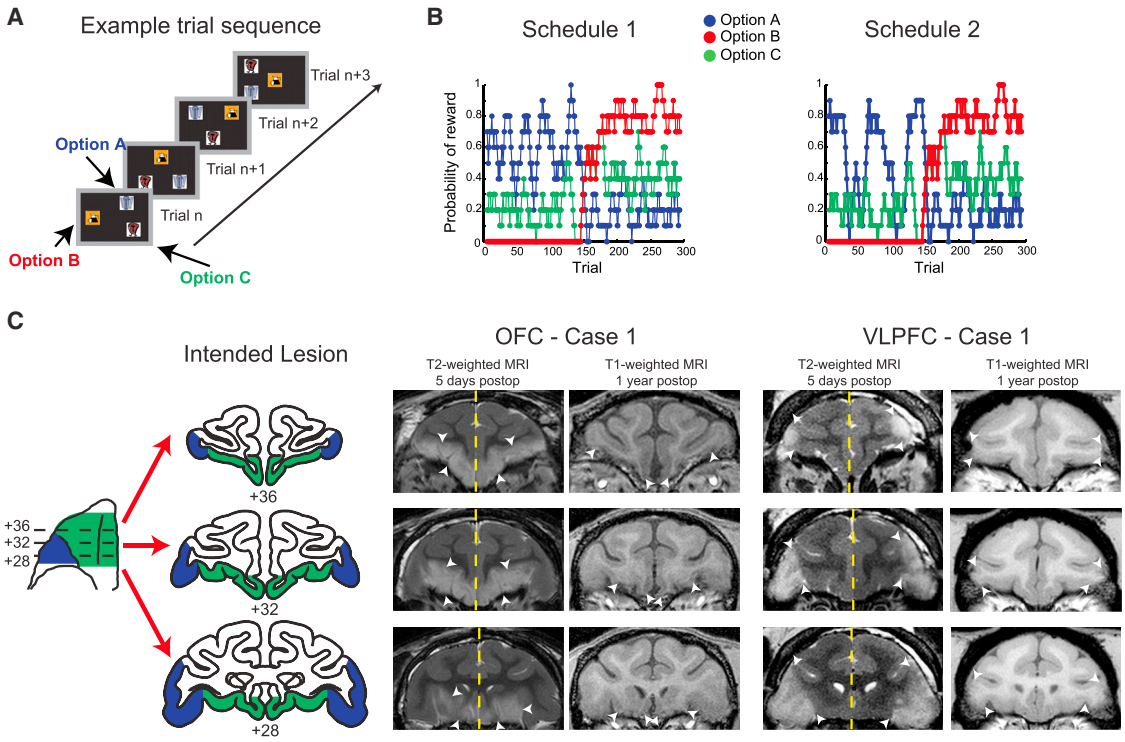


Figure 1. The Three-Choice Probabilistic Learning Task, Reward Schedules, and Lesions Extents

(A) Task sequence. On each trial, monkeys were presented with three stimuli for choice and, through trial and error, could learn which stimulus was associated with the highest probability of reward.

(B) Reward delivery was dependent on the underlying reward schedules shown here and the ones illustrated in Figure S3.

(C) Schematic of OFC (green) and VLPFC lesions (blue). For both OFC and VLPFC lesions, T2-weighted MRI images taken within 1 week of surgery were used to estimate the extent of the lesions. The white hypersignal in the T2-weighted images—set off by arrowheads—is associated with edema that follows injections of excitotoxins and indicates the likely extent of the lesion. For T2-weighted images, the left and right sides of the MR images are from different scans and have been placed together for ease of viewing. Yellow dashed lines indicate where images from two different postoperative scans have been joined. MR images from T1-weighted scans acquired at least a year after surgery confirm the loss of cortex in the intended regions. Numbers indicate the distance in millimeters from the interaural plane. The MRI images are from levels matching the drawings of coronal sections.

cortex in the inferior frontal gyrus in humans has also been associated with inhibitory control (Aron et al., 2004). We therefore examined whether lesions of the VLPFC or OFC resulted in perseveration; i.e., a decrease in the likelihood of switching choices. As can be seen for schedule 2, monkeys with VLPFC lesions were much more likely to change their choice from one trial to the next compared with before lesions were made (effect of surgery, $F(1,3) = 45.53$, $p < 0.01$; Figure 2C, bottom). In contrast, monkeys with OFC lesions did not differ from controls in this regard (group, $F(1,10) = 0.07$, $p > 0.8$; Figure 2C, top).

To further probe this effect, we evaluated the influence of positive (reward) and negative (no reward) feedback on subsequent choices across all schedules. Unoperated controls and monkeys with OFC lesions showed a similar pattern of behavior; both groups were less likely to switch choices after a rewarded choice (positive feedback) than after an unrewarded one (Figure 3B, top row; effect of group, $F(1,10) = 0.04$, $p > 0.8$; effect of reward, $F(1,10) = 225.56$, $p < 0.001$). Following VLPFC lesions, however, the effect of positive feedback on choice was reduced (Figure 3B; effect of surgery, $F(1,3) = 9.05$, $p = 0.057$). Thus, the deficit in monkeys with VLPFC lesions appears to be characterized by an inability to assign feedback to the previously chosen stimulus.

To directly test this hypothesis, we conducted a logistic regression analysis to assess how monkeys used the outcomes they received for choosing a particular option on each trial, either reward or no reward, to guide future choices. This analysis goes beyond those conducted above because it allows us to determine not just the effect of the most recent choice and outcome but also longer-term effects of reward history and choice history on current choices. Our analysis was identical to the one conducted on the choice behavior of monkeys with aspiration lesions of the OFC (Walton et al., 2010) and was conducted on choices and outcomes from all four schedules. To specifically look at how choice and outcome history influenced behavior, this analysis included all of the possible combinations of choices and outcomes—i.e., whether monkeys received a reward or not—from the five preceding trials (n-1 to n-5, Figure 4A). We also included the n-6 trial in the analysis as a confounding variable for longer choice and reward histories (see STAR Methods for full details).

By computing the influence of all combinations of choices and outcomes from the recent past in this way, we were able to probe a monkey's ability to credit an outcome to the choice made directly before. This type of learning is often referred to as

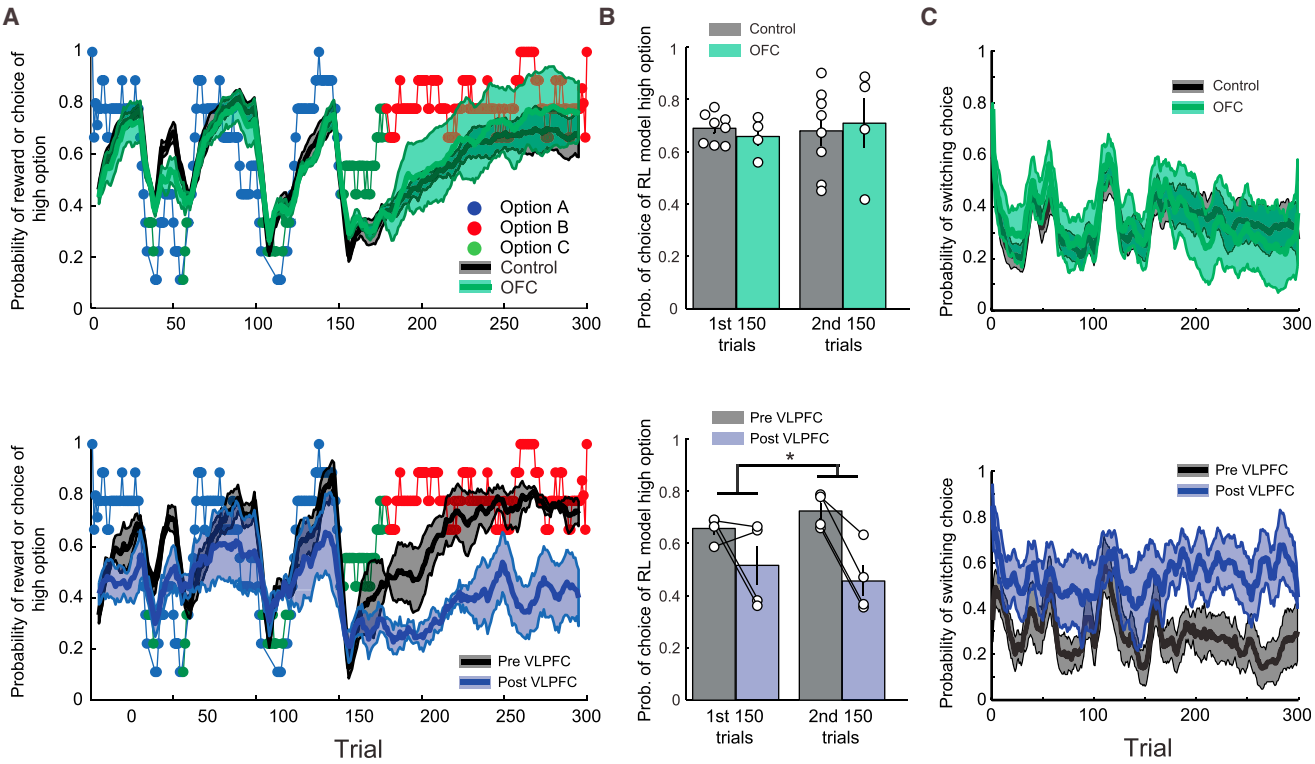


Figure 2. VLPFC, but Not OFC, Lesions Disrupt the Ability to Choose According to Outcome Probability on the Three-Choice Probabilistic Learning Task

(A) Mean (± SEM) choice behavior of unoperated controls (gray, top, n = 8), monkeys with OFC lesions (green, top, n = 4), and monkeys before (gray, n = 4) and after (blue, n = 4) VLPFC lesions (bottom) on schedule 2. Note that in (A) (top), the gray curve and shading (control) are largely obscured by the overlying green curve and shading (OFC). Colored points represent the identity and probability of receiving a reward for selection of the high reward option.

(B) Mean (± SEM) probability of choice of reinforcement learning estimated high reward option in the first and second sets of 150 trials for unoperated controls (n = 8) and monkeys with OFC lesions (top, n = 4) and monkeys before and after VLPFC lesions (bottom, n = 4) on schedule 2. Symbols show scores of individual subjects.

(C) Mean (± SEM) probability of switched choice options from trial to trial for unoperated controls (n = 8) and monkeys with OFC lesions (top, n = 4) and monkeys before and after VLPFC lesions (bottom, n = 4) on schedule 2.

*$p < 0.05$. Error bars represent SEM. Also see Figures S2 and S3.

"contingent learning," in the sense that a causal association is made between a particular choice and its contingent outcome ("law of effect"; Thorndike, 1933). In such learning, positive outcomes that follow a choice will increase the likelihood of that choice being repeated, the converse for negative outcomes. In the example in Figure 4A, higher weightings on the diagonal of the matrix of past choices and outcomes would indicate that monkeys are learning contingently (Figure 4A, red squares in the matrix). In addition, this approach also allowed us to probe noncontingent learning mechanisms ("spread of effect"; Thorndike, 1933): how past outcomes can influence choices made nearby in time but that did not causally lead to that outcome. In Figure 4A, noncontingent learning is associated with higher weighting in off-diagonal parts of the matrix, most notably on the vertical or horizontal from the previous trial, corresponding to the influence of both previous choices and reward, respectively (Figure 4A, blue and green squares in the matrix).

The choices of unoperated monkeys, both control and preoperative monkeys, were strongly influenced by recently chosen stimuli and the outcome, either rewarded or unrewarded, associated with each of those choices, as evidenced by the higher weightings on the diagonal of the matrix of past choices and reward (Figure 4A, red shading; Figure 4B, left; and Figure 4D). Such a pattern indicates that monkeys were making contingent associations between their specific choices and subsequent outcomes. This effect diminished with increasing distance from the current trial, suggesting that monkeys preferentially used the most recent feedback to guide future choices (effect of trial; unoperated controls, $F(5,35) = 12.65$, $p < 0.01$; preoperative VLPFC lesion monkeys, $F(5,15) = 27.74$, $p < 0.001$). In keeping with the findings of Walton et al. (2010), there was also evidence of monkeys learning from noncontingent choices and outcomes, as evidenced by higher weightings in matrix squares away from the diagonal (Figure 4B). Specifically, there was an influence of recent reward on previous choices (Figure 4A, blue shading; Figure 4C, controls and preoperative VLPFC effect of trial F values > 10, $p < 0.01$) as well as an influence of previous reward on recent choices (Figure 4A, green shading; Figure 4E, F values > 14, $p < 0.005$), and both affected subsequent choices.

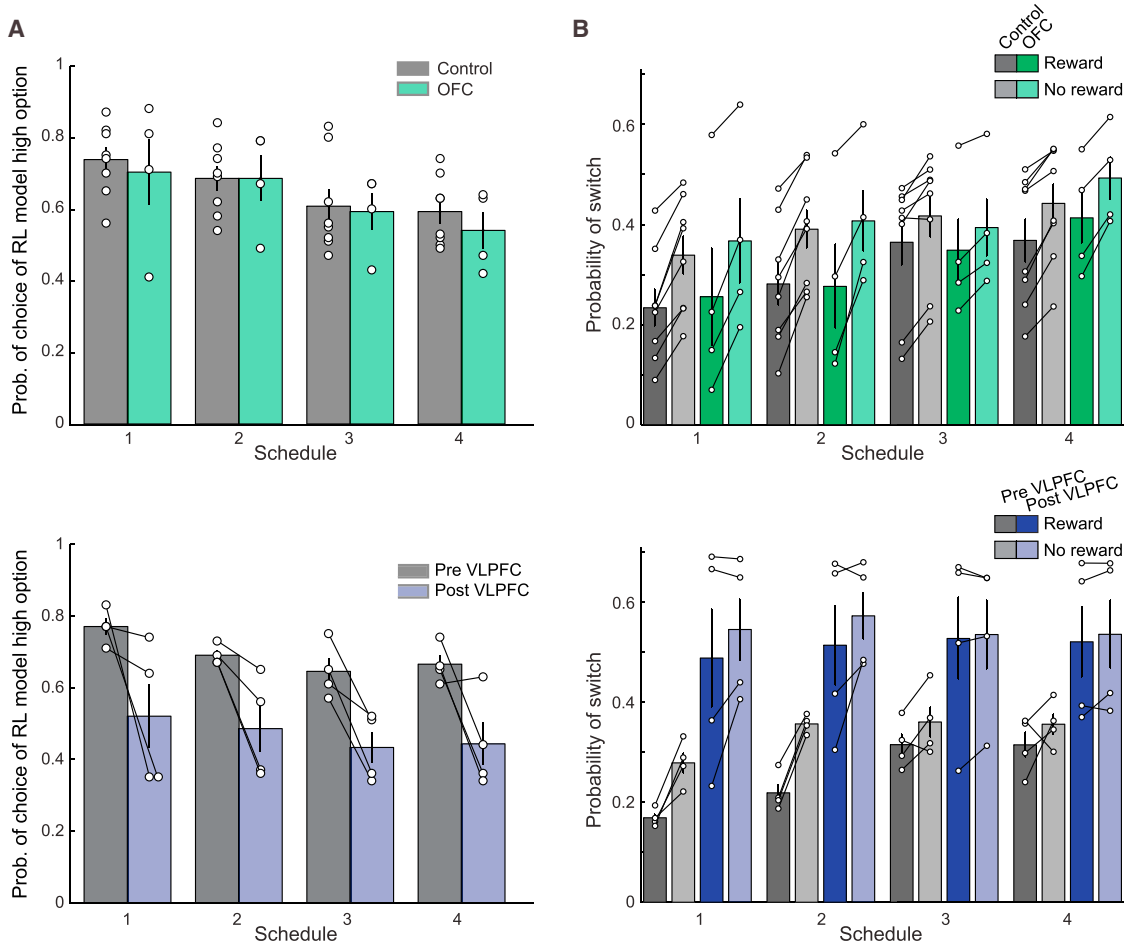


Figure 3. VLPFC, but Not OFC, Lesions Disrupt Probabilistic Learning

(A) Mean (± SEM) probability of choice of the option associated with the highest probability of reward, as defined by a reinforcement learning model fit to monkeys' choices in each of the 4 schedules for unoperated controls (gray, top, n = 8) and monkeys with OFC lesions (green, top, n = 4) and monkeys before and after VLPFC lesions (gray and blue, respectively; bottom; n = 4).

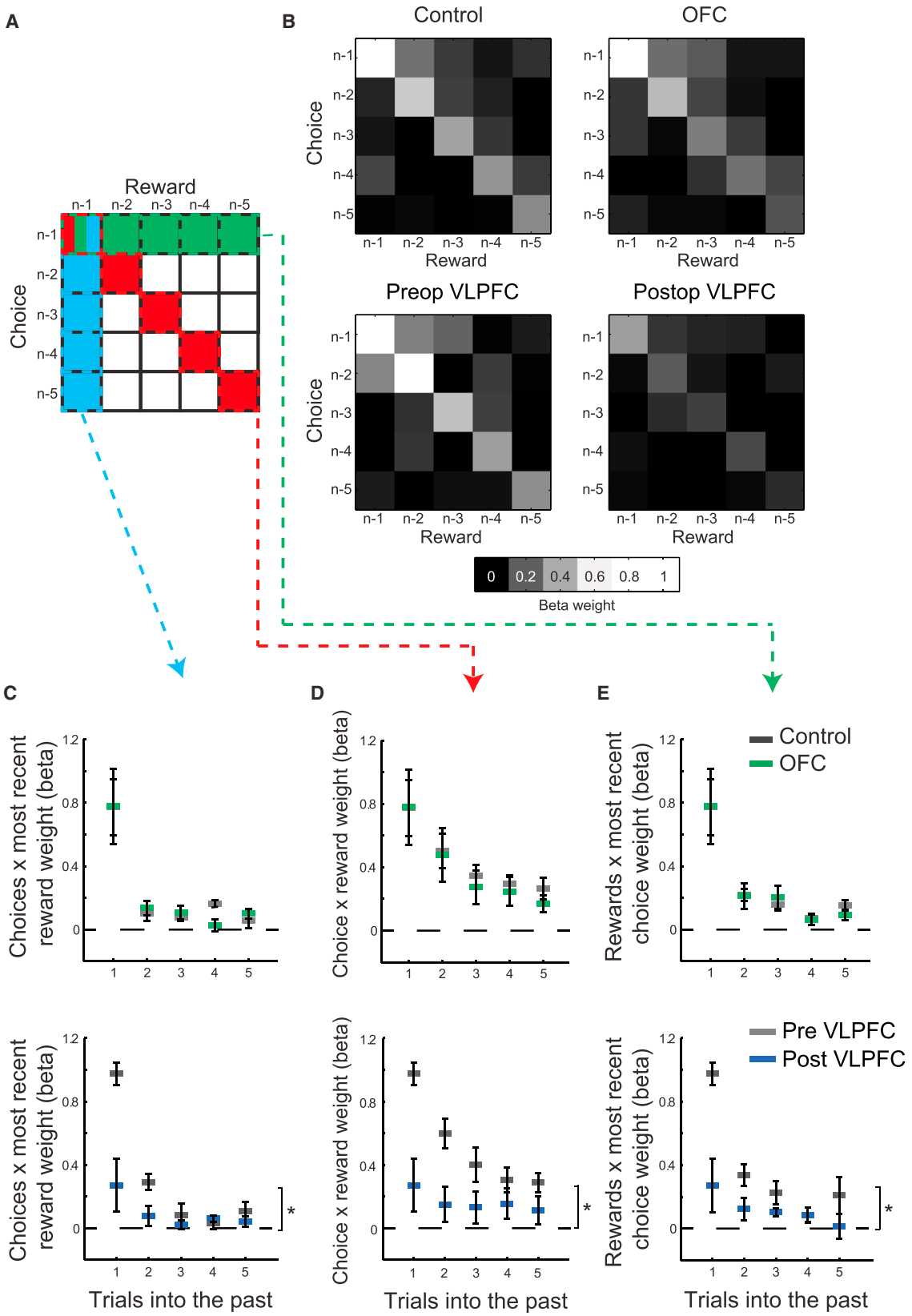
(B) Mean (± SEM) probability of switching on rewarded (darker shading) or unrewarded trials (lighter shading) for unoperated controls (gray, n = 8), monkeys with OFC lesions (green, n = 4), and monkeys before (gray, n = 4) and after VLFPC lesions (blue, n = 4).

Error bars represent SEM. Symbols show scores of individual subjects.

Monkeys with lesions of the OFC exhibited a pattern almost identical to that of the unoperated control monkeys; their current choices were strongly influenced by previous choices and their contingent outcomes (compare left and right of the top of Figure 4B and also see top of Figure 4D). Not only were these monkeys able to use contingent associations between choices and outcomes to guide subsequent choices (unoperated controls versus OFC, effect of group or group by trial interaction, F values < 0.3, $p > 0.6$; Figures 4B and 4D, top), but their choices were also influenced by noncontingent associations (either comparison effect of group or group by trial interaction, F values < 1, $p > 0.6$; Figures 4C and 4E, top). This pattern of results suggests that both contingent and noncontingent learning mechanisms were intact in monkeys with excitotoxic lesions of the OFC.

In contrast, monkeys with lesions of the VLPFC had a profound impairment in contingent learning (Figures 4B and 4D, bottom). The association between previous choices and the outcomes that contingently followed had virtually no influence on monkeys' subsequent choices (preoperative versus postoperative VLPFC, surgery by trial effect, $F(5,15) = 6.94$, $p < 0.01$; postoperative VLPFC, effect of trial, $F(5,15) = 1.61$, $p > 0.25$; Figure 4D, bottom). Lesions of the VLPFC also affected noncontingent learning mechanisms. This was true for both associations between previous choices and the most recent outcome (Figure 4C; surgery by trial interaction, $F(5,15) = 10.81$, $p < 0.01$) as well as between the most recent choices and previous outcomes (Figure 4E; surgery by trial interaction, $F(5,15) = 5.76$, $p < 0.01$).

In three additional experiments, we confirmed that the deficit exhibited by monkeys with VLPFC lesions on the three-choice probabilistic learning task was stable over time and could not be attributed to the order of testing (retest over a year after the initial lesion, contingent learning – preoperative versus postoperative test 2, test by trial effect, $F(5,15) = 9.4$, $p < 0.001$, Supplemental Information; Figure S4); that the deficit was not simply to



(legend on next page)

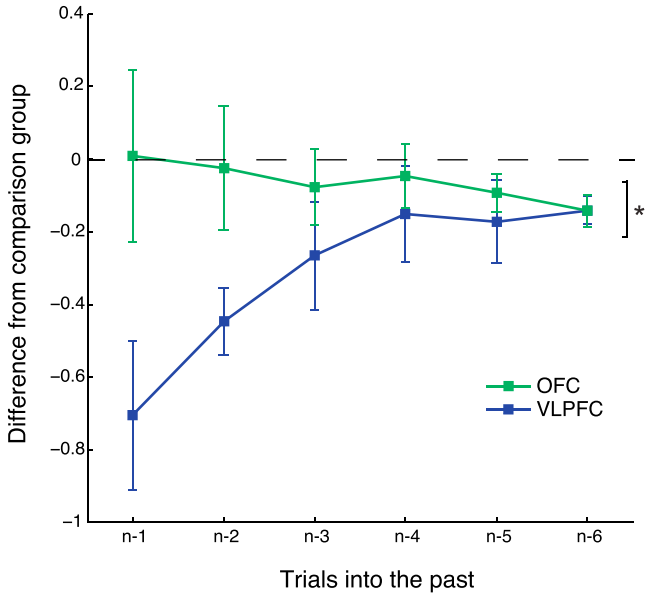


Figure 5. Direct Comparison of Contingent Learning in Monkeys with OFC and VLPFC Lesions

Mean (± SEM) contingent learning difference score for monkeys with OFC (green, n = 4) and VLPFC lesions (blue, n = 4). For each subject, we computed difference scores based on the beta weights from the logistic regression that reflect contingent associations (red cells in Figure 4A) as follows: for monkeys with excitotoxic OFC lesions, the control group mean was subtracted from each OFC lesion monkey's individual score, whereas, for monkeys that received VLPFC lesions, difference scores were computed as the difference between each subject's preoperative and postoperative test scores. Negative scores reflect a decrease in performance relative to controls/preoperative data. *p < 0.05. Error bars represent SEM.

the result of an inability to flexibly alter stimulus-outcome associations, as indexed by the good performance of this group on an object discrimination reversal learning task with deterministic feedback (Supplemental Information; effect of group, $F(1,10) = 0.06$, $p > 0.8$; Figure S5); and that monkeys with VLPFC lesions were able to learn the prevailing stimulus-outcome associations when the difference between the three options was set at the extreme probabilities (1.0, 0.0, 0.0) and were stable over trials (Figure S6; see also Figure 6A).

Finally, to confirm that there was a dissociation between monkeys with OFC and VLPFC lesions in the ability to contingently associate choices and outcomes, we conducted an additional analysis directly comparing performance. To account for the additional training in the VLPFC group, we computed difference scores based on the beta weights from the logistic regression that reflect contingent associations (red squares in Figure 4A) as follows: for monkeys with excitotoxic OFC lesions, the control group mean was subtracted from each OFC lesion monkey's individual score, whereas, for monkeys that received VLPFC lesions, difference scores were computed as the difference between each subject's preoperative and postoperative test scores. Comparison of these difference scores revealed that monkeys with VLPFC lesions differed from monkeys with OFC lesions (Figure 5, difference score OFC versus VLPFC, group by trial interaction, $F(5,30) = 5.05$, $p < 0.005$). Taken together, these data show that the VLPFC, but not the OFC, is required for choosing the best option when choices are guided by reward probability.

Experiment 2: Updating the Desirability of Predicted Outcomes

To determine how the OFC and VLPFC contribute to choices based on desirability, monkeys were tested on a stimulus-based reinforcer devaluation task (Málková et al., 1997). This task measures the ability of monkeys to choose between visual stimuli associated with different food reward based on current biological needs. In contrast to the probabilistic learning task, in which the history of choices and outcomes provides information about the best option and the value of the outcome (a single food pellet) is stable, in the devaluation task, the current value of the food outcome guides choices between visual stimuli. In this experiment, we used three-dimensional objects as visual stimuli.

Over a number of weeks, monkeys learned to discriminate 60 pairs of objects for a food reward. One of the objects in each pair was always rewarded with either food 1 (e.g., peanuts, 30 objects) or food 2 (e.g., M&Ms, 30 objects). Despite a profound impairment in learning probabilistic associations, monkeys with VLPFC lesions learned to discriminate the object pairs at the same rate as controls and monkeys with OFC lesions (effect of group, $F(2,13) = 2.03$, $p > 0.1$; Figure 6A).

We then employed a selective satiation procedure intended to devalue one of the two foods and tested whether monkeys were able to shift their choices of objects to obtain the higher-value outcome. Specifically, following the selective satiation procedure, monkeys were presented with pairs of objects, one object each associated with food 1 and food 2. The effects of devaluation were quantified by calculating the extent to which monkeys shifted their choices toward objects associated with the higher-value food relative to baseline choices. A higher proportion of shifted choices reflects a greater sensitivity to the current value of the foods, updated on the basis of recent and

Figure 4. VLPFC, but Not OFC, Lesions Disrupt Contingent and Noncontingent Learning

(A) Schematic of the full matrix of five previous choices and corresponding rewards received for those choices. The matrix components highlighted in red along the diagonal represent the influence of previous choices and their contingent rewards on subsequent choices. Components highlighted in green represent the influence of reward from the previous five trials and the most recent choice on subsequent choices, whereas those highlighted in blue represent the influence of the five previous choices and the rewards from the previous trial on the subsequent choices.

(B) Matrix plots showing the influence (beta weightings from logistic regression) of all combinations of the five previous choices and rewards on subsequent choice for control monkeys (n = 8), monkeys with OFC lesions (top, n = 4), and monkeys before and after VLPFC (bottom, n = 4). Lighter shading is associated with higher beta weights.

(C–E) Mean (±SEM) raw beta weights from the vertical (C, blue), diagonal (D, red), and horizontal (E, green) parts of the matrix for controls (gray, n = 8) and monkeys with OFC lesions (green, top, n = 4) and monkeys before (gray) and after VLPFC lesions (blue, bottom, n = 4). *p < 0.05. Error bars represent SEM.

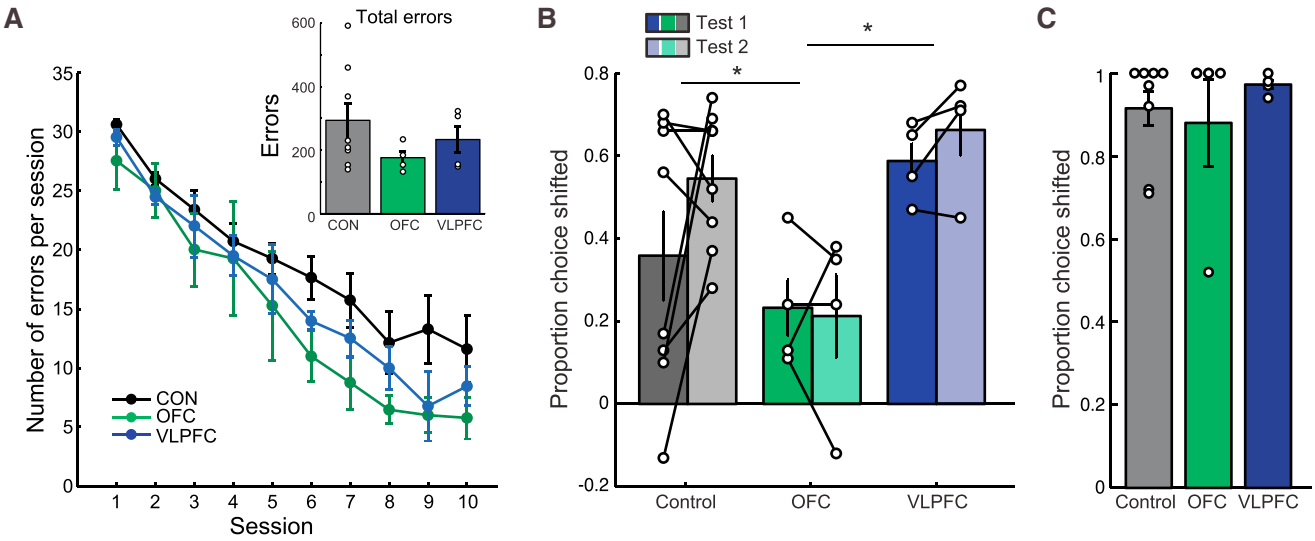


Figure 6. OFC Lesions, but Not VLPFC Lesions, Disrupt the Ability to Choose According to Outcome Value on the Reinforcer Devaluation Task

(A) Mean (± SEM) number of errors for each group during the first 10 sessions of the 60-pair discrimination learning. The inset shows the total errors to criterion for unoperated controls (n = 8), monkeys with OFC lesions (n = 4), and monkeys with VLPFC lesions (n = 4).

(B and C) Mean (± SEM) proportion shifted for unoperated controls (gray bars, n = 8), monkeys with OFC lesions (green bars, n = 4), and monkeys with VLPFC lesions (blue bars, n = 4) during (B) the two reinforcer devaluation tests and (C) a control test where only foods (no objects) were presented for choice. Symbols show scores of individual subjects. *p < 0.05. Error bars represent SEM.

selective satiation. For example, a proportion-shifted score of 0 corresponds to no change in object choice, whereas a score of 1 corresponds to all object choices being shifted away from the devalued food. Two tests, carried out approximately a month apart, were conducted (test 1 and test 2). Each test took into account choices after food 1 and food 2 were devalued, which was assessed in separate sessions.

Both unoperated control monkeys and monkeys with lesions of the VLPFC were able to update and use the current biological value of food reward to guide their choices (Figure 6B). In contrast, monkeys with lesions of the OFC chose stimuli associated with the sated food at a much higher rate, as reflected by the lower proportion of shifted choices (effect of group ($F(2,13) = 4.59$, $p = 0.031$; post hoc least significant difference [LSD], control versus OFC: $p = 0.044$; OFC versus VLPFC: $p = 0.011$; VLPFC versus control: $p = 0.258$). Because monkeys with VLPFC lesions were tested both before and after lesions were made, we also compared their pre- and postoperative performance. This further confirmed that lesions did not affect the ability to update the value of a specific food reward to guide choices (effect of surgery, $F(1,3) = 0.41$, $p > 0.8$).

A control test revealed that, when given the opportunity to make visual choices between two foods after selective satiation, monkeys consistently chose the higher-value (nonsated) food (Figure 6C; effect of group, $F(2,13) = 0.315$, $p = 0.735$). Thus, the deficit in monkeys with OFC lesions was due to an inability to link objects with the current value of the food (or some feature of the food) as opposed to an inability to discriminate the foods or a disruption of satiety mechanisms. In summary, lesions of the OFC, but not of the VLPFC, affected the ability to use the current, updated desirability of a predicted outcome to guide choices.

Comparison of Performance in Experiments 1 and 2

Finally, to provide strong evidence for a double dissociation of function between the OFC and VLPFC, we directly compared performance across the two tasks. Here we conducted an ANOVA using the difference scores computed for the two most recent trials (n-1 and n-2) from the direct comparison of contingent learning in the OFC and VLFPC groups in the three-choice learning task (Figure 5) and the proportion-shifted scores from the two devaluation tests (Figure 6B). This confirmed a double dissociation of function between the VLPFC and OFC on the three-choice probabilistic learning and reinforcer devaluation tasks (task by group interaction, $F(1,6) = 13.86$, $p < 0.05$).

DISCUSSION

The present findings reveal selective and independent contributions of two parts of the granular prefrontal cortex (PFC) in primates. In experiment 1, we found that the VLPFC, but not the OFC, is necessary for updating representations of stimulus-outcome probabilities (Figures 2, 3, 4, and 5). In experiment 2, we found that the OFC, but not the VLPFC, is necessary for updating representations of stimulus-outcome desirability based on current biological states and needs (Figure 6B). Taken together, our findings indicate that, although both the VLPFC and OFC guide choices based on representations of outcome values, they contribute to updating these representations in different ways. The VLPFC is critical for guiding choices based on updated outcome probability, a property that reflects the potential availability of beneficial outcomes, whereas the OFC is necessary for guiding choices based on current biological value, a property that reflects the desirability of a specific outcome.

VLPFC

Neurons in the VLPFC, especially those in area 12, encode different aspects of outcomes during decision-making, including risk (Kobayashi et al., 2010), and a number of studies have suggested that neural activity in this area is linked to external task variables or attentional processes (Kennerley and Wallis, 2009; Rich and Wallis, 2014). Consistent with this idea, fMRI studies in macaques have reported activations in the VLPFC that reflect stimulus value in a two-choice probabilistic learning task (Kaskan et al., 2017). Furthermore, activations in the VLPFC encode adaptive responding (a win-stay, lose-shift strategy) in the context of object reversal learning (Chau et al., 2015).

A straightforward account for the impairment on the three-choice probabilistic learning task is that the VLPFC is important for associating, at the time of feedback, particular visual stimuli (or the choice of a given stimulus) with the outcome that occurs on a specific trial. Walton et al. (2010), referred to these contingent associations in terms of credit assignment, suggesting that the OFC is necessary for updating valuations based on memories of individual events. Although the concept of credit assignment has several variants, Walton et al. (2010) emphasized the correct attribution of a beneficial outcome to the stimulus or choice. On the basis of our results, we embrace many of their conclusions but substitute the VLPFC for the OFC, and the same substitution probably applies to human performance on probabilistic reward tasks as well (Camille et al., 2011; Hornak et al., 2004).

VLPFC lesions affected both contingent and noncontingent learning (Figure 4), but only under conditions of dynamic, stochastic stimulus-outcome associations. When the association between stimuli and outcomes was deterministic (Figure 6A) or static (Figure S5), or when there was clearly a best option (first 150 trials of schedule 1; Figure S2), monkeys with VLPFC lesions were not impaired. This was true even when such deterministic associations between stimuli and outcomes were reversed (Figure S5). This latter finding means that neither the OFC nor the VLPFC are required for object discrimination reversal learning in macaques. This indicates that the ''classic'' impairment seen after aspiration lesions of the OFC is not due to a single area but likely caused by disconnection of a number of areas from the PFC, potentially including the medial striatum, mediodorsal thalamus, and/or neuromodulatory systems (Clarke et al., 2004, 2008; Groman et al., 2013; Iversen and Mishkin, 1970; Roberts et al., 1990).

We also note that our findings are qualitatively and quantitatively different from those following excitotoxic VLPFC lesions in marmosets (Rygula et al., 2010). Specifically, Rygula et al., (2010) reported that the VLPFC was required for reversing new postoperatively acquired associations, but not associations learned before lesions, in a deterministic reversal learning task. Because the deficits in marmosets are seen only during the reversal phase on the task, they are clearly different from what we report here: VLPFC lesions disrupted probabilistic learning before any reversal in stimulus-reward contingencies. Further, the findings of Rygula et al. (2010) would predict that VLPFC lesions should degrade object discrimination reversal learning performance when monkeys had to learn and reverse associations with novel stimuli; however, we observed no deficit in this situation (Figure S5).

It is more difficult to explain why, despite being primates that have a comparably differentiated prefrontal cortex (Burman and Rosa, 2009; Carmichael and Price, 1994), we see differences between the effects of VLPFC lesions in macaques and marmosets. We note that VLPFC is one of the brain areas where the greatest differential expansion has occurred between these two lineages (Chaplin et al., 2013). One possibility is that—since their last common ancestor more that 30 million years ago—the VLPFC has developed divergent functions in macaques and marmosets, partly as a consequence of independent and differential expansion and partly as a consequence of corresponding changes in anatomical connections. This possibility is bolstered by the knowledge that an expansion of the prefrontal cortex occurred independently in macaques more recently than 15 million years ago (Gonzales et al., 2015). A related possibility is that the foraging niche of the two species has driven these areas to subserve divergent functions. Common marmosets (*Callithrix jacchus*) primarily eat tree sap and insects, foods that require more localized foraging in home ranges of between 1–6 hectares (Hubrecht, 1985; Scanlon et al., 1989). Rhesus macaques (*Macaca mulatta*), in contrast, feed on seeds, bark, cereals, buds, and fruit, which requires more distant foraging. Consequently, their home ranges are much larger than those of marmosets: up to 1,500 hectares (Lindburg, 1971). It is possible that a difference in foraging range placed dissimilar selective pressures on the VLPFC in macaques and marmosets, a point we take up later.

Although our results from VLPFC lesions resemble most of the effects that Walton et al. (2010) attributed to OFC lesions, they differ with respect to noncontingent learning. Their aspiration lesions of the OFC affected contingent learning but left noncontingent mechanisms intact. In contrast, our VLPFC lesions affected both contingent and noncontingent learning (Figure 4). One possible explanation for this difference is that their aspiration lesions of the OFC only disrupted fibers connected to the VLPFC that coursed through the uncinate fascicle and left intact the gray matter of the VLPFC as well as many of its connections. Accordingly, the remaining functionality of the VLPFC might have been sufficient to support noncontingent learning.

Additional possibilities involve the ITI, which was slightly longer in our case than in the experiment of Walton et al. (2010) (5 versus 2 s) and the fact that, in our experiment, the chosen stimulus was not re-presented in the absence of the other stimuli after choice. We think that these differences in task parameters provide unlikely explanations for the difference in findings on noncontingent learning, but they merit further investigation because they suggest a mnemonic component to the deficit following VLPFC lesions.

Our conclusions regarding the role of the VLPFC agree with the known anatomical connections of this area, which receives highly processed visual information from the inferior temporal cortex (IT) as well as inputs from the amygdala, OFC, and other outcome-related structures (Carmichael and Price, 1995a, 1995b). In this view, VLPFC underlies the ability to link the kinds of representations housed in the IT—mid-level visual feature conjunctions of color, shape, glossiness, translucence, and texture—with the memory of an outcome that appeared to be caused by the choice of a stimulus that had these features.

A related role of the VLPFC in probabilistic learning involves its role in top-down selective attention. VLPFC damage has been linked to reduced attentional selection, as evidenced by impairments in shifting between stimulus dimensions (Buckley et al., 2009; Dias et al., 1996), reduced performance on tasks requiring allocation of attention to specific visual cues (Rossi et al., 2007; Rushworth et al., 2005), and poor implementation of vision-based rules in the absence of either discrimination or working memory impairments (Baxter et al., 2009; Bussey et al., 2001; Rushworth et al., 1997). Accordingly, the impairment we observed on the three-choice probabilistic learning task could result from a deficit in attentional selection, learning, or some combination of the two. A recent study in humans supports the idea that the VLPFC plays a role in attentional selection (Vaidya and Fellows, 2016).

Notably, activation related to the win-stay, lose-shift rule in macaques was found in a relatively restricted region of the VLPFC immediately lateral to the lateral orbital sulcus, in area 12o (Chau et al., 2015). As Figure 1 shows, this area was included in our VLPFC lesion, although, in a descriptive sense, it lies mostly on the orbital surface of the primate frontal lobe. So inclusion of area 12o as part of the OFC can lead to different conclusions about the OFC than the ones advanced here. In prior work, the OFC has usually been defined as areas 11, 13, and 14, and we adhere to that view here. However, if a part of the VLPFC, area 12o, is included in the OFC, then conclusions about its functional specializations will need to be adjusted to take this redefinition into account. We do not know which parts of our VLPFC lesions caused the impairment reported here, and additional work might be directed to a more precise identification of the crucial region or regions.

Granular OFC

A number of neurophysiological studies have shown that OFC neurons carry signals related to previous choices, outcome history, or both (Kennerley et al., 2011; Simmons and Richmond, 2008; Tsujimoto et al., 2009). The findings reported here indicate that these signals are not necessary for learning about stimulus-outcome probabilities. Instead of the OFC, the VLPFC is required for updating these representations. Our current findings augment those already in the literature on the OFC by demonstrating intact learning of a three-choice probabilistic stimulus-outcome task after complete, excitotoxic OFC removals as well as by showing impaired devaluation-based choice shifts in the same group of monkeys. This pattern of spared and impaired abilities after excitotoxic OFC lesions helps establish a double dissociation of function between the OFC and VLPFC.

The performance of one of the monkeys with an OFC lesion (case 1) differed from the others in the group on the three-choice probabilistic learning task (Figures 2 and 3; Figures S2 and S3). However, although different from the other monkeys that received excitotoxic OFC lesions, this subject rarely scored outside of the range of the unoperated controls. Further, there was no relation between lesion volume and performance, again suggesting that the poor performance of this monkey was not related to the OFC lesion.

VLPFC-OFC Cooperativity

The separate processing of outcome availability and desirability in the VLPFC and OFC, respectively, has implications for models of PFC function during choice behavior. Notably, within the OFC, our previous work shows that lateral OFC areas 11 and 13, not medial OFC area 14, are essential for registering changes in the value of outcomes (Rudebeck and Murray, 2011). In addition, the medial PFC, including the medial OFC (area 14) and medial frontal pole cortex (the medial part of area 10), are involved in comparing different options for choice (Blanchard et al., 2015; Fellows and Farah, 2007; Noonan et al., 2010; Rudebeck and Murray, 2011). Accordingly, one possibility is that the medial PFC receives converging signals from the VLPFC and OFC. The former could convey information about outcome probabilities; the latter would provide information about the current desirability of a specific outcome. The combination of these types of information, along with other valuation-related variables such as magnitude and effort costs, could then guide foraging choices. In line with this idea, the medial PFC receives projections from both the OFC and VLPFC (Carmichael and Price, 1996). In addition, fMRI studies in macaques and humans have found activations in the medial PFC that are modulated not only by outcome contingency (Kaskan et al., 2017; Tanaka et al., 2008), delays in receiving an outcome (Kable and Glimcher, 2007), and the current biological value of outcomes (Howard and Kahnt, 2017) but also by comparison between alternative outcomes (Boorman et al., 2009).

Interpretational Limitations

As is always the case in lesion experiments, the interpretation of results can be compromised by neuroplastic adaptations in the remaining brain areas and connections. However, the effects of VLPFC lesions were evident over a year after surgery (Figure S4), and our earlier work showed that lesions of the OFC produce enduring effects on the devaluation task (Rhodes and Murray, 2013; Rudebeck et al., 2013). The remaining brain structures that contribute to updating outcome valuations, either desirability or availability, appear to have a poor ability, if any, to compensate for the loss of representations established, updated, and maintained by neuronal networks that depend on either the OFC or VLPFC. Therefore, these parts of the granular PFC seem to provide a significant advantage over the remainder of the brain. We close with a consideration of this topic.

Comparative Analysis

Comparative neuroanatomy indicates that the granular OFC and VLPFC arose at different times during primate evolution (Preuss and Goldman-Rakic, 1991) and that they have different connectional fingerprints (Neubert et al., 2014, 2015; Passingham et al., 2002). According to Preuss and Goldman-Rakic (1991), the granular OFC emerged early in primate evolution, and it connects preferentially with the perirhinal cortex and agranular OFC (Kondo et al., 2005; Saleem et al., 2008). The former provides it with visual representations at the level of whole objects and the latter with olfactory, gustatory, and visceral signals (Carmichael and Price, 1996). These inputs suggest that the granular OFC represents conjunctions of outcome features, such as visual appearance and taste, an assumption confirmed by neurophysiological studies in macaque monkeys (Rolls and Baylis, 1994). This enhanced capacity, and especially the contribution from fine-grained visual features of outcomes, probably

provided early primates with a selective advantage in making local foraging choices.

In contrast to the emergence of the granular OFC in early primates, the VLPFC evolved later, sometime during anthropoid evolution (Preuss and Goldman-Rakic, 1991). Rather than the perirhinal cortex, the VLPFC is preferentially connected with the IT (Kondo et al., 2005; Saleem et al., 2014), which supplies it with visual signals at a level of hierarchy between that of whole objects and low-order feature conjunctions or elemental features. Accordingly, the VLPFC probably provided a selective advantage for foraging choices made at a distance, a mode of decision-making that became especially important as anthropoids became large, far-ranging animals (Murray et al., 2017). As we noted earlier, modern macaques differ from marmosets in that the former forage over large home ranges, whereas the latter forage locally. When foraging at a distance, information about a resource's fine-grained visual properties, smell, and taste are less important (because of distance) or unavailable. A role for the VLPFC in representing reward probability may have arisen because this area provided an advantage in estimating resource availability at distant locations, based on visual signals from the IT or acoustic signals from the superior temporal cortex.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Apparatus and Materials
  - Surgery
  - Lesion Assessment
  - Behavioral Testing
  - Three-Choice Probabilistic Learning Task
  - Food Preference Testing
  - Reinforcer Devaluation
  - Object Discrimination Reversal Learning
  - Data Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2017.07.042.

AUTHOR CONTRIBUTIONS

P.H.R. and E.A.M. devised and designed the study. D.A.L., P.H.R., and E.A.M. conducted testing and analyzed the data. R.C.S., P.H.R., and E.A.M. performed the surgeries and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Emily Moylan for assistance with data collection and Rachel Reoli for help with performing surgery. We also thank Vincent D. Costa and Steven P. Wise for comments on an earlier version of the manuscript. This work was supported by the Intramural Research Program of the National Institute of Mental Health (ZIAMH002887 to E.A.M.), an NIMH BRAINS award (R01 MH110822 to P.H.R.), a NARSAD Young Investigator Award (#23638 to P.H.R.), a Rosen Family scholarship (to P.H.R.), and generous seed funds from the Icahn School of Medicine at Mount Sinai (to P.H.R.).

Received: March 3, 2017
Revised: June 20, 2017
Accepted: July 31, 2017
Published: August 30, 2017

REFERENCES

Aron, A.R., Robbins, T.W., and Poldrack, R.A. (2004). Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. *8*, 170–177.

Averbeck, B.B., Djamshidian, A., O'Sullivan, S.S., Housden, C.R., Roiser, J.P., and Lees, A.J. (2013). Uncertainty about mapping future actions into rewards may underlie performance on multiple measures of impulsivity in behavioral addiction: evidence from Parkinson's disease. Behav. Neurosci. *127*, 245–255.

Baxter, M.G., Gaffan, D., Kyriazis, D.A., and Mitchell, A.S. (2009). Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. Eur. J. Neurosci. *29*, 2049–2059.

Blanchard, T.C., Hayden, B.Y., and Bromberg-Martin, E.S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron *85*, 602–614.

Boorman, E.D., Behrens, T.E., Woolrich, M.W., and 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.

Buckley, M.J., Mansouri, F.A., Hoda, H., Mahboubi, M., Browning, P.G., Kwok, S.C., Phillips, A., and Tanaka, K. (2009). Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. Science *325*, 52–58.

Burke, K.A., Franz, T.M., Miller, D.N., and Schoenbaum, G. (2008). The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. Nature *454*, 340–344.

Burman, K.J., and Rosa, M.G. (2009). Architectural subdivisions of medial and orbital frontal cortices in the marmoset monkey (Callithrix jacchus). J. Comp. Neurol. *514*, 11–29.

Bussey, T.J., Wise, S.P., and Murray, E.A. (2001). The role of ventral and orbital prefrontal cortex in conditional visuomotor learning and strategy use in rhesus monkeys (Macaca mulatta). Behav. Neurosci. *115*, 971–982.

Camille, N., Tsuchida, A., and Fellows, L.K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. J. Neurosci. *31*, 15048–15052.

Carmichael, S.T., and Price, J.L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. J. Comp. Neurol. *346*, 366–402.

Carmichael, S.T., and Price, J.L. (1995a). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J. Comp. Neurol. *363*, 615–641.

Carmichael, S.T., and Price, J.L. (1995b). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol. *363*, 642–664.

Carmichael, S.T., and Price, J.L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol. *371*, 179–207.

Chaplin, T.A., Yu, H.H., Soares, J.G., Gattass, R., and Rosa, M.G. (2013). A conserved pattern of differential expansion of cortical areas in simian primates. J. Neurosci. *33*, 15120–15125.

Chau, B.K., Sallet, J., Papageorgiou, G.K., Noonan, M.P., Bell, A.H., Walton, M.E., and Rushworth, M.F. (2015). Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. Neuron *87*, 1106–1118.

Clarke, H.F., Dalley, J.W., Crofts, H.S., Robbins, T.W., and Roberts, A.C. (2004). Cognitive inflexibility after prefrontal serotonin depletion. Science *304*, 878–880.

Clarke, H.F., Robbins, T.W., and Roberts, A.C. (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. J. Neurosci. *28*, 10972–10982.

Dias, R., Robbins, T.W., and Roberts, A.C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. Nature *380*, 69–72.

Fellows, L.K., and Farah, M.J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? Cereb. Cortex *17*, 2669–2674.

Gonzales, L.A., Benefit, B.R., McCrossin, M.L., and Spoor, F. (2015). Cerebral complexity preceded enlarged brain size and reduced olfactory bulbs in Old World monkeys. Nat. Commun. *6*, 7580.

Groman, S.M., James, A.S., Seu, E., Crawford, M.A., Harpster, S.N., and Jentsch, J.D. (2013). Monoamine levels within the orbitofrontal cortex and putamen interact to predict reversal learning performance. Biol. Psychiatry *73*, 756–762.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., and Polkey, C.E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. J. Cogn. Neurosci. *16*, 463–478.

Howard, J.D., and Kahnt, T. (2017). Identity-specific reward representations in orbitofrontal cortex are modulated by selective devaluation. J. Neurosci. *37*, 2627–2638.

Hubrecht, R.C. (1985). Home-range size and use and territorial behavior in the common marmoset, Callithrix jacchus jacchus, at the Tapacura Field Station, Recife, Brazil. Int. J. Primatol. *6*, 533–550.

Iversen, S.D., and Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp. Brain Res. *11*, 376–386.

Izquierdo, A., and Murray, E.A. (2007). Selective bilateral amygdala lesions in rhesus monkeys fail to disrupt object reversal learning. J. Neurosci. *27*, 1054–1062.

Izquierdo, A., Suda, R.K., and Murray, E.A. (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J. Neurosci. *24*, 7540–7548.

Izquierdo, A., Suda, R.K., and Murray, E.A. (2005). Comparison of the effects of bilateral orbital prefrontal cortex lesions and amygdala lesions on emotional responses in rhesus monkeys. J. Neurosci. *25*, 8534–8542.

Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. *10*, 1625–1633.

Kaskan, P.M., Costa, V.D., Eaton, H.P., Zemskova, J.A., Mitz, A.R., Leopold, D.A., Ungerleider, L.G., and Murray, E.A. (2017). Learned value shapes responses to objects in frontal and ventral stream networks in Macaque monkeys. Cereb. Cortex *27*, 2739–2757.

Kennerley, S.W., and Wallis, J.D. (2009). Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus. J. Neurophysiol. *102*, 3352–3364.

Kennerley, S.W., Behrens, T.E., and Wallis, J.D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat. Neurosci. *14*, 1581–1589.

Kobayashi, S., Pinto de Carvalho, O., and Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. J. Neurosci. *30*, 534–544.

Kondo, H., Saleem, K.S., and Price, J.L. (2005). Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. J. Comp. Neurol. *493*, 479–509.

Lindburg, D.G. (1971). The rhesus monkey in North India: an ecological and behavioral study. Primate Behavior *2*, 1–106.

Lindgren, B.W. (1993). Statistical Theory, Fourth Edition (Chapman & Hall).

Málková, L., Gaffan, D., and Murray, E.A. (1997). Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. J. Neurosci. *17*, 6011–6020.

Mishkin, M. (1964). Perseveration of central sets after frontal lesions in monkeys. In The Frontal Granular Cortex and Behavior, J.M. Warren and K. Akert, eds. (McGraw-Hill), pp. 219–241.

Mobini, S., Body, S., Ho, M.Y., Bradshaw, C.M., Szabadi, E., Deakin, J.F., and Anderson, I.M. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology (Berl.) *160*, 290–298.

Murray, E.A., Wise, S.P., and Graham, K.S. (2017). The Evolution of Memory Systems: Ancestors, Anatomy, and Adaptations (Oxford University Press).

Neubert, F.X., Mars, R.B., Thomas, A.G., Sallet, J., and Rushworth, M.F. (2014). Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. Neuron *81*, 700–713.

Neubert, F.X., Mars, R.B., Sallet, J., and Rushworth, M.F. (2015). Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. Proc. Natl. Acad. Sci. USA *112*, E2695–E2704.

Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., and Rushworth, M.F. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc. Natl. Acad. Sci. USA *107*, 20547–20552.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. Annu. Rev. Neurosci. *34*, 333–359.

Passingham, R.E., Stephan, K.E., and Kötter, R. (2002). The anatomical basis of functional localization in the cortex. Nat. Rev. Neurosci. *3*, 606–616.

Preuss, T.M., and Goldman-Rakic, P.S. (1991). Myelo- and cytoarchitecture of the granular frontal cortex and surrounding regions in the strepsirhine primate Galago and the anthropoid primate Macaca. J. Comp. Neurol. *310*, 429–474.

Rhodes, S.E., and Murray, E.A. (2013). Differential effects of amygdala, orbital prefrontal cortex, and prelimbic cortex lesions on goal-directed behavior in rhesus macaques. J. Neurosci. *33*, 3380–3389.

Rich, E.L., and Wallis, J.D. (2014). Medial-lateral organization of the orbitofrontal cortex. J. Cogn. Neurosci. *26*, 1347–1362.

Roberts, A.C., Robbins, T.W., Everitt, B.J., Jones, G.H., Sirkia, T.E., Wilkinson, J., and Page, K. (1990). The effects of excitotoxic lesions of the basal forebrain on the acquisition, retention and serial reversal of visual discriminations in marmosets. Neuroscience *34*, 311–329.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. Cereb. Cortex *10*, 284–294.

Rolls, E.T., and Baylis, L.L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. J. Neurosci. *14*, 5437–5452.

Rolls, E.T., Hornak, J., Wade, D., and McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J. Neurol. Neurosurg. Psychiatry *57*, 1518–1524.

Rossi, A.F., Bichot, N.P., Desimone, R., and Ungerleider, L.G. (2007). Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. J. Neurosci. *27*, 11306–11314.

Rudebeck, P.H., and Murray, E.A. (2011). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. J. Neurosci. *31*, 10569–10578.

Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., and Rushworth, M.F. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. J. Neurosci. *28*, 13775–13785.

Rudebeck, P.H., Saunders, R.C., Prescott, A.T., Chau, L.S., and Murray, E.A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. Nat. Neurosci. *16*, 1140–1145.

Rushworth, M.F., Nixon, P.D., Eacott, M.J., and Passingham, R.E. (1997). Ventral prefrontal cortex is not essential for working memory. J. Neurosci. *17*, 4829–4838.

Rushworth, M.F., Buckley, M.J., Gough, P.M., Alexander, I.H., Kyriazis, D., McDonald, K.R., and Passingham, R.E. (2005). Attentional selection and action selection in the ventral and orbital prefrontal cortex. J. Neurosci. *25*, 11628–11636.

Rygula, R., Walker, S.C., Clarke, H.F., Robbins, T.W., and Roberts, A.C. (2010). Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. J. Neurosci. *30*, 14552–14559.

Saleem, K.S., Kondo, H., and Price, J.L. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. J. Comp. Neurol. *506*, 659–693.

Saleem, K.S., Miller, B., and Price, J.L. (2014). Subdivisions and connectional networks of the lateral prefrontal cortex in the macaque monkey. J. Comp. Neurol. *522*, 1641–1690.

Scanlon, C.E., Chalmers, N.R., and da Cruz, M.A.M. (1989). Home range use and the exploitation of gum in the marmoset Callithrix jacchus jacchus. Int. J. Primatol. *10*, 123–136.

Simmons, J.M., and Richmond, B.J. (2008). Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. Cereb. Cortex *18*, 93–103.

Tanaka, S.C., Balleine, B.W., and O'Doherty, J.P. (2008). Calculating consequences: brain systems that encode the causal effects of actions. J. Neurosci. *28*, 6750–6755.

Thorndike, E.L. (1933). A proof of the law of effect. Science *77*, 173–175.

Tsujimoto, S., Genovesio, A., and Wise, S.P. (2009). Monkey orbitofrontal cortex encodes response choices near feedback time. J. Neurosci. *29*, 2569–2574.

Vaidya, A.R., and Fellows, L.K. (2016). Necessary contributions of human frontal lobe subregions to reward learning in a dynamic, multidimensional environment. J. Neurosci. *36*, 9843–9858.

Wallis, J.D. (2007). Orbitofrontal cortex and its contribution to decision-making. Annu. Rev. Neurosci. *30*, 31–56.

Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H., and Rushworth, M.F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron *65*, 927–939.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Ibotenic acid	Sigma-Aldrich	Cat #: I2765; CAS #: 2552-55-8
Ibotenic acid	Tocris	Cat #: 0285; CAS #: 2552-55-8
N-Methyl-D-aspartic acid	Sigma-Aldrich	Cat #: M3262; CAS #: 6384-92-5
Experimental Models: Organisms/Strains		
Rhesus macaque (Macaca mulatta)	National Institute of Mental Health	https://www.nimh.nih.gov/
Software and Algorithms		
MATLAB v.2014a	MathWorks	https://www.mathworks.com/products/matlab.html
SPSS v.22	IBM	https://www.ibm.com/us-en/marketplace/spss-statistics
Ryklin Software	Ryklin Software	http://www.ryklinsoftware.com/behavioral-tasks/
Other		
Pellet dispenser	Med Associates	http://www.med-associates.com/product/pedestal-pellet-dispenser-for-rat-or-primate-190mg/
190 mg food pellets	Bio-Serv	https://www.bio-serv.com/product/DPP_PGB.html

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and request for resources should be directed to and will be fulfilled by the Lead Contact, Dr. Peter H. Rudebeck (peter.rudebeck@mssm.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Sixteen adult rhesus monkeys (*Macaca mulatta*), 14 male and 2 female, served as subjects. Monkeys weighed between 5.1–10.0 kg and all were at least 4.5 years old at the start of testing. Each animal was individually or pair housed, was kept on a 12-h light dark cycle and had access to water 24 hr a day. All experiments were conducted during the light phase. For the first experiment, four monkeys sustained bilateral excitotoxic lesions of OFC and the remaining eight were retained as unoperated controls (CON). For the second experiment, four monkeys that had previously served as unoperated controls received bilateral excitotoxic lesions of VLPFC and were retested on the 3-choice probabilistic learning task. For the reinforcer devaluation task, eight monkeys served as unoperated controls; four had been tested on the three-choice probabilistic learning task, and the other four had not. Data from the monkeys with excitotoxic lesions of OFC on the devaluation task have previously been reported (Rudebeck et al., 2013). Monkeys were randomly assigned to each group. The testing order in which tasks were administered is shown in Figure S7. No statistical test was run to determine the sample size a priori. The sample sizes we chose are similar to those used in previous publications. All procedures were reviewed and approved by the National Institute of Mental Health (NIMH) Animal Care and Use Committee.

METHOD DETAILS

Apparatus and Materials

All apparatus and materials were identical to those described in previous reports on the effects of lesions within the macaque OFC on probabilistic learning and reinforcer devaluation tasks (Izquierdo and Murray, 2007; Izquierdo et al., 2004, 2005; Rudebeck and Murray, 2011).

For the probabilistic learning task, monkeys sat in primate chairs in front of a touch sensitive monitor on which visual stimuli could be presented and monkeys' choices recorded. Reward pellets (190 mg Noyes pellets) were delivered from an automated food dispenser (MED Associates) into a centrally located cup. A computer running custom software (Ryklin Software, New York, USA) controlled stimulus presentation, timing, contingency, and reward delivery.

For reinforcer devaluation, all testing was conducted in a modified Wisconsin General Test Apparatus (WGTA) inside a darkened room. Monkeys occupied a wheeled transport cage in the animal compartment of the WGTA. The test compartment of the WGTA held the test tray, which contained two food wells spaced 235 mm apart. Test material for reinforcer devaluation consisted of 120 objects that varied in size, shape, color, and texture. Food reward for the devaluation task consisted of two of the following six foods: M&M's (Mars candies, Hackettstown, NJ), half peanuts, raisins, craisins (Ocean Spray, Lakeville-Middleboro, MA), banana-flavored pellets (Noyes, Lancaster, NH) and fruit snacks (Giant Foods, Landover, MD).

Two additional novel objects were used for object discrimination reversal learning. For object reversal learning a half peanut served as the food reward.

Surgery

Standard aseptic surgical procedures were used throughout (Rudebeck and Murray, 2011). Under isoflurane anesthesia, a large bilateral bone flap was raised over the region of the prefrontal cortex and a dura flap was reflected toward the orbit to allow access to the orbital surface in one hemisphere. For the excitotoxic OFC lesion, a series of injections was made into the cortex corresponding to Walker's areas 11, 13 and 14 in each hemisphere using a hand-held Hamilton syringe with a 30-gauge needle. Surgery was carried out in two stages, one hemisphere at a time. Injections were made into the cortex on the orbital surface between the fundus of the lateral orbital sulcus and the rostral sulcus on the medial surface of the hemisphere. The rostral boundary of the injections was an imaginary line at the level of the rostral end of the medial orbital sulcus. The caudal boundary of the injections was an imaginary line at the caudal end of the medial orbital sulcus (Figure 1C). For the VLPFC lesion injections were made into the cortex corresponding to Walker's areas 12, 45, and ventral 46 in each hemisphere (Figure 1C). For cases 1, 3 and 4, surgery was carried out in two stages, one hemisphere at a time. For case 2, surgery was completed in a single stage. The lateral boundary of the lesion was just ventral to the lower lip of the principal sulcus and the medial boundary was the fundus of the lateral orbital sulcus. On the inferior frontal convexity, the rostral boundary of the lesion was the rostral tip of the principal sulcus, and the caudal boundary was the caudal end of the principal sulcus. The lesion therefore avoided the frontal eye fields but included the cortex on the anterior bank of the inferior limb of the arcuate sulcus. On the orbital surface, the rostral limit of the lesion was the anterior tip of the lateral orbital sulcus and the caudal limit was the caudal end of the lateral orbital sulcus. At each site 1.0 μl of ibotenic acid (10-15 μg/μl; Sigma or Tocris) or a cocktail of ibotenic acid and N-Methyl-D-aspartic acid (NMDA) (ibotenic acid 10 μg/μl, NMDA 10 μg/μl; Sigma) was injected into the cortex as a bolus. The needle was then held in place for 2–3 s to allow the toxin to diffuse away from the injection site. Injections were spaced approximately 2 mm apart. For OFC, the mean number of injections per hemisphere was ± SEM: 98 ± 9 (Range: 71 – 119), whereas for VLPFC, the mean number of injections per hemisphere was ± SEM: 92 ± 4 (Range: 76 – 102).

Lesion Assessment

Injections of excitotoxins into OFC and VLPFC resulted in hypersignal – visible in T_2-weighted MR scans – in the cortex on the orbital and ventrolateral surface, respectively. For the monkeys with injections into the OFC, hypersignal extended from the fundus of the lateral orbital sulcus, laterally, to the rostral sulcus, medially (Figure 1C; Figure S1). For the VLPFC group, hypersignal extended from the fundus of the lateral orbital sulcus laterally to the principal sulcus (Figure 1C; Figure S1). For brain regions studied so far, the location and extent of excitotoxic lesions is reliably indicated by white hypersignal on T_2-weighted scans. Accordingly, for each operated monkey the extent of hypersignal on coronal MR images between approximately 40 to 26 mm anterior to the interaural plane was plotted onto a standard set of drawings of coronal sections from a macaque brain. The volume of the lesions was then estimated using a digitizing tablet (Wacom, Vancouver, WA).

Behavioral Testing

Prior to surgery all animals were habituated to the WGTA and were allowed to retrieve food from the test tray. For experiment 1, following preliminary training and initial food preference testing, monkeys either received excitotoxic lesions of OFC or were retained as unoperated controls. Following surgery, monkeys were tested on reinforcer devaluation and then the three-choice probabilistic learning task. For the second experiment, four unoperated controls from the first experiment received excitotoxic VLPFC lesions and were retested on the three-choice probabilistic learning task. They were then tested on the reinforcer devaluation task. Over a year after receiveing excitotoxic VLPFC lesions they were retested on the 3-choice probabilistic learning task. Testers conducting the behavioral experiments were, where possible, blind to group assignments.

Three-Choice Probabilistic Learning Task

All testing was conducted while monkeys sat in a primate chair positioned in front of a touch sensitive monitor. In each test session, animals were presented with 3 novel stimuli, which they had never previously encountered, assigned to the three options (A-C). Stimuli could be presented in one of four spatial configurations and each stimulus could occupy any of the three positions specified by the configuration (see Rudebeck et al., 2008). Configuration and stimulus position was determined randomly on each trial thereby ensuring that animals used stimulus identity rather than action- or spatially-based values to guide their choices.

The start of each trial was signaled by the presentation of three stimuli. Animals made their selections by touching one of the stimuli on the screen. The stimuli then disappeared and reward was delivered, or not, according to the programmed schedule. Intertrial intervals were 5 s.

Reward was delivered stochastically on each option according to four predefined schedules (Figure 1B; Figure S3): stable, variable, forward, and backward which have previously been used to probe stimulus-reward learning in macaques (Rudebeck et al., 2008; Walton et al., 2010). The schedules are a predetermined series of reward/no-reward outcomes for each option on each trial of the 300-trial testing session. The likelihood of receiving a reward for choosing an option on each trial was calculated using a moving 20-trial window (±10 trials) and this is what is shown in Figure 1B and Figure S3. The highest probability of receiving a reward on each trial was determined by taking the envelope of these reward probability functions. Whether or not reward was delivered for selecting one option was independent of the other alternatives. Available reward on unchosen alternatives were not held over for subsequent trials. Each animal completed ten sessions for each schedule. Monkeys completed a single 300-trial testing session each day. Testing proceeded at the rate of one session per day for 5-6 days per week. Novel stimuli were used each day. For the four schedules the sessions were interleaved (i.e., day 1, stable1; day 2, variable1; etc.) to ensure the subjects could not learn the underlying reward schedules.

To confirm that the deficit in learning probabilistic reward associations was stable over time and could not be overcome by compensatory mechanisms, monkeys with VLPFC lesions were retested on the three-choice probabilistic learning task over one year after the initial testing. Each monkey completed 5 sessions of each of the four schedules after an initial period where they were re-familiarized with the task (5 completed sessions with stimuli associated with stationary probabilities of 0.8, 0.5, 0.2 of receiving a reward. There was one exception: Monkey 1 was unable to complete testing on schedules 1 and 2 during the retest, meaning that analyses for this monkey only compared performance on schedules 3 and 4 before and over a year after excitotoxic VLPFC lesions (retest). Otherwise, the data were analyzed using identical methods to those used previously.

Food Preference Testing

After habituation to the WGTA, each monkey's preference for six different foods was assessed over a 15-day period. Every day monkeys received 30 trials consisting of pairwise presentation of the six different foods, one each in the left and right wells of the test tray. The left-right position of the foods was counterbalanced. Preferences were determined by analyzing choices within each of the 15 possible pairs of foods over the final five days of testing.

Reinforcer Devaluation

The behavioral methods used were highly similar to those reported before (Rudebeck et al., 2013). The procedure employed object discrimination learning, which set up particular object-outcome associations, followed by reinforcer devaluation tests, in which probe trials gauged the monkeys' ability to link objects with current food value. For the operated groups, all testing was conducted postoperatively.

Object Discrimination Learning

Monkeys were trained to discriminate 60 pairs of novel objects. For each pair, one object was randomly designated as the positive object (S+, rewarded) and the other was designated as negative (S–, unrewarded). Half of the positive objects were baited with food 1. The other half were baited with food 2. For each monkey, the identity of foods 1 and 2 was based on the monkey's previously determined food preferences. The foods selected were those that the monkey valued highly and which were roughly equally palatable as judged by choices in the food preference test.

On each trial, monkeys were presented with a pair of objects, one each overlying a food well, and were allowed to choose between them. If they displaced the S+ they were allowed to retrieve the food. The trial was then terminated. If they chose the S–, no food was available, and the trial was terminated. The left-right position of the S+ followed a pseudorandom order. Training continued until monkeys attained the criterion of a mean of 90% correct responses over 5 consecutive days (i.e., 270 correct responses or greater in 300 trials).

Reinforcer Devaluation Test 1

Monkey's object choices were assessed under two conditions: after one of the foods was devalued, and in normal (baseline) conditions. On separate days we conducted four test sessions, each consisting of 30 trials. Only the positive (S+) objects were used. On each trial, a food-1 object and a food-2 object were presented together for choice; each object covered a well baited with the appropriate food. With the constraint that a food-1 object was always paired with a food-2 object, the object pairs were generated randomly for each session.

Preceding two of the test sessions a selective satiation procedure, intended to diminish the value of one of the foods, was conducted. For the other two test sessions, which provided baseline scores, monkeys were not sated on either food before being tested. The order in which the test sessions occurred was the same for all monkeys and was as follows: 1) baseline test 1; 2) food 1 devalued by selective satiation prior to test session; 3) baseline test 2; 4) food 2 devalued by selective satiation prior to test session.

For the selective satiation procedure a food box filled with a pre-weighed quantity of either food 1 or food 2 was attached to the front of the monkey's home cage. The monkey was given a total of 30 min to consume as much of the food as it wanted, at which point the experimenter started to observe the monkey's behavior. Additional food was provided if necessary. The selective satiation procedure was deemed to be complete when the monkey refrained from retrieving food from the box for 5 min. The amount of time taken in the selective satiation procedure and the total amount of food consumed by each monkey was noted. The monkey was then taken to the WGTA within 10 min and the test session conducted.

Reinforcer Devaluation Test 2

A second devaluation test, identical to the first, was conducted between 44 and 90 days after reinforcer devaluation test 1. Monkeys were retrained on the same 60 pairs to the same criterion as before. After relearning, the reinforcer devaluation test was conducted in the same manner as before.

Reinforcer Devaluation Test 3: Food Choices after Selective Satiation

Shortly after reinforcer devaluation test 2, we assessed the effect of selective satiation on monkey's choices of foods alone (object-based reinforcer devaluation test 3, Figure S7). This test was conducted to evaluate whether satiety transferred from the home cage to the WGTA, and whether behavioral effects of the lesion (if any) were due to an inability to link objects with food value as opposed to an inability to discriminate the foods. This test was identical to both reinforcer devaluation tests 1 and 2, but with the important difference that no objects were presented over the two wells where foods were placed. On each trial of the 30-trial sessions, monkeys could see the two foods and were allowed to choose between them. As was the case for reinforcer devaluation tests 1 and 2, there were four critical test sessions; two were preceded by selective satiation and two were not.

Object Discrimination Reversal Learning

Monkeys with VLPFC lesions were tested postoperatively and their behavior compared to unoperated controls. A single pair of objects, novel at the start of testing, was used throughout object discrimination reversal learning testing. To prevent object preferences from biasing learning scores, both objects were either baited (for half the monkeys in each group) or unbaited on the first trial of the first session of acquisition of the object discrimination. If the object chosen on the first trial was rewarded, it was designated the S+; if not, it was designated the S-. Through trial and error monkeys learned which object was associated with a food reward. Monkeys were tested for 30 trials per daily session for 5-6 days per week. Criterion was set at 93% (i.e., 28 correct responses in 30 trials) for one day followed by at least 80% (i.e., 24/30) the next day. Once monkeys had attained criterion on the initial object discrimination problem, the contingencies were reversed and animals were trained to the same criterion as before. This procedure was repeated until a total of nine serial reversals had been completed. Data were analyzed with repeated-measures ANOVA with factors of surgery (within subject effect, 2 levels) and reversal (within subjects effect, 9 levels).

Data Analysis

To obtain a trial-by-trial estimate of whether monkeys were choosing the best option based on their prior history of choices and reward, we fit a reinforcement-learning model to monkey's choices in schedules 1 and 2. The model was fit separately to the choice behavior from each session producing estimates of stimulus value and choice probability, for each stimulus on each trial, as well as the learning rate and the inverse temperature for each session. The model updates the value, v, of a chosen option, i, based on reward feedback, r in trial t as follows:

$$v(t) = v_i\ (t-1) + \alpha(r(t) - v_i\ (t-1)). \qquad \text{Equation 1}$$

Thus, the updated value of an option is given by its old value, $v_i(t-1)$ plus a change based on the reward prediction error $(r(t) - v_i(t-1))$, multiplied by the learning rate parameter, α. At the beginning of each session, the value, v, of all three novel stimuli is set to zero. The free parameters (the learning rate parameter, α, and the inverse temperature, β, which estimates how consistently animals choose the highest valued option), were fit by maximizing the likelihood of the choice behavior of the monkeys, given the model parameters. Specifically, we calculated the choice probability $d_i(t)$ using the following:

$$d_i(t) = \frac{exp(\beta v_i\ (t))}{\sum_{k=1}^{3} exp(\beta v_k\ (t))}. \qquad \text{Equation 2}$$

And then calculated the log-likelihood as follows:

$$ll = \sum_{t=1}^{T} log \sum_{k=1}^{3} C_k(t) d_k(t). \qquad \text{Equation 3}$$

Where $c_k(t) = 1$ when the subject chooses option k in trial t and $c_k(t) = 0$ for all unchosen options, meaning that the model maximizes the choice probability ($d_k(t)$) of the actual choices the monkeys made. T is the total number of trials that monkeys completed in a session, usually 300. Model parameters were fitted using methods as described in Averbeck et al. (2013). In brief, parameters were optimized by minimizing the log likelihood of the subject's choices using the fminsearch function in MATLAB. Learning rate parameters were drawn from a normal distribution with a mean of 0.5 and a standard deviation of 3. The inverse temperature parameter was drawn from a normal distribution with a mean of 1 and a standard deviation of 5. These distributions were chosen because learning rates in probabilistic settings should be considerably less than 1, given the stochastic nature of reward delivery, and positive inverse temperatures indicate that choices are biased toward higher reward values. Model fits were repeated 1000 times to avoid local minima and no constraints were placed on the estimated parameters. The maximum log-likelihood across the 1000 fits was used as the model's estimate. We then took the choice probabilities on each trial and determined whether monkeys chose the image with the highest choice probability in either the first or second 150 trials in all schedules.

Logistic regression analysis of monkeys choices in the three-choice probabilistic learning task used methods identical to those used in Walton et al. (2010). These analyses were conducted on the data from all 4 reward schedules (Figure 1B; Figure S3). To determine how recently made choices and recently received reward influenced subsequent choices, we conducted three separate logistic regression analyses, one for each potential stimulus (A,B,C) that the monkey could select. From here on we describe the logistic regression analyses for "A" choices, but the same was done for for stimuli B and C. We first constructed vectors for whenever the monkey chose stimulus A, (vector set to 1) and when they chose stimuli B or C (vector set to 0). We then formed explanatory variables (EVs) based on all possible combinations of choices and reward from the recent past, trials n-1 to n-6. For each choice-outcome interaction, the EV was set to 1 when the monkey chose stimulus A and was rewarded. The same EV was set to −1 when either stimulus B or C was chosen and rewarded and set to 0 when no reward was delivered for any choice. A standard logistic regression was then fit to these 36 EVs (i.e., 6 by 6 matrix of all combination of previous choices and outcomes from preceding 6 trials). Of these, the 25 EV constructed from the five most recent trials were of interest whereas the remaining 11 that involved n-6 trials were included as confounding regressors in order to remove the influence of longer term choice/reward trends.

Ultimately, this analysis produced estimates of $\widehat{\beta}_A$ and $\widehat{C}_A$. The analysis was repeated for stimuli B and C, which produced regression weights for each stimulus, $\widehat{\beta}_A$, $\widehat{\beta}_B$, $\widehat{\beta}_c$ and a corresponding set of covariances, $\widehat{C}_A$, $\widehat{C}_B$, $\widehat{C}_C$. Regression weights for each stimulus were combined into a single weight vector using the variance-weighted mean (Lindgren, 1993):

$$\widehat{\beta} = \left(\widehat{C}_A^{-1} + \widehat{C}_B^{-1} + \widehat{C}_C^{-1}\right)^{-1}\left(\widehat{C}_A^{-1}\widehat{\beta}_A + \widehat{C}_B^{-1}\widehat{\beta}_B + \widehat{C}_C^{-1}\widehat{\beta}_C\right)^{-1}. \quad \text{Equation 4}$$

Regression weights from the different groups were then compared using repeated-measures ANOVAs to determine the differential influence of previous choices, outcomes, and combinations between the two within and across the groups of monkeys.

For the reinforcer devaluation task, the proportion shifted relative to baseline for each subject was computed using the following equation:

$$\textit{Proportion shifted} = \frac{(F1_N - F1_D) + (F2_N - F2_D)}{F1_N + F2_N}. \quad \text{Equation 5}$$

F1 and F2 represent the choices of the objects paired with the two food reward in sessions where the foods were devalued (D) and when they were not (N). Nondevalued choices ($F1/2_N$) were based on the average of two baseline sessions conducted in the week prior to the devaluation sessions. Proportion shifted scores for each monkeys were analyzed using repeated-measures ANOVA with factors of test (two levels, within subjects effect), group (three levels, between subjects effect), and interaction effects where appropriate.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were conducted in both SPSS version 22 and MATLAB version 2014a. Unless otherwise stated, we used repeated-measures ANOVA to compare the performance of different groups. Analysis of the control and OFC group data was conducted with group as a between subjects factor, whereas for the monkeys that received VLPFC lesions, data were analyzed with surgery as a within subjects factor. Other within subjects factors used were phase (first versus second 150 trials of each session, 2 levels), schedule (4 levels), reward (2 levels), trial (trials into the past, 6 levels), test (devaluation test 1 versus test 2), reversal (9 levels), and task (2 levels). In all figures, error bars reflect the standard error of the mean.