Supporting Information

Noonan et al. 10.1073/pnas.1012246107

SI Methods

Animals. Four male rhesus macaque monkeys (*Macaca mulatta*) were tested before and after bilateral aspiration lesions of medial orbitofrontal cortex (mOFC). They were compared with three animals with lateral OFC (lOFC) lesions. Results of investigations of the three lOFC animals have previously been reported (1), but the lesions are also shown in Fig S1*B*. Examination of the histology confirmed that the OFC lesions did not include much medial to the medial orbital sulcus apart from in the most posterior sections (Fig. S1). Unlike in a previous report, we have taken care here to show sections from a more extended range of anterior-posterior sections so that the limited nature of the lOFC lesion is clear. In contrast, the mOFC lesions extended between the medial orbital sulcus on the orbital surface and the rostral sulcus on the medial surface. There were no significant differences in preoperative measures of value assignment and value-based decision making (Figs. 2–4). In addition, in Experiment 2, an additional comparison was made between the lOFC lesion animals and a control group that had undergone testing of the fixed and varying schedules in an identical order in matched pre- and postoperative sessions to the lOFC lesion group. All macaques were male, aged between 4 and 10 y, and weighed between 7 and 13.5 kg. All animals were maintained on a 12-h light/dark cycle and had 24-h ad libitum access to water, apart from when they were testing. All experiments were conducted in accordance with the United Kingdom Scientific Procedures Act (1986).

Apparatus. Each monkey sat in a testing room, unrestrained, in a wheeled transport cage placed 20 cm from a touch-sensitive monitor (38 cm wide × 28 cm high) on which visual stimuli could be presented (eight-bit color clipart bitmap images, 128 × 128 pixels) and responses recorded. Rewards (190-mg Noyes pellets) were delivered from a dispenser (MED Associates) into a food well immediately to the right of the touch screen. A large metal food box, situated to the left below the touch screen, contained each individual's daily food allowance (given in addition to the reward pellets), consisting of proprietary monkey food, fruit, peanuts, and seeds, delivered immediately after testing each day. This food was supplemented by a forage mix of seeds and grains given ∼6 h before testing in the home cage. Stimulus presentation, experimental contingencies, reward delivery, and food box opening was controlled by a computer using in-house software.

Surgery and Histology. At least 12 h before surgery, macaques were treated with an antibiotic [8.75 mg/kg amoxicillin, intramuscularly (i.m.)] and a steroidal anti-inflammatory (20 mg/kg methylprednisolone, i.m.) to reduce the risk of postoperative infection, edema, and inflammation. Additional supplements of steroids were given at 4- to 6-h intervals during surgery. On the morning of surgery, animals were sedated with ketamine (10 mg/kg, i.m.) and xylazine (0.5 mg/kg, i.m.), and given injections of atropine (0.05 mg/kg), an opioid (0.01 mg/kg buprenorphine), and a nonsteriodal anti-inflammatory (0.2 mg/kg meloxicam) to reduce secretions and provide analgesia, respectively. The monkeys were also treated with an H2 receptor antagonist (1 mg/kg ranitidine) to protect against gastric ulceration, which might have occurred as a result of administering both a steroid and nonsteroidal antiinflammatory treatments. Macaques were then moved to the operating theater where they were intubated, switched onto isoflurane anesthesia (1–2%, to effect in 100% oxygen), and placed in a head holder. The head was shaved and cleaned using antimicrobial scrub and alcohol. A midline incision was made, the tissue retracted in anatomical layers, and a bilateral bone flap removed. All lesions were made by aspiration with a fine-gauge sucker. Throughout the surgery, heart rate, respiration rate, blood pressure, expired CO_2, and body temperature were continuously monitored. At the completion of the lesion, the wound was closed in anatomical layers. Nonsteroidal anti-inflammatory analgesic (0.2 mg/kg meloxicam, orally) and antibiotic (8.75 mg/kg amoxicillin, orally) treatment were administered for at least 5 d postoperatively. All surgery was carried out under sterile conditions with the aid of a binocular microscope. The wound was closed in anatomical layers. At least 2 wk were allowed for recovery before testing resumed.

The mOFC lesion was made by removing the cortex from the medial orbitofrontal sulcus to the rostral sulcus and included mainly Walker's area 14 (2), but may have included in addition some parts of area 10. In contrast, the lOFC lesion was made by removing the cortex between the medial and lateral orbitofrontal sulci (predominantly Walker's areas 11 and 13 and parts of area 12).

When the animals had completed their testing, they were anesthetized with sodium pentobarbitone and perfused with 90% saline and 10% formalin. The brains were then removed and placed in 10% sucrose formalin until they sank. The brains were blocked in the coronal plane at the level of the most medial part of the central sulcus. Each brain was cut in 50-μm coronal sections. Every tenth section was retained for analysis and stained with Cresyl violet.

Five coronal sections through the frontal lobes are shown in Fig. S1 *A* and *B* in examples of mOFC- and lOFC-lesioned animals, respectively. Actual lesions are shown adjacent to schematics of the intended lesions (see Fig. 1*C* for all animal's lesion positions). As can be observed, the lesions were made as intended. The mOFC lesion included mainly Walker area 14 and the lOFC lesions included mainly Walker areas 11 and 13. The lOFC lesions did, however, include some of the more posterior part of the medial orbitofrontal cortex, although anterior ventromedial prefrontal and mOFC was consistently spared (Fig. S1).

Training Histories. Three lOFC lesion animals were trained, tested, and compared with three animals that acted as unoperated controls for the varying schedules for a study recently published by Walton et al. (1). For the present study, all animals had extensive training on varying reward schedules before testing. With the addition of another animal, three of the unoperated control animals from this experiment went on to become part of the current mOFC-lesion group. Approximately 18 mo separated testing in the Walton et al. (1) experiment and training in the present study. Before testing in the current experiment, all animals in the mOFC lesion group were brought to a criterion of 80% correct on three choice-reversal schedules. We ensured they were at roughly the same preoperative performance level as they were when they acted as unoperated controls. Critically, the two lesion groups did not differ in terms of their preoperative performances (Figs. 2– 4). This finding suggests that any differences in postoperative performance between the two groups cannot be attributed to differences in training histories.

Schedules. During a day's testing session, three novel clipart stimuli were presented on a touch-screen computer. In each trial, three stimuli were presented in a computer-randomized order in one of four possible spatial configurations, and the location of each clipart stimulus within each configuration was also similarly randomized. Each schedule was used five times but the stimuli

were novel in each session. The animals selected a stimulus by touching the screen in the stimulus location. Registration of a response delivered a sucrose pellet according to the probability defined for that stimulus for the schedule under which the animal was currently operating. During the feedback period, the selected stimulus remained onscreen as the unselected options were extinguished. An intertrial interval period preceded the subsequent trial. All animals had extensive experience of three choice touchscreen tasks before surgery and were trained to a criterion of choosing the best, V1, option on 80% of trials in several varying reward probability conditions (Fig. 1B, discussed below).

Fixed. The data from the fixed reward schedules were analyzed as a function of the objective value for all three stimuli and, on that basis, stimuli were defined as V1, V2, and V3 for the best, mid, and worst option, respectively (Fig. 3 *A*–*C*). The number of trials to reach a criterion of 70% of V1 stimulus choices was calculated for each session for each animal, pre- and postoperatively. These totals were averaged across the five sessions and subjected to a two-way repeated-measures ANOVA with factors of Testing Session (pre- and postoperative) and Condition, with levels of V2_HIGH, V2_MID, and V2_LOW. Animals with mOFC and lOFC lesions were compared postoperatively on the same measure in a two-way ANOVA of Condition (V2_HIGH, V2_MID, and V2_LOW) and the between-subjects factor of Group (mOFC and lOFC).

Varying schedules. Three monkeys completed five additional three-choice tasks in which the reward probability associated with each stimulus varied over the course of the session, pre- and postoperatively. As a result of an experimental oversight, the other monkey only completed four varying schedules. All trials completed are included in the regression analysis and softmax analysis. All other analyses are based only on the four schedules that all monkeys completed (Fig. S2).

Credit Assignment Analyses. ***Analysis of the gross deficit induced by an mOFC lesion.*** The data from the varying reward schedules were analyzed as a function of the subjective expected values of all three different stimuli, which were derived using a standard Rescorla-Wagner learning model with a Boltzmann action selection rule. The reward learning rate (α) was fitted individually to each animal's pre- and postoperative data using standard nonlinear minimization procedures. These data were used to estimate the expected value of each of the three stimuli on every trial (the same learning rate was used for all three stimuli). The aim was to identify the best (V1), second best (V2), and worst (V3) stimulus for every trial and determine the probability the animals chose the best option. The log-transformed proportion of V1 choices as function of the total choices was calculated and averaged across monkeys for each third of a testing session. These values were subjected to a repeated-measures ANOVA of Lesion (two levels: before and after) × Split (three levels: trial 1–100, 101–200, 201–300) (Fig. S3).

Effect of past-reward history and past-choice history on the assignment of reward to stimulus choices. The data from the varying schedules were pooled into two condition schedules and reported using parametric repeated-measures analyses, with within-subjects factors of Surgery (two levels: pre- or postsurgery), Condition (two levels), and Reward (two levels: rewarded and unrewarded). When comparing across different reward histories (Fig. 2*B* and Fig. S5), a fourth factor of Reward history was included (four levels: $A^{?}AAB$, $AA^{?}AB$, $AAA^{?}AB$, and $AAAA^{?}B$), where the "?" refers to the presence or absence of reward at the point in the history. Alternatively, when choice history was considered, this additional factor was Choice history (Fig. 2*A* and Fig. S4) with three levels (AB, $A^{2\text{-}3}B$, and $A^{4\text{-}7}B$). Comparing the mOFC and lOFC lesion groups required an additional between-subjects factor of Group (two levels: mOFC and lOFC). Note that options "A," "B," and "C" do not necessarily directly refer to stimuli A, B, and C, as depicted in Fig. 1*B*, but instead to sequences of similar choices (i.e., an "AAB" history could be made up of choice sequences of stimuli AAB, AAC, BBA, CCB, BBC, or CCA). Also note that to correct for positive skew in the data, values were log-transformed before analysis.

For effect of stimulus choice history on credit assignment for the current stimulus choice, see Fig. S4.

For effect of reward proximity on the current choice, see Fig. S5.

Learning rates. The data from the varying reward schedules were analyzed as a function of the subjective expected values of all three different stimuli, which were derived using a standard Rescorla-Wagner learning model with a Boltzmann action selection rule. The reward learning rate (α) was fitted individually to each animal's pre- and postoperative data using standard nonlinear minimization procedures. This process was used to estimate the expected value of each of the three stimuli on every trial. The aim was to identify the best (V1), second best (V2), and worst (V3) stimulus for every trial and determine the probability the animals chose the best option (Fig. S6).

Value Comparison Analyses. ***Analysis of varying schedule data in terms of value differences used in Experiment 1.*** Trials across all varying schedules in which the subjective difference in value between V1 and V2 and between V2 and V3 resembled the value differences in the three fixed schedules were extracted (Fig. 3 *J*–*O*). The fixed V2_HIGH schedule is approximately equivalent to trials drawn from the varying schedules with a small V1V2 value difference (0.1–0.3 difference in reward probability) and a large V2V3 value difference (0.2–0.5 difference in reward probability). Fixed V2_MID was approximated by trials from varying schedules with V1V2 and V2V3 value differences of 0.3 to 0.5 and 0.0499 to 0.2, respectively, but fixed V2_LOW was approximated by trials in which V1V2 and V2V3 value differences were 0.5 to 0.7 and 0 to 0.0499, respectively. Trials were binned according to these divisions in value differences and the probability of choosing V1 was compared across the three value difference bins in a two-way repeated-measure ANOVA of Testing Session (pre- and postoperative) and Value Difference [three levels: three Value Bin (V1V2 small/V2V3 large, V1V2 mid/V2V3 mid, and V1V2 large/V2V3 small)]. The percentage of trials included in this analysis for the mOFC group was 26 and 23% for before and after, respectively, and 32 and 19% for the lOFC group.

Logistic softmax function. To determine whether the value comparisons the animals were making were influenced by the context in which they were made, logistic softmax functions were fitted to each animal's choice data (Fig. 4 *B* and *C*). This process tested whether the animals were choosing optimally in their environment. We defined three positions in the decision context. A stimulus could either be chosen or unchosen. If it was an unchosen stimulus it could act as the irrelevant alternative or the "3rd option". Each trial was only used once in the analysis. The proportions of trials on which a subject chose as a function of its value difference with respect to another option (this could either be positive or negative) were plotted separately for situations when 3rd option was high (top 33% of 3rd option values) and low (bottom 33% of 3rd option values) in value. These distributions were then fitted with a logistic sigmoid function. The function of the logistic curves for when the other option, 3rd option, was a high and low value were compared independently in paired sample *t* tests for the pre- and postoperative data. The behavioral analysis was performed using Matlab 6.5 (MathWorks).

We also investigated both the mOFC- and lOFC-lesioned animals' distribution of choices in the first 150 trials of each day's learning session, pre- and postoperatively, in two independent two-way repeated-measures ANOVAs of Lesion (two levels: pre- and postoperative) and Value Differences (three levels: low,

mid, and high). The groups were compared directly in a three-way repeated-measures ANOVA, which included the additional factor of Group (two levels: mOFC and lOFC).

We also carried out the same analysis after normalizing the values (Fig. S7). The results remained essentially unchanged.

Reward and Error Sensitivity Analyses. For reward and error sensitivity analysis, see Fig. S8.

Average Value and Response Vigor Analyses. For average value and response vigor analysis, see Fig. S9.

1. Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, Rushworth MF (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* 65:927–939.

2. Walker EA (1940) A cytoarchitectural study of the prefrontal area of the macaque monkey. *J Comp Neurol* 73(1):59–86.

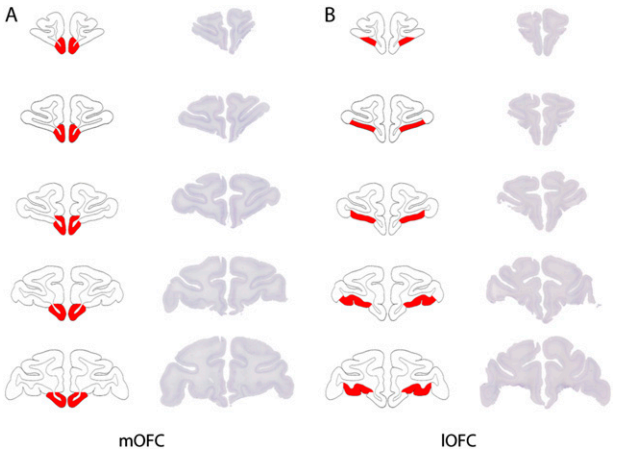

Fig. S1. Intended mOFC and lOFC lesion positions (*Left* of *A* and *B*, respectively) and actual coronal lesion sections for two typical mOFC and lOFC lesioned animals (*Right* of *A* and *B*, respectively).

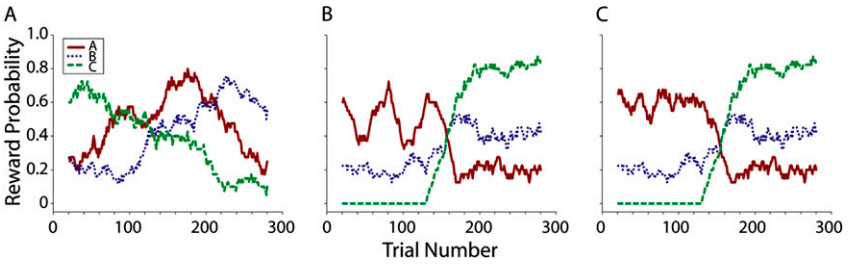

Fig. S2. Schematic representation of the reward probability associated with each stimulus in the three additional varying schedules averaged over 20 running trials. Three stimuli (*A–C*; represented by brown, blue, and green lines respectively) were presented on a touch-sensitive screen, each of which was associated with a probability of reward, which varied across the 300 trial session.

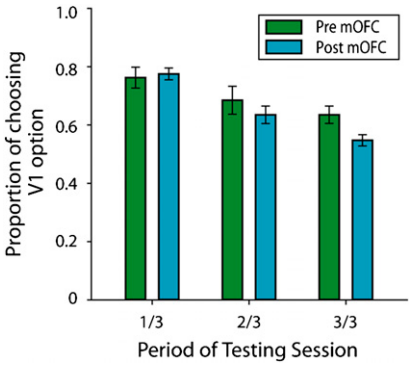

Fig. S3. Gross mOFC lesion deficit on V1 choices as a function of total choice. As similarly reported for lOFC lesions, mOFC lesions lead to a decrement in the proportion of V1 choices, particularly in the later periods of testing sessions (Surgery × Session period interaction: $F_{1,6} = 8.41$, $P = 0.0430$).

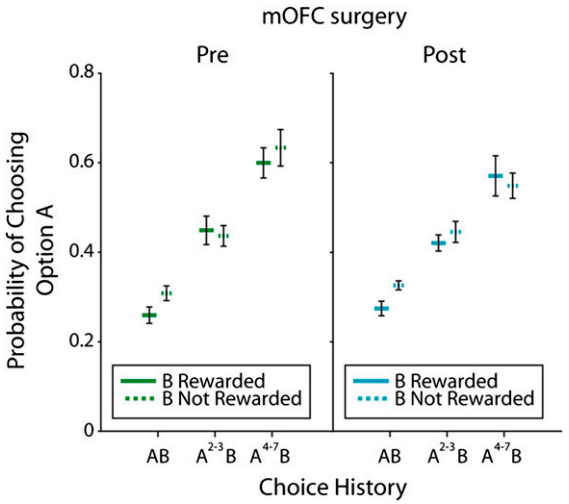

Fig. S4. Influence of past choices of one option (A) on current choice behavior (trial n) in changeable three-armed bandit tasks as a function of reward received for choosing option B on the previous trial (trial n-1). The plots shown in Fig. 2*A* were composed by calculating the difference between rewarded and unrewarded trials shown here. The plots show the likelihood of choosing option A on trial n after either receiving a reward (filled line) or not receiving a reward (dashed line) for choosing option B on the previous trial (n-1). Pre- (green) and postoperative (blue) results are plotted separately (*Left* and *Right*, respectively). Medial OFC and lOFC groups were compared directly in a five-way repeated-measure ANOVA, as described in *SI Methods: Credit Assignment Analysis*, but only mOFC results are presented here; lOFC lesioned data are published elsewhere (1).

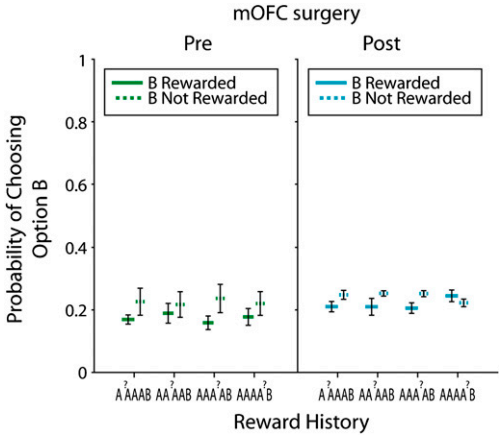

Fig. S5. Likelihood of choosing a particular option on the current trial (n) after having chosen option A on four past trials (n-2 to n-5) and then option B on the previous trial (n-1), plotted as a function of presence or absence of reinforcement on one particular A option in the past (A?). Top row represents the likelihood of choosing option B on trial n when a previous A choice (A?) was either rewarded (filled line) or not rewarded (dashed line). Fig 2*B* was prepared by comparing the difference in probability of choosing option B for rewarded and unrewarded trials, plotted here independently. Pre- (green) and postoperative (blue) results are plotted separately (*Left* and *Right*, respectively). Medial OFC and lOFC groups were compared directly in a five-way repeated measure ANOVA, as described in *SI Methods: Credit Assignment Analysis*, but only mOFC results are presented here; lOFC lesioned data are published elsewhere (1).

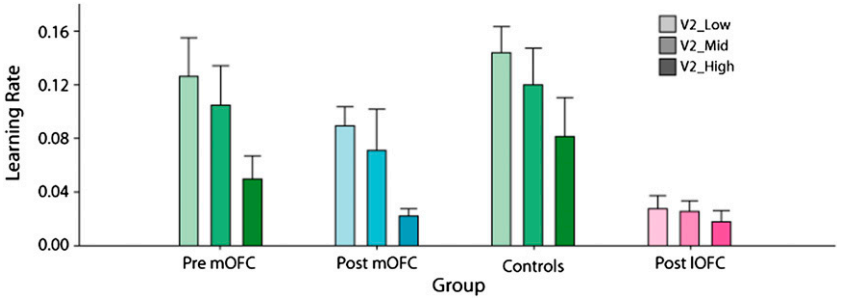

Fig. S6. The reward learning rate in the three fixed reward schedules. Increasing V2 reward probability (i.e., decreasing V1V2 value differences) is represented with increasing darkness of color for the preoperative (green for both mOFC and lOFC) and postoperative animals (blue and pink for mOFC and lOFC, respectively). The reward learning rate (α) increased with increasing V1-V2 value differences in controls and mOFC-lesioned animals but not in lOFC-lesioned animals.

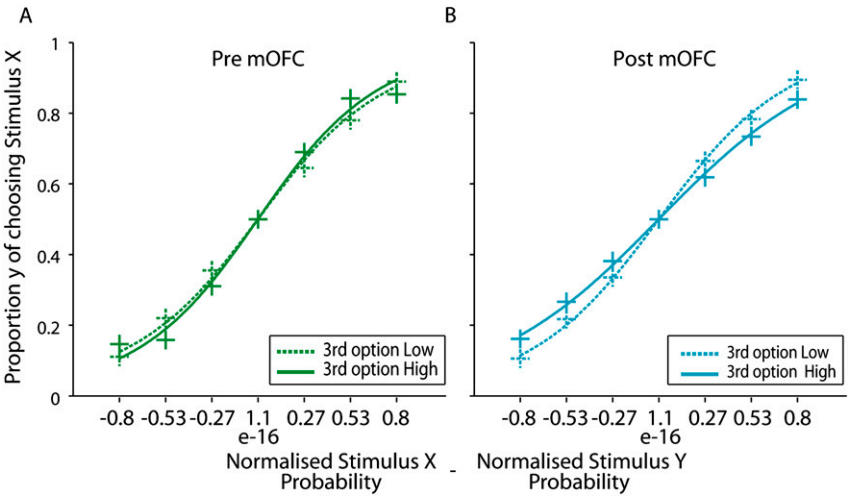

Fig. S7. Proportion of trials on which monkeys chose options as a function of the normalized value difference with respect to one other option in the context of a high (solid line) or low (dashed line) value third option before (*A*) and after (*B*) mOFC lesion.

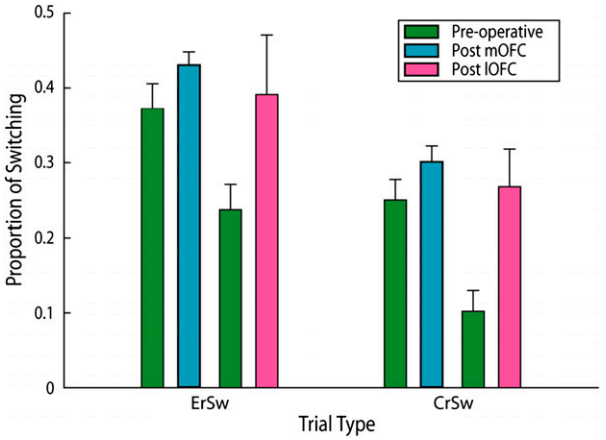

Fig. S8. Error and Reward sensitivity in preoperted control animals (green), mOFC- (blue), and lOFC- (pink) lesioned animals. Percentages switching after trials on which a reward was not delivered (an error; ErSw) and percentage of switching on the stimulus choice after a reward (CrSw) was calculated for all three groups in the four variable reinforcement schedules. Groups were compared in two-way ANOVA of Surgery (two levels: pre- and post-) and Trial type (two levels: ErSw and CrSw) with the between subjects factor of Group (two levels: lOFC and mOFC). This analysis showed that both lesions caused a general increase switching (main effect of surgery: $F_{1,5} = 13.09$, $P = 0.015$).

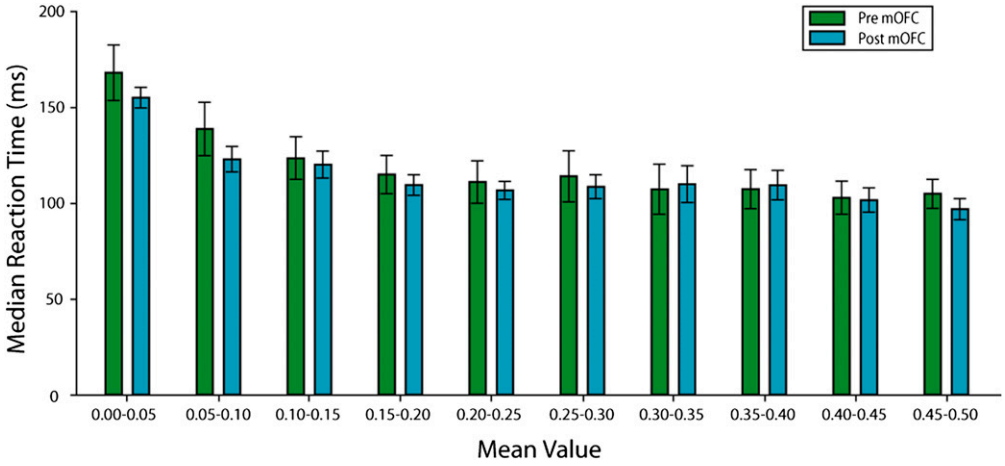

Fig. S9. Niv et al. have shown that response vigor, as indexed by decreased average reaction time, increases with the average value of an environment (1). Slower reacton times constitute greater opportunity costs in higher value environments. If mOFC lesions deprive macaques of any sense of value, then this relationship should be abolished or at least compromised. Pre- and postoperative trials were divided into 10 bins on the basis of the mean value of all three stimuli on that trial, as estimated by the reinforcement learning model described above. Average stimulus value increased by 0.05 from one bin to the next. A comparison of median reaction times on trials in the 10 average value bins before and after surgery revealed a main effect of average value ($F_{7,\ 21} = 10.25$, $P = 0.001$) but no effect of the lesion or interaction between lesion and average value ($P > 0.1$). Essentially the same results were obtained even if the first two bins, when reaction times were longest, were excluded from the analysis.

1. Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.