

ORIGINAL ARTICLE

Effects of Amygdala Lesions on Object-Based Versus Action-Based Learning in Macaques

Craig A. Taswell, Vincent D. Costa, Benjamin M. Basile, Maia S. Pujara, Breonda Jones, Nihita Manem, Elisabeth A. Murray and Bruno B. Averbeck

Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892-4415, USA

Address correspondence to Bruno B. Averbeck, Ph.D. Laboratory of Neuropsychology, NIMH/NIH, Building 49 Room 1B80, 49 Convent Drive MSC 4415, Bethesda, MD 20892-4415, USA. Email: bruno.averbeck@nih.gov.

Abstract

The neural systems that underlie reinforcement learning (RL) allow animals to adapt to changes in their environment. In the present study, we examined the hypothesis that the amygdala would have a preferential role in learning the values of visual objects. We compared a group of monkeys (*Macaca mulatta*) with amygdala lesions to a group of unoperated controls on a two-armed bandit reversal learning task. The task had two conditions. In the What condition, the animals had to learn to select a visual object, independent of its location. And in the Where condition, the animals had to learn to saccade to a location, independent of the object at the location. In both conditions choice-outcome mappings reversed in the middle of the block. We found that monkeys with amygdala lesions had learning deficits in both conditions. Monkeys with amygdala lesions did not have deficits in learning to reverse choice-outcome mappings. Rather, amygdala lesions caused the monkeys to become overly sensitive to negative feedback which impaired their ability to consistently select the more highly valued action or object. These results imply that the amygdala is generally necessary for RL.

Key words: Bayesian modeling, monkeys, reversal learning

Introduction

Learning to execute actions or select objects that lead to rewards is critical for survival. While formal models of reinforcement learning (RL) do not distinguish between these (Averbeck and Costa 2017), there is considerable evidence to support the view that separate neural circuits mediate learning about the value of actions versus objects. Starting in the visual system, there is a distinction between spatial vision and object vision, that has been referred to as the dorsal (spatial) and ventral (object) visual streams hypothesis (Ungerleider and Mishkin 1982). A related view suggests that the distinction between the two systems involves processing information for action versus perception (Goodale and Milner 1992). The anatomical separation between these systems continues into prefrontal cortex (Barbas 1988; Webster et al. 1994) and also through the frontal-basal ganglia-thalamo-cortical loops (Haber et al. 2006; Averbeck et al. 2014). There is also interaction between these circuits

(Serenio and Maunsell 1998), especially when object information is required to select spatially directed actions (Bussey et al. 2002). But to some extent these processing streams are segregated. The anatomy, therefore, suggests that learning to associate rewards with actions may rely more on dorsal circuitry, and learning to associate rewards with objects may rely more on ventral circuitry (Neftci and Averbeck 2019).

RL has often been linked to the ventral striatum (VS). This suggests that the VS underlies both learning to associate rewards with actions and stimuli. There is considerable evidence for the role of the VS in object-based RL (O'Doherty et al. 2004; Costa et al. 2016), particularly when it comes to learning to choose between two positive outcomes that vary in magnitude (Taswell et al. 2018). There has been less evidence for the role of the VS in action selection. When action and object learning have been studied in the same experiment, monkeys with

lesions to the VS had deficits in object but not action-based RL (Rothenhoefer et al. 2017). Other work has shown that the dorsal striatum (DS) plays a role in learning to associate actions (Samejima et al. 2005; Lau and Glimcher 2008; Parker et al. 2016) and action sequences (Seo et al. 2012; Lee et al. 2015) with rewards. This suggests that different neural circuits underlie these two different types of learning, at least in the striatum.

The amygdala has also been shown to play an important role in visual object-based RL (Cardinal et al. 2002; Paton et al. 2006; Hampton et al. 2007; Belova et al. 2008; Costa et al. 2016; Rudebeck, Ripple, et al. 2017) and other forms of reward learning (Baxter and Murray 2002; Salzman and Fusi 2010). Studies in monkeys have shown that lesions of the amygdala lead to learning deficits in a probabilistic reversal learning task (Costa et al. 2016) and a reward magnitude learning task (Rudebeck, Ripple, et al. 2017). And, amygdala lesions lead to a decrease in the information about stimuli associated with rewards, relative to prelesion recordings, in the orbital prefrontal cortex (Rudebeck et al. 2013; Rudebeck, Ripple, et al. 2017). Therefore, there is considerable evidence that supports a role for the amygdala in learning to associate objects with rewards. Furthermore, the amygdala has strong anatomical connections with the ventral visual pathway, and less pronounced connections with dorsal pathway structures (Amaral and Price 1984; Amaral et al. 1992; Neftci and Averbach 2019). Although the amygdala contributes to object reward learning, and has strong links to the ventral visual pathway, it also has links to the dorsal pathway. For example, single neurons in the amygdala code the locations of chosen objects independent of reward expectation (Peck et al. 2014; Costa et al. 2019). In addition, the amygdala projects to cingulate motor areas (Morecraft et al. 2007), which provides a potential route for the amygdala to influence action learning. Whether the amygdala makes a causal contribution to learning to choose rewarded locations, however, has not been directly examined.

To determine the amygdala's role in action- versus object-based RL, we tested four monkeys with excitotoxic lesions of the amygdala on a two-arm bandit reversal learning task used previously to examine learning following VS lesions (Rothenhoefer et al. 2017). This task involved two different types of learning, carried out in blocks of trials. In one block type, the monkeys had to learn to pick the location (action based) that yielded the most rewards, and in the other block type monkeys had to learn to choose the stimuli (object based) that led to the most rewards. We found that the amygdala plays a role in both action- and object-based RL.

Materials and Methods

Subjects

The subjects included 10 male rhesus macaques with weights ranging from 6 to 11 kg. Four of the male monkeys received bilateral excitotoxic lesions of the amygdala. The remaining six monkeys served as unoperated controls. Four out of the six unoperated control monkeys were the same monkeys used in a previous study (Costa et al. 2016). Five out of the six unoperated control monkeys were the same monkeys from an additional study (Rothenhoefer et al. 2017). All remaining monkeys were not previously used in the studies mentioned above. In particular, none of the amygdala lesion monkeys ($n = 4$) were previously used in the studies mentioned above. For the duration of the study, monkeys were placed on water control. On testing days,

monkeys earned their fluid from their performance on the task. Experimental procedures for all aspects of the study were performed in accordance with the Guide for the Care and Use of Laboratory Animals and were approved by the National Institute of Mental Health Animal Care and Use Committee.

Surgery

Four monkeys received two separate stereotaxic surgeries, one for each hemisphere, which targeted the amygdala using the excitotoxin ibotenic acid (for details, see Costa et al. 2016). Injection sites were determined based on structural magnetic resonance (MR) scans obtained from each monkey prior to surgery. After both lesion surgeries had been completed, each monkey received a cranial implant of a titanium head post to facilitate head restraint. Unoperated controls received the same cranial implant. Behavioral testing for all monkeys began after they had recovered from the implant surgery.

Lesion Assessment

Lesion volume estimates were taken by first transforming each subject's T2-weighted scan acquired 1 week postoperatively to the standard NMT (NIMH macaque template; Seidlitz et al. 2018) using AFNI's 3dAllineate function (Cox 1996). We then applied thresholding to identify the area of hyperintensity on the transformed T2-weighted object to isolate a binary mask that corresponded to the area of damage. The masks were visually inspected and manually edited to ensure that they fully captured the areas of hyperintensity on the T2-weighted object. A lesion overlap map was created by summing the binary masks for each hemisphere and displaying the output on the NMT (Fig. 1C).

As intended, all operated monkeys sustained extensive damage to the amygdala, bilaterally; the estimated percent damage ranged from 86% to 95% (Supplementary Table 1, Fig. 1C). Surrounding structures, mainly the entorhinal cortex, sustained inadvertent damage that varied widely in extent (Supplementary Table 1). Based on prior work, the percent damage to entorhinal cortex as estimated from the T2-weighted scans is almost certainly an overestimate (Basile et al. 2017).

Task and Apparatus

We tested rhesus macaques (*Macaca mulatta*) on a probabilistic two-arm bandit reversal learning task. During the experiment, animals were seated in a primate chair facing a computer screen. Eye movements were used as behavioral readouts. In each trial, monkeys first acquired central fixation (Fig. 1A,B). After a fixation hold period of 500 ms, we presented two objects, left and right of fixation. Monkeys made saccades to one of the two objects to indicate their choice. After holding their choice for 500 ms, a reward was stochastically delivered according to one of three reward schedules: 80%/20%, 70%/30%, and 60%/40%. In an 80%/20% reward schedule, one of the choices led to a reward 80% of the time and the other choice led to a reward 20% of the time. The reward schedule and stimuli were used for a total of 80 trials, which constituted one training block. At the beginning of each block, two novel objects were introduced and the block was randomly assigned a reward schedule; this assignment remained constant throughout the entire block. In addition, on each trial the location of "best" object, left or right of fixation, was randomized.

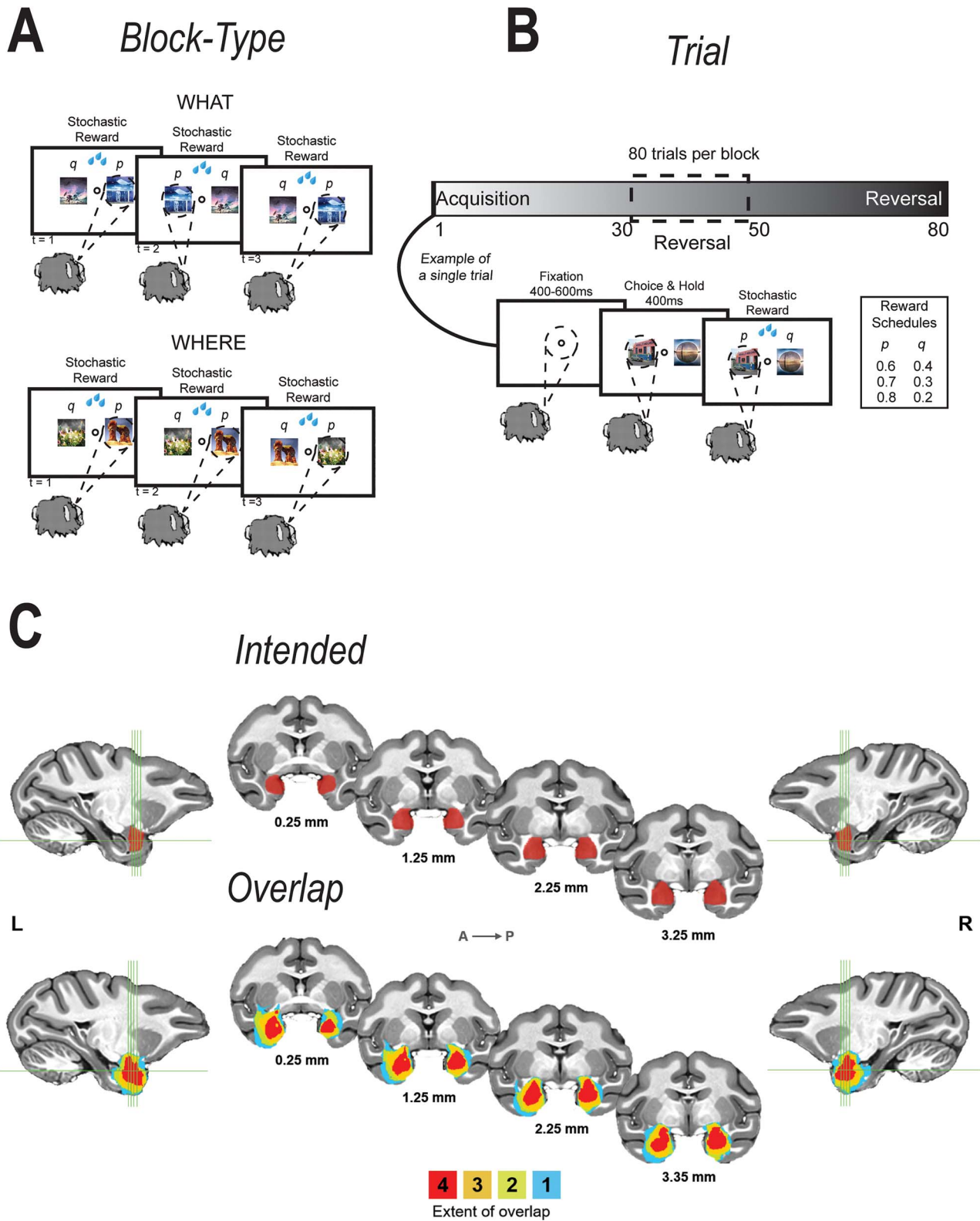


Figure 1. Task and lesion extent. (A, B) What and Where Task. The task was divided into 80-trial blocks. At the beginning of each 80-trial block, we introduced two new objects that the animal had never seen before. Each block of trials was either a What block or a Where block. If it was a What block, we assigned a high-reward probability to one object and a low-reward probability to the other. If it was a Where block, we assigned a high-reward probability to one of the locations and a low-reward probability to the other. We used three different reward schedules (80/20, 70/30, and 60/40). The reward schedules were randomly assigned to the block and remained fixed for the block. The block type was also randomized and remained fixed for the entire block. There was no cue to indicate block type. Therefore, the animals had to infer the block type. In addition, on a randomly chosen trial between 30 and 50, we reversed the choice-outcome mapping, such that the better choice became the worse choice and vice-versa. (C) Extent of lesion and number of animals with shown extent, overlaid on a standardized macaque brain template.

There were two different block types: What and Where. In “what” blocks, the higher probability option was one of the two objects independent of which side it was presented on. In “where” blocks, the higher probability option was one of the two saccade directions independent of the object that was selected. There was no cue to indicate block type; monkeys determined block type by making choices and getting feedback. As with the reward schedule, the block type remained constant throughout the entire 80-trial block. In each block, on a randomly selected trial between 30 and 50, inclusive, the reward mapping was reversed, making the previously lower probability option the higher probability option. The reversal trial was not cued; monkeys had to learn through trial and error that the reward mapping switched.

Task Training

All animals were trained on the task using the same procedure. Eight out of 10 monkeys (5 of the 6 controls and 3 of the 4 lesion monkeys) had a more extensive training history. These monkeys completed other tasks before beginning training for the current task. In the previous tasks, they learned only object-based reward associations. After the remaining two monkeys (1 control and 1 lesion) learned to make saccades to fixate on targets, they were trained on a simple two arm bandit RL task in which they learned only object-based reward associations.

Next, all monkeys were trained with a deterministic schedule (100/0) in both the What and Where conditions. Monkeys were first introduced to one block type, either What or Where, with block type randomly assigned and balanced across the group. Once monkeys could successfully perform 15–24 blocks per session, we introduced the other block type by itself, and then upon stabilized performance in that block type, we mixed the two block types. Once the monkeys reached stable performance in the deterministic setting, we gradually introduced probabilistic outcomes; probabilities were lowered until the final schedules of 80/20, 70/30, 60/40 were reached.

Objects and Eye Tracking

Objects provided as choice options were normalized for luminance and spatial frequency using the SHINE toolbox for MATLAB (Willenbockel et al. 2010). All objects were converted to grayscale and subjected to a 2D FFT to control spatial frequency. To obtain a goal amplitude spectrum, the amplitude at each spatial frequency was summed across the two object dimensions and then averaged across objects. Next, all objects were normalized to have this amplitude spectrum. Using luminance histogram matching, we normalized the luminance histogram of each color channel in each object so it matched the mean luminance histogram of the corresponding color channel, averaged across all objects. Spatial frequency normalization always preceded the luminance histogram matching. Each day before the monkeys began the task, we manually screened each object to verify its integrity. Any object that was unrecognizable after processing was replaced with an object that remained recognizable. Eye movements were monitored and the object presentation was controlled by PC computers running the Monkeylogic (version 1.1) toolbox for MATLAB (Asaad and Eskandar 2008) and Arrington Viewpoint eye-tracking system (Arrington Research).

Bayesian Model of Reversal Learning

We fit a Bayesian model to estimate probability distributions over several features of the animals’ behavior as well as ideal observer estimates over these features (Costa et al. 2015, 2016; Jang et al. 2015; Rothenhoefer et al. 2017). The Bayesian ideal observer model inverts the causal model for the task, so it is the optimal model. Using the ideal model, we estimated probability distributions over reversal points to estimate when a reversal occurred.

To estimate the Bayesian model, we fit a likelihood function given by:

$$f(x, y|r, p, h, b) = \prod_{k=1}^T q(k) \quad (1)$$

where r is the trial on which the reward mapping is reversed ($r \in 0-81$) and p is the probability of reward of the high reward option. The variable h encodes whether option 1 or option 2 is the high reward option at the start of the block ($h \in 1, 2$) and b encodes the block type ($b \in 1, 2$ —What or Where). The variable k indexes trial number in the block and T is the current trial. The variable k indexes over the trials up to the current trial so, for example, if $T = 10$, then $k = 1, 2, 3, \dots, 10$. The variable r ranges from 0 to 81 because we allow the model to assume that a reversal may not have happened within the block, and that the reversal occurred before the block started or after it ended. In either scenario where the model assumes the reversal occurs before or after the block, the posterior probability of reversal would be equally weighted for r equal to 0 or 81. The choice data are given in terms of x and y , where elements of x are the rewards ($x_i \in 0, 1$) and elements of y are the choices ($y_i \in 1, 2$) in trial, i . The variable p is varied from 0.51 to 0.99 in steps of 0.01. It can also be indexed over just the exact reward schedules (i.e., 0.8, 0.7, and 0.6), although this makes little difference as we marginalize over p for all analyses.

For the ideal observer model used to estimate the blocktype in the Bayesian analysis, we estimated the blocktype probability at the current trial T , based on the outcomes from the previous trials. Thus, the estimate is based on the information that the monkey had when it made its choice in the current trial. For each schedule, the following mappings from choices to outcomes gave us $q(k)$. For estimates of What (i.e., $b = 1$), targets 1 and 2 refer to the individual objects and saccade direction is ignored; whereas for Where (i.e., $b = 2$), targets 1 and 2 refer to the saccade direction and the object is ignored. For $k < r$ and $h = 1$, (when target 1 is the high-probability target and the trial is prior to the reversal) choose 1 and get rewarded $q(k) = p$, choose 1 and receive no reward $q(k) = 1 - p$, choose 2 and get rewarded $q(k) = 1 - p$, choose 2 and have no reward $q(k) = p$. For $k \geq r$, these probabilities are flipped. For $k < r$ and $h = 2$, the probabilities are complementary to the values where $k < r$ and $h = 1$. To estimate reversal, all values were filled in up to the current trial T .

For the animal’s choice behavior, used to estimate the posterior over b for each group, the model is similar, except the inference is only over the animal’s choices, and not whether it is rewarded. This model assumes that the animal had a stable choice preference, which switched at some point in the block from one object to the other. Given the choice preference, the animals chose the wrong object (i.e., the object inconsistent with their choice preference) at some lapse rate $1 - p$. Thus, for $k < r$ and $h = 1$ choosing option 1: $q(k) = p$, choosing option 2: $q(k) = 1 - p$. For $k \geq r$ and $h = 1$, choosing option 1: $q(k) = 1 - p$, choosing option 2: $q(k) = p$. Correspondingly for $k < r$ and $h = 2$, choosing option 2: $q(k) = p$, etc. The choice behavior model is

therefore similar to the ideal observer, except p indexes reward probability in the ideal observer model and $1 - p$ indexes the lapse rate in the behavioral model. The reward outcome also does not factor into the behavior model.

Using these mappings for $q(k)$, we then calculated the likelihood as a function of r , p , h , and b for each block of trials. The posterior is given by:

$$p(r, p, h, b | x, y) = f(x, y | r, p, h, b) p(r) p(p, h, b) / p(x, y) \quad (2)$$

For r , p , h , and b , the priors were flat. There is general agreement between the ideal observer estimate of the reversal point and the actual programmed reversal point (Costa et al. 2015, 2016).

With these priors, we calculated the posterior over the reversal trial by marginalizing over p , h , and b .

$$p(r | x, y, M) = \sum_{p, h, b} p(r, p, h, b | x, y) \quad (3)$$

The posterior over block type could correspondingly be calculated by marginalizing over r , p , and h .

Reinforcement Learning Model of Choice Behavior

We fit six different RL models that varied in the number of parameters used to model the data. In the results, we focus on the two models that most often accounted for the behavior. All models were based on a Rescorla-Wagner (RW), or stateless RL value update equation given by:

$$v_i(k+1) = v_i(k) + \alpha_f (R - v_i(k)) \quad (4)$$

We then passed these value estimates through a logistic function to generate choice probability estimates:

$$d_j(k) = \left(1 + e^{\beta(v_i(k) - v_j(k) + h_i(k) - h_j(k))} \right)^{-1}, d_i(k) = 1 - d_j(k) \quad (5)$$

The variable v_i is the value estimate for option i , R is the reward feedback for the current choice for trial k , and α_f is the learning rate parameter, where f indexes whether the current choice was rewarded ($R = 1$) or not ($R = 0$). For each trial, α_f is one of two-fitted values used to scale prediction errors based on the type of reward feedback for the current choice. Note that models M1, M2, and M3 described below do not have the h_j factors. The variable $h_j(k)$ implemented a choice autocorrelation function, which increased the value of a cue that had occurred in the same location, recently. This allows us to model a tendency to repeat a given choice, independent of whether it was rewarded. Because we wanted to use the same model across both What and Where, we implemented the choice autocorrelation functions as repetitions of choices when the same object occurs in the same location, which results in autocorrelations across four terms (i.e., stim 1 left, stim 1 right, stim 2 left, and stim 2 right). A model which used only choice repetition across location could not be fit to the Where condition, since the animals should “perseverate” on location. Thus, the use of object-location terms for the autocorrelation allows us to use the same model in both tasks.

The autocorrelation function was defined as follows:

$$h_i(k) = \kappa e^{-\lambda(k - k_{l(i)})} \quad (6)$$

where the variable κ and λ were free parameters scaling the size of the effect and the decay rate, respectively. The variable $k_{l(i)}$ indicates the last trial on which a given object was chosen in a given location, i . There were four separate values for $k_{l(i)}$ as it tracked two cues across locations. The values entered into equation (5) were the two (of the 4) that corresponded to the object/location pairs actually presented in the current trial. These parameters allowed us to characterize choice-perseveration across the interaction of object and action choices.

The likelihood was given by:

$$f(x, y | \beta, \alpha, \kappa, \lambda) = \prod_k [d_1(k)c_1(k) + d_2(k)c_2(k)] \quad (7)$$

where $c_1(k)$ had a value of 1 if option 1 was chosen on trial k and $c_2(k)$ had a value of 1 if option 2 was chosen. Conversely, $c_1(k)$ had a value of 0 if option 2 was chosen, and $c_2(k)$ had a value of 0 if option 1 was chosen for trial k . We used standard function optimization methods to maximize the likelihood of the data given the parameters. Note, not all parameters were present in all models.

Three of the models (M1, M2, and M3) had different numbers of learning rate and inverse temperature parameters. M1 had one inverse temperature and two learning rate parameters (indexed by the subscript f on α), one for positive feedback and one for negative feedback. M2 had one inverse temperature and one learning rate parameter. M3 had two inverse temperatures, one for the acquisition phase and one for the reversal phase, and four learning rates, two for the acquisition phase (one for positive feedback and one for negative feedback), and two for the reversal phase (one for positive feedback and one for negative feedback). The remaining three models are the plus versions of the models discussed above (M1+, M2+, and M3+). The plus models have the same number of parameters as the basic (i.e., M1, M2, and M3) model with the addition of two more parameters, one for the coefficient on the autocorrelation factor, κ , and one for the decay factor on the autocorrelation, λ . Models M2 and M2+ predicted behavior most often across groups, so to simplify presentation we show results for these two models only, and the plots for these models show only the number of times these were the best model.

ANOVA Models

To quantify the difference between choice behavior in each group, we first flipped the data following the reversal when we entered it into the ANOVA. Data are plotted unflipped. Next, we performed an arcsine transformation on the choice accuracy values from each session, as this transformation normalizes the data (Zar 1999). Data were then averaged across sessions within monkey. We then carried out an N -way ANOVA (ANOVAN). Monkey was included as a random factor. All other factors were fixed effects. For all reported ANOVAs, we always ran an omnibus model with all factors and interactions of all order. Nonreported interactions were not significant. The ANOVA on win-stay lose-switch, entropy, and reversal trial difference were done in the same way as above without the arcsine transformation. For the choice strategy model, we entered both win-stay and lose-switch

as dependent variables and included a factor in the model for choice-strategy (i.e., either win-stay or lose-switch). Effect size is reported using ω^2 (Olejnik and Algina 2000).

Results

We tested rhesus macaques on a two-armed bandit reversal learning task with three different stochastic reward schedules: 80%/20%, 70%/30%, and 60%/40%. In addition to the three different reward schedules, there were two different block types: What and Where. In “what” blocks, the higher probability option was one of the two objects independent of the chosen location. In “where” blocks, the higher probability option was one of the two saccade directions independent of the chosen object. There was no cue to indicate block type. Therefore, monkeys determined block type by making choices and getting feedback. In each block, on a randomly selected trial between 30 and 50 (inclusive), the reward mapping was reversed, making the previously lower probability option the higher probability option and vice versa. The reversal trial was not cued and therefore monkeys had to learn through trial and error that the reward mapping switched.

Choice Behavior

We began by analyzing the monkeys’ choice behavior. Because the reversal trial differed across blocks, we first aligned each block to the true reversal point and interpolated the trials in the acquisition and reversal phases so there were 40 “trials” in each. We then carried out ANOVAs on this data, where the dependent variable was the fraction of times the animals chose the best initial option (Fig. 2). We first carried out an ANOVA across both block types (What and Where). There was no average difference in performance across block type (block type; $F(1, 8) = 0.05$, $P = 0.83$, $\omega^2 = 0$). We did however, find differences in reward schedule (schedule; $F(2, 16) = 107$, $P < 0.001$, $\omega^2 = 0.171$) on choices. There were also differences in these factors by trial (Block type \times Trial; $F(78, 624) = 4.1$, $P < 0.001$, $\omega^2 = 0.007$ and Schedule \times Trial; $F(156, 1248) = 23.4$, $P < 0.001$, $\omega^2 = 0.026$), which reflects both the initial learning, and the reversal of choices after the reversal in reward mapping. We also found that control monkeys performed better than amygdala monkeys and this varied by trial (Group \times Trial; $F(79, 632) = 2.3$, $P < 0.001$, $\omega^2 = 0.014$), and by schedule (Group \times Schedule \times Trial; $F(156, 1248) = 1.3$, $P = 0.01$, $\omega^2 = 0.001$). The groups did not, however, differ by block type (Group \times Block Type \times Trial; $F(79, 632) = 0.6$, $P = 0.99$, $\omega^2 = 0.001$).

Although there were no group differences by block type, we carried out planned comparisons on the data from each condition. In What blocks, both groups chose more accurately in the richer reward schedules (Schedule \times Trial; $F(156, 1248) = 16.3$, $P < 0.001$, $\omega^2 = 0.027$). We also found that control monkeys performed better than the amygdala lesioned monkeys and this varied by trial (Group \times Trial; $F(78, 624) = 2.1$, $P < 0.001$, $\omega^2 = 0.013$), and by schedule (Group \times Schedule \times Trial; $F(156, 1248) = 1.4$, $P = 0.003$, $\omega^2 = 0.002$). Similarly, in Where blocks, both groups chose more accurately in the richer reward schedules (Schedule \times Trial; $F(156, 1248) = 8.9$, $P < 0.001$, $\omega^2 = 0.029$). Control monkeys performed better than the amygdala monkeys which varied by trial (Group \times Trial; $F(79, 632) = 1.6$, $P < 0.001$, $\omega^2 = 0.018$), but there was no difference across schedule (Group \times Schedule \times Trial; $F(156, 1248) = 1.1$, $P = 0.27$, $\omega^2 = 0.003$).

To further characterize the learning behavior, we analyzed the win-stay, lose-switch performance (Fig. 3). Win-stay is the probability that the animals chose the same option after a

positive outcome in the previous trial and lose-switch is the probability that they chose the other option after a negative outcome in the previous trial. For purposes of the ANOVA, we analyzed only win-stay and lose-switch probabilities. The difference between win-stay and lose-switch was coded as a choice-strategy effect. We found differences across block types (Block type; $F(1, 8) = 23.7$, $P = 0.001$, $\omega^2 = 0.006$). Consistent with the decreased overall accuracy of the lesioned animals, they also had lower win-stay strategies relative to higher lose-switch than controls (Group \times Choice Strategy; $F(1, 8) = 6.4$, $P = 0.036$, $\omega^2 = 0.035$). These group strategies did not differ by block type (Group \times Choice Strategy \times Block type; $F(1, 8) = 1.8$, $P = 0.217$, $\omega^2 = 0.004$). We then ran the analysis separately for win-stay and lose-switch strategies and found that there were no group differences for win-stay (Group; $F(1, 8) = 3.6$, $P = 0.095$, $\omega^2 = 0.101$). However, lesioned animals more often switched following a negative outcome (Group; $F(1, 8) = 11.1$, $P = 0.010$, $\omega^2 = 0.393$). Therefore, across block types, lesioned animals switched after a negative outcome more frequently than the control animals. The groups also differed by block type (Group \times Block type; $F(1, 8) = 16.3$, $P = 0.004$, $\omega^2 = 0.004$). When we ran the analysis separately for each group, controls did not differ by block type (Block type; $F(1, 5) = 0.6$, $P = 0.46$, $\omega^2 = 0$) but the lesioned animals did (Block type; $F(1, 3) = 21.7$, $P = 0.018$, $\omega^2 = 0.37$).

Next, we looked at the probability that while animals were in one block type they were making choices consistent with the other block type (Fig. 4). For What blocks, we quantified the probability of choosing the most frequently chosen location and for Where blocks we quantified the probability of choosing the most frequently chosen object. On average, monkeys should be choosing each action at chance levels in What blocks, and each object at chance levels in Where blocks. However, even if they infer the correct block type, at the beginning of the block, they may make choices consistent with the wrong block type for several trials, and this can persist into the block.

We started by analyzing both block types together and found that animals were closer to chance, and therefore were making choices more consistent with the appropriate block type in easier schedules (Schedule; $F(2, 16) = 2.6$, $P < 0.001$, $\omega^2 = 0.009$). In addition, we found that the groups differed across schedule and block type (Group \times Schedule \times Block type; $F(2, 16) = 11.2$, $P < 0.001$, $\omega^2 = 0.018$). The amygdala-lesioned animals were making relatively more location choices in What blocks than control animals, when compared with object choices in Where blocks and this differed by schedule. To examine this in more detail, we analyzed each block type separately. In What blocks (Fig. 4A), we found that animals performed better in easier schedules (Schedule; $F(2, 16) = 37.5$, $P < 0.001$, $\omega^2 = 0.35$). We also found that the groups differed across schedule (Group \times Schedule; $F(2, 16) = 8.2$, $P = 0.003$, $\omega^2 = 0.039$). In Where blocks (Fig. 4B), we again found that animals performed better in easier schedules (Schedule; $F(2, 16) = 19.6$, $P < 0.001$, $\omega^2 = 0.042$). There were, however, no group differences across schedules (Group \times Schedule; $F(2, 16) = 2.9$, $P = 0.081$, $\omega^2 = 0.006$).

Reinforcement Learning Model

To further investigate why the monkeys with amygdala lesions behave differently, we fit several RL models which varied in the number of free parameters used to model the choice behavior (see Methods). We used the Bayesian Information Criterion (BIC) to assess which model fit best in each session for each animal (Fig. 5). Across monkeys, the model which most frequently fit

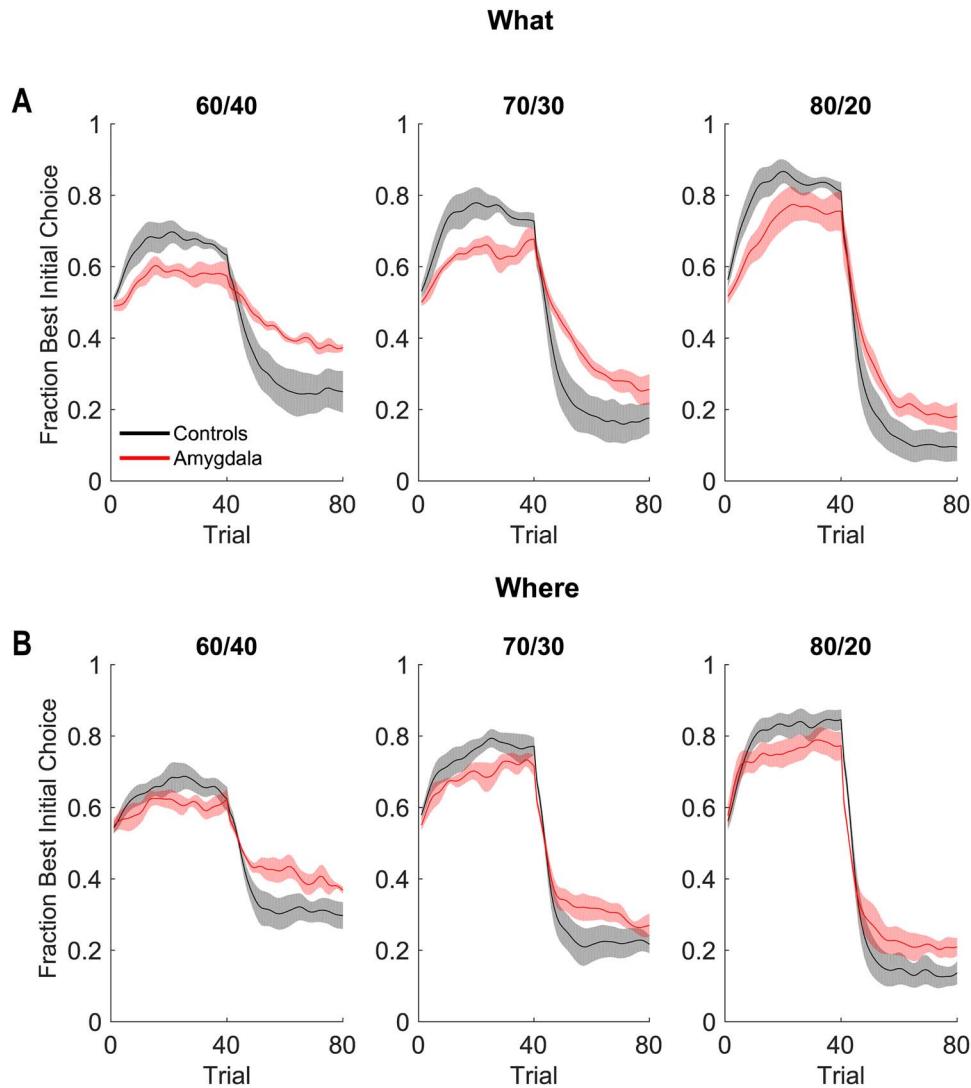


Figure 2. Behavioral performance in What and Where conditions. (A) Fraction of times the animals chose the best initial cue in the What condition. Shaded region indicates ± 1 standard error mean (SEM), where the N = the number of animals in each group (4 lesion, 6 control). (B) Same as A for the Where condition.

best had four parameters (M2+). The M2+ model had one learning rate, one inverse temperature, one autocorrelation parameter, and one decay parameter which decays the autocorrelation perseveration effects. The autocorrelation and decay parameters characterized the tendency to perseverate on choices, independent of whether they were appropriate to the current block. The model which fit second most frequently had two parameters (M2). The M2 model had one learning rate and one inverse temperature parameter but no perseveration parameters. The M2+ model captures perseverative choice biases driven by choices consistent with the opposite block type. Therefore, a preference for the plus model suggests the monkeys choices are not driven by the choice-outcome effects for the current blocktype, to the same extent.

The relative preference for the M2+ model was larger in amygdala animals than controls in the What condition than the Where condition (Fig. 5; Group \times Block type \times Model; $F(1, 8) = 6$, $P = 0.040$, $\omega^2 = 0.083$). Next, we split the analysis by block type. In the What condition, there was a preference for the M2+ model

(Model; $F(1, 8) = 21.5$, $P = 0.001$), but there was no preference in the Where condition (Model; $F(1, 8) = 0.7$, $P = 0.434$). However, there were no group effects or interactions with group in either the What or the Where conditions ($P > 0.05$). Therefore, there was a shift towards the M2 model, relative to the M2+ model, in the amygdala animals in the Where conditions. Since the plus model captures a tendency to repeat a response to an object at a specific location, independent of reward, this suggests that there is a shift towards a less object dependent strategy in the Where condition, relative to a location dependent strategy in the What condition, in the amygdala animals.

Next, we examined the parameters for the M2+ model. The only parameter that the groups differed on was the autocorrelation coefficient, κ (Fig. 6), which was larger in controls across both conditions (Group; $F(1, 8) = 15.6$, $P = 0.004$, $\omega^2 = 0.582$). The autocorrelation factor captures the tendency to repeat choices of objects at specific locations, and therefore captures perseveration across actions and objects (Lau and Glimcher 2005; Gershman et al. 2009).

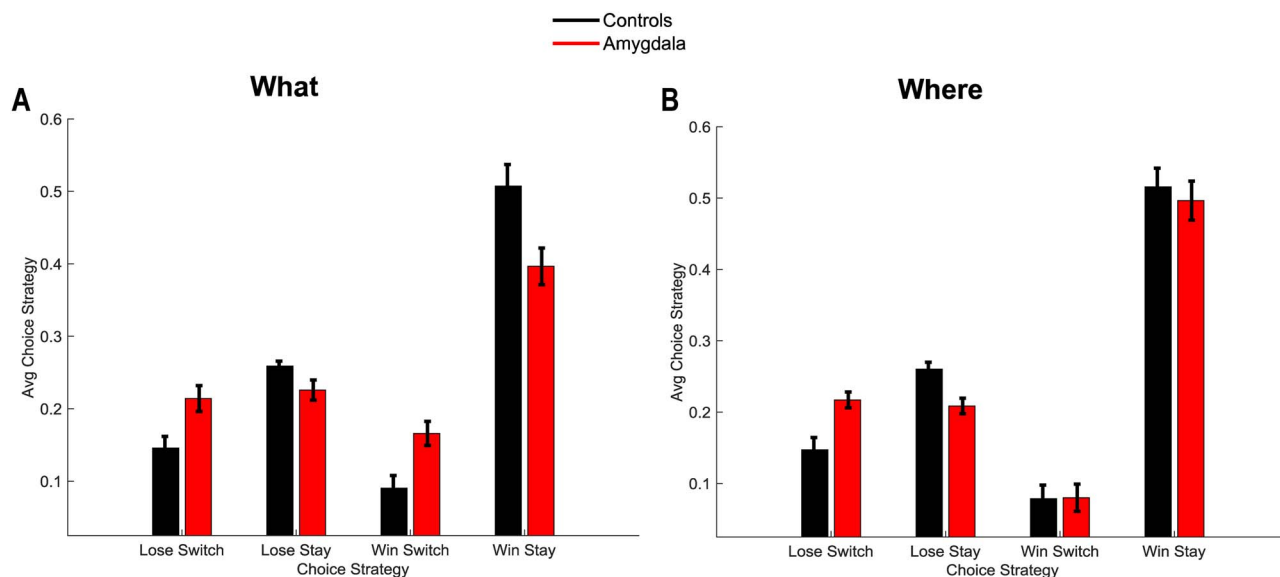


Figure 3. Win-stay, Lose-switch. (A) Win-stay, lose switch performance for the two groups in the What condition, averaged across schedules. (B) Win-stay, lose-switch performance for the two groups in the Where condition, averaged across schedules. Error bars are ± 1 SEM ($N = 6$ control, 4 lesion).

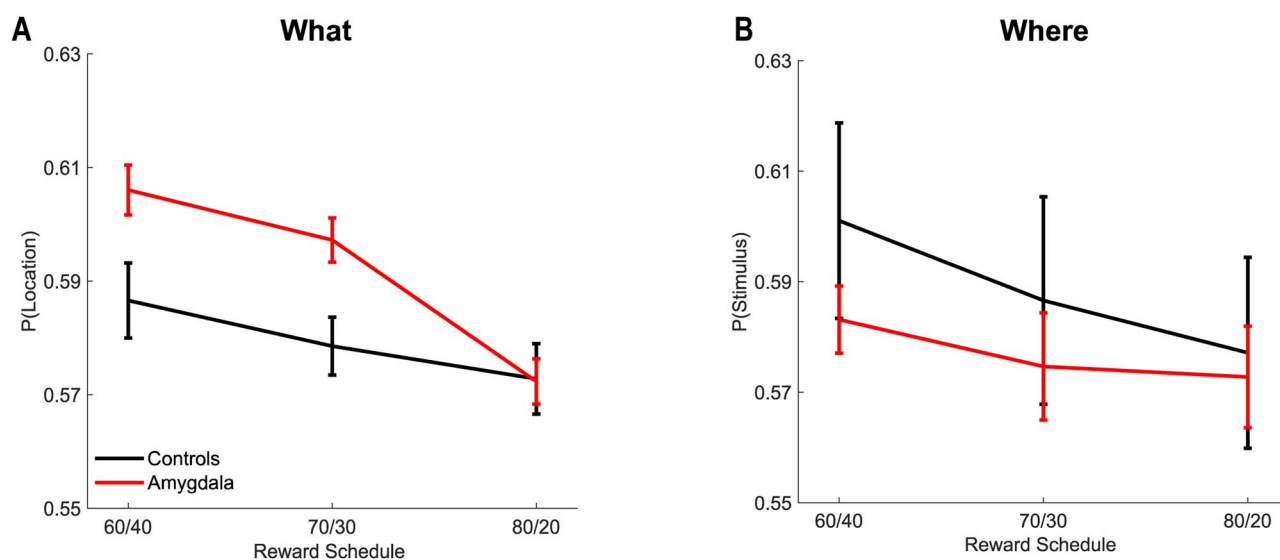


Figure 4. Cross condition choice frequencies. (A) Probability of choosing the most frequently chosen location in the What condition, averaged across reversals (with reversal data flipped). (B) Probability of choosing the most frequently chosen object in the Where condition. Error bars are ± 1 SEM, where the N = the number of animals in each group (4 lesion, 6 control).

Reversals

Learning in this task is governed by three processes, which may or may not map onto different neural systems. Monkeys have to infer the block type, they have to infer the correct option within each block type, and they have to reverse this preference when the outcome mapping reverses. The animals have extensive experience on the task before we collect behavioral data and, at least control animals, learn that reversals happen in the middle of the block (Costa et al. 2015). The monkeys use the acquired task knowledge to improve performance on the task. The results above show that monkeys with amygdala lesions have deficits in both the What and Where conditions. However, it is not clear

whether animals with amygdala lesions have general deficits in forming associations between actions or objects and rewards, or whether they have deficits in reversing these preferences. Therefore, we next addressed the reversal performance directly.

We used a Bayesian model to analyze the reversal behavior. The model assumes that the animals develop an initial preference for one option, and then reverse this preference at some point in the block. Because behavior is stochastic, the animals do not pick one option exclusively, and then switch at some point in the middle of the block to picking the other option. However, they tend to pick one of the options more often and this tendency switches in the middle of the block (Fig. 2). The model generates the probability that the animal reversed its choice

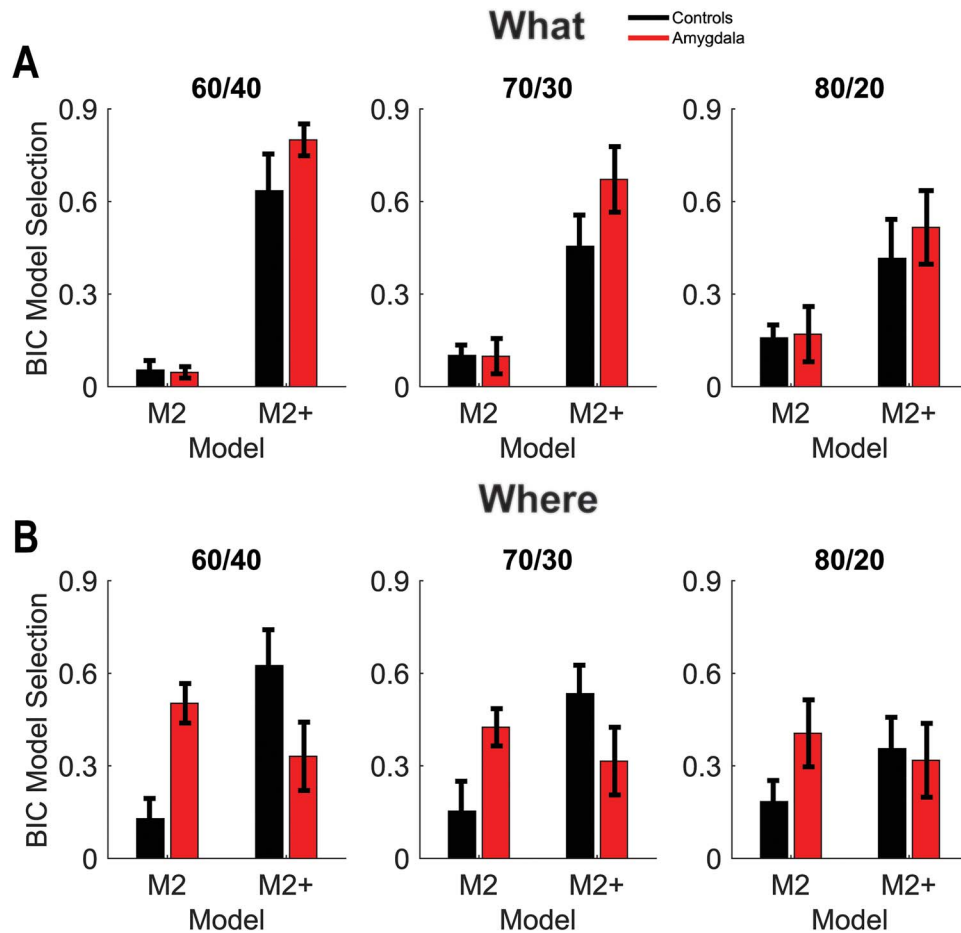


Figure 5. BIC model selection. (A) Percentage of sessions BIC selected models M2 and M2+ (out of all 6 models) in the What condition. Error bars are ± 1 SEM, where the N = the number of animals in each group (4 lesion, 6 control). (B) Same as A for the Where condition.

behavior on each trial of the block—a probability distribution over reversal trial ($p(r)$, Fig. 7). On average, these probability distributions were better centered around the actual reversal points for the easier than harder reward schedules (Fig. 7A,B,D,E). We tested this by taking the value of the probability distribution on the average expected reversal trial (40). When we compared these values across both block types, we found that animals had higher values in easier schedules (Schedule; $F(2, 16) = 11.4$, $P < 0.001$, $\omega^2 = 0.127$).

To characterize the distributions and examine group differences, we calculated the entropy (i.e., $\hat{h} = -\sum_{r=0.81} p(r) \log p(r)$) of the posterior distribution over reversals in each block (Fig. 7C,F). The entropy generalizes the concept of variance to non-Gaussian distributions. It is a measure of how concentrated the distribution is around the mean or mode. Higher entropy indicates broader reversal distributions and therefore noisier, less precise reversals. When we compared the entropy across both block types, we found an overall effect of schedule on entropy (Schedule; $F(2, 16) = 47.5$, $P < 0.001$, $\omega^2 = 0.107$). Therefore, the switch in choice preference was more clearly defined for the easy than hard schedules. We also found that the entropy for control monkeys was significantly lower than for the monkeys with amygdala lesions (Group; $F(1, 14) = 9.4$, $P = 0.015$, $\omega^2 = 0.197$). Next, we analyzed the What and Where blocks separately. In What blocks, we found an overall effect of schedule

(Schedule; $F(2, 16) = 25.2$, $P < 0.001$, $\omega^2 = 0.07$) and a group effect (Group; $F(1, 8) = 9$, $P = 0.017$, $\omega^2 = 0.252$). In Where blocks, we also found effects of schedule (Schedule; $F(2, 16) = 31.1$, $P < 0.001$, $\omega^2 = 0.157$) and group (Group; $F(1, 8) = 6.9$, $P = 0.030$, $\omega^2 = 0.147$).

Two distributions can have different entropy but the same mean. Therefore, we next examined whether the estimated reversal trial differed between lesion and control groups, to see if the groups tended to reverse on the same trial. To do this, we calculated the expected value (i.e., the mean) of the reversal distribution in each block (i.e., $\hat{r} = \sum_{r=0.81} r p(r)$). This gives us a single number for each block, estimating the trial on which the animal reversed its choice preference. This number can be compared with where the actual reversal occurred, which we refer to as the reversal trial difference (i.e., $\hat{r} - r_{\text{actual}}$; Fig. 8A,B). The variable \hat{r} characterizes our estimate of where the monkey reversed and r_{actual} is the programmed reversal trial. When we analyzed both block types together, we found no effect of block type on the difference between the reversal trial of the animals and the actual reversal trial (Block Type; $F(1, 8) = 0.01$, $P = 0.92$, $\omega^2 = 0$). However, there was an effect of schedule (Schedule; $F(2, 16) = 17.1$, $P < 0.001$, $\omega^2 = 0.24$), with animals reversing before the actual reversal trial in the harder conditions, consistent with previous work (Costa et al. 2015; Rothenhoefer et al. 2017). There were no group differences (Group; $F(1, 8) = 1.2$, $P = 0.3$, $\omega^2 = 0.033$). When we analyzed the where block by itself, we found that the

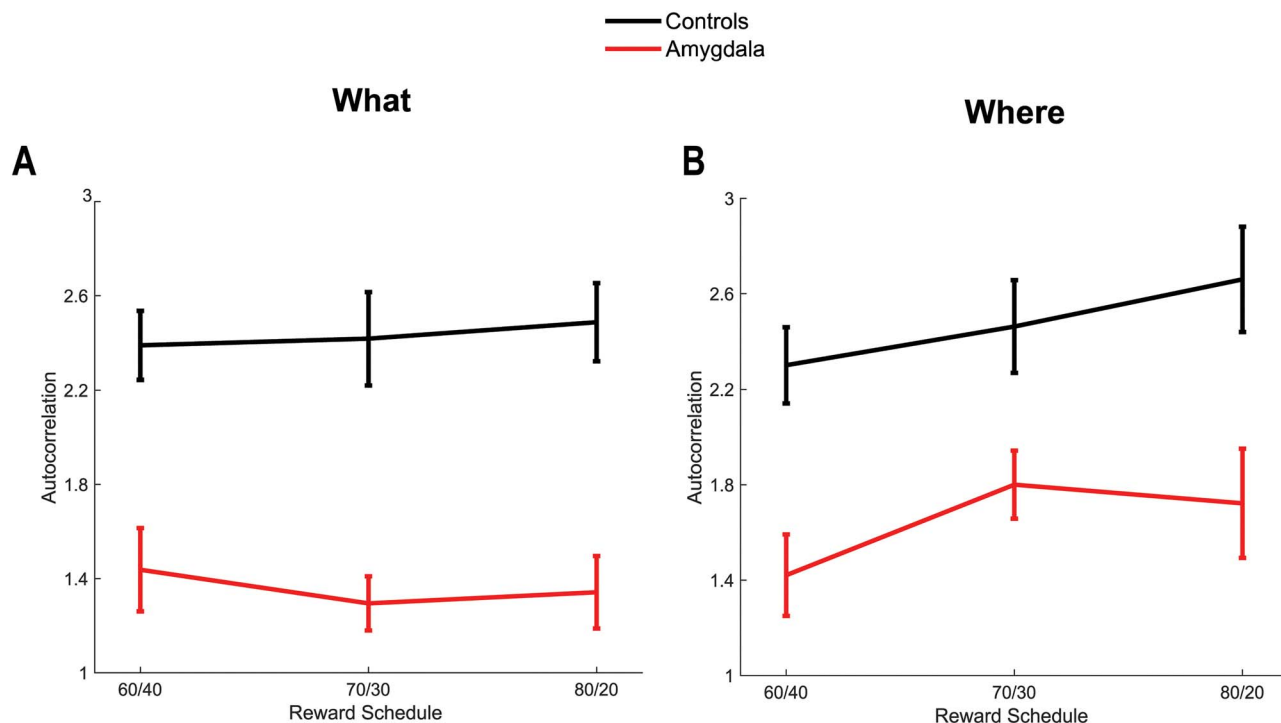


Figure 6. Autocorrelation parameter for model M2+. (A) Autocorrelation parameter in the What condition. Error bars are ± 1 SEM, where the N = the number of animals in each group (4 lesion, 6 control). (B) Same as A for the Where condition.

groups differed in reversal behavior across schedule, reflecting the difference in the 60/40 condition (Group \times Schedule; $F(2, 16) = 5.2, P = 0.018, \omega^2 = 0.121$).

Following this, we looked at the absolute value of the difference between the monkey reversal and the actual reversal (i.e., $|\hat{r} - r_{\text{actual}}|$; Fig. 8A,B). Unlike the difference in reversal trials (Fig. 8A,B), the absolute value of the difference (Fig. 8C,D) characterizes how close the animals were to the actual reversal, either before or after. As above we found an effect of schedule (Schedule; $F(2, 16) = 64.0, P < 0.001, \omega^2 = 0.35$). There were, however, no differences across block type (Block Type; $F(1, 8) = 0.03, P = 0.87, \omega^2 = 0$) and there were no differences between groups (Group; $F(1, 8) = 0.15, P = 0.71, \omega^2 = 0.01$). Overall, therefore, despite their generally noisier behavior, the monkeys with amygdala lesions tended to reverse on the same trial as the controls, and they were as close to the actual reversal, in absolute value.

The increased entropy of the reversal distribution may be driven by noisy choice behavior. The algorithm assumes that any choice not consistent with the dominant choice in a phase is possibly a reversal. Therefore, noisy choices broaden the reversal distribution. To characterize this in more detail, we first calculated the average fraction of correct choices for each animal relative to the currently most rewarded object, across the block. We then correlated the average fraction of correct choices with the average entropy of the reversal distribution (Fig. 9A), the average reversal trial difference (Fig. 9C) and the average absolute value of the reversal trial difference (Fig. 9B). The correlation between fraction correct and the entropy was large and negative across animals ($\rho = -0.920, P < 0.001$), as would be expected as entropy depends on choice accuracy. The correlation was also significant with the absolute value of the difference between the monkey and actual reversal trial ($\rho = -0.773,$

$P < 0.01$). However, the correlation between the fraction correct and the signed difference in the reversal trial was not significant ($\rho = 0.117, P > 0.05$). The correlation between fraction correct and entropy was significantly larger than the correlation between fraction correct and the reversal trial difference ($Z = 3.19, P = 0.001$), but the difference between the correlation of the fraction correct and the entropy, and the correlation between the fraction correct and the absolute value, was not different ($Z = 1.05, P = 0.300$). It is not surprising that these correlations do not differ, because the absolute value is related to the entropy. The entropy characterizes the width of the posterior over reversal trials, and the absolute value characterizes how far samples from this distribution are, from the mean, on average.

The Bayesian model also estimates whether the monkey's choices were more consistent with choosing one of the objects (What block) or one of the saccade directions (Where blocks). These estimates provide evidence for the block type the monkeys thought they were in based on their choice strategy (Fig. 10). Across conditions there was a fourth-order interaction (Group \times Trial \times Schedule \times Block type; $F(158, 1264) = 1.3, P = 0.007, \omega^2 = 0.001$). To examine this in detail, we analyzed each block type separately. In What blocks, we found that posteriors were higher for easier schedules (Schedule; $F(2, 16) = 42.7, P < 0.001, \omega^2 = 0.12$). We also found that group differences varied across schedules and trials (Group \times Trial \times Schedule; $F(158, 1264) = 2.4, P < 0.001, \omega^2 = 0.004$). When we analyzed effects in the What blocks separately for each schedule, we found the groups differed across trials in all schedules (Group \times Trial; 60/40 $F(79, 632) = 3.5, P < 0.001, \omega^2 = 0.028$); 70/30 $F(79, 632) = 6.6, P < 0.001, \omega^2 = 0.036$); 80/20 $F(79, 632) = 1.31, P = 0.044, \omega^2 = 0.009$). In Where blocks (Fig. 10B), we found that posteriors were

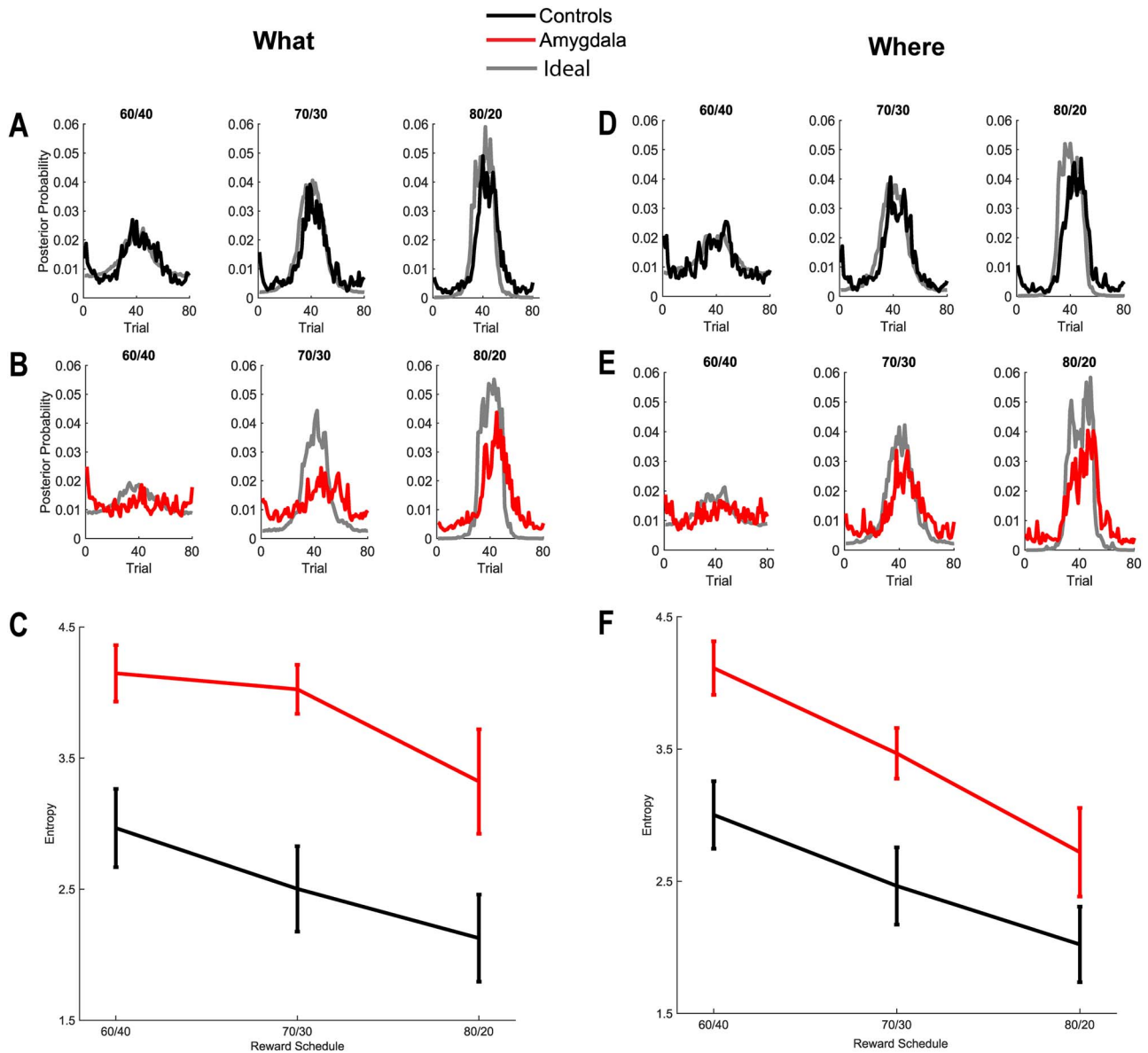


Figure 7. Posterior distributions and entropy of posterior. (A) Posterior distribution for control group in the What condition, overlaid on ideal observer posterior for each schedule of the What condition. (B) Same as A for the lesion group. (C) Entropy of posterior distribution for both groups for the What condition, broken out by schedule. Error bars are ± 1 SEM ($N = 6$ control, 4 lesion). (D) Same as A for the Where condition. (E) Same as B for the Where condition. (F) Same as C for the Where condition.

higher in easy schedules, reflecting increased consistency in the monkey's choice behavior (Schedule; $F(2, 16) = 29.9$, $P < 0.001$, $\omega^2 = 0.117$). There were, however, no group differences. Therefore, in What blocks, the monkeys with amygdala lesions were less consistently choosing one of the objects relative to the controls. However, in Where blocks the groups did not differ.

Discussion

In the present study, we found that lesions of the amygdala affected learning to select rewarding stimuli (What) and rewarding actions (Where). In both block types, we found that controls more often chose the better option than the monkeys with amygdala lesions. The choice accuracy deficit in the animals with amygdala lesions was not significantly different in one

block type versus the other. When we analyzed win-stay, lose-switch measures of the monkey's choices, we found that the lesioned animals more often switched after a negative outcome, which decreased performance due to the stochastic schedules. Therefore, much of their decreased accuracy, overall, followed from switching after negative outcomes, and these effects were significant in the What condition, although we did not find that the groups differed significantly across conditions. We also found that the animals with amygdala lesions tended to consistently select locations more often in What blocks, relative to control animals in harder schedules. This was consistent with the finding that animals with amygdala lesions were better fit by a model with object by location perseveration in the What condition than the Where condition. Because the perseveration term in the RL model is independent of reward, it will tend

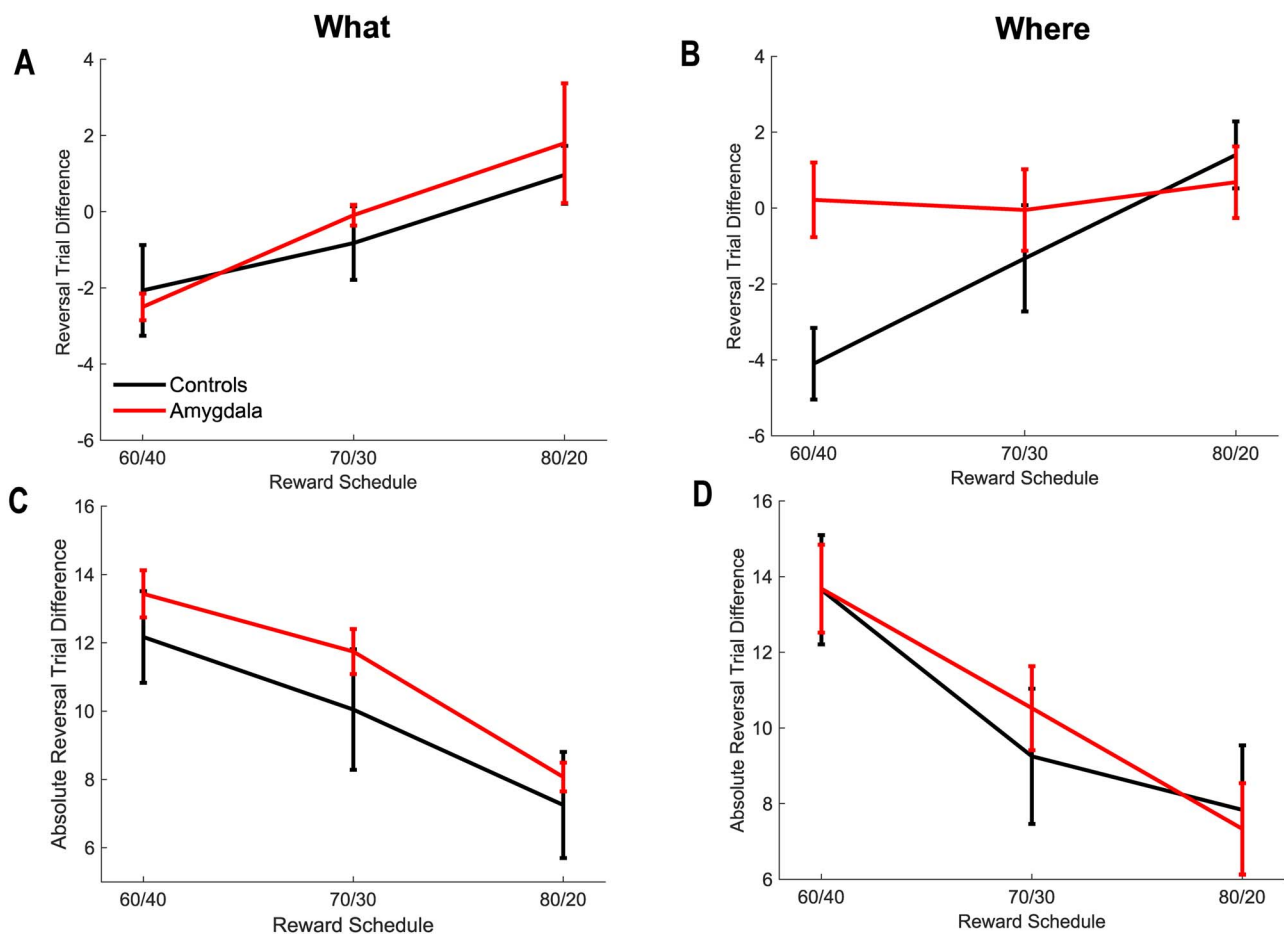


Figure 8. Relative and absolute difference in reversal behavior. Error bars are ± 1 SEM ($N = 6$ control animals and 4 lesion animals). (A) Relative reversal trial in the What condition. The relative reversal trial is given by the difference between the point estimate of the monkey's reversal trial in each block and the actual reversal trial. Negative numbers indicate that the monkey reversed before the actual reversal trial. (B) Relative reversal trial in the Where condition. (C) Absolute value of the difference in the reversal trial in the What condition. In each block, we computed the difference between the estimated reversal trial of the animal, and the actual reversal trial. We then took the absolute value of this difference.

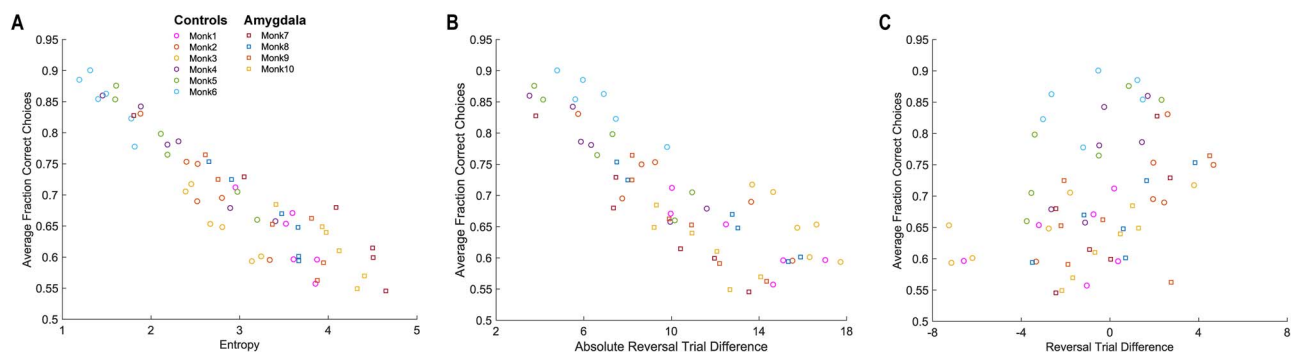


Figure 9. Correlation plots for all 10 monkeys (6 control, 4 lesion). (A) The correlation between fraction correct and entropy. (B) Same as A, but the correlation is between fraction correct and absolute reversal trial difference. (C) Same as A, but the correlation is between fraction correct and reversal trial difference.

to lower performance and therefore it lowered performance relatively more in the What condition than the Where condition.

Because the operated monkeys sustained a variable amount of damage to structures adjacent to the amygdala, in addition to the substantial, planned damage to the amygdala, we considered the possibility that the behavioral impairments arose from the inadvertent, extra-amygdala damage. Notably, there was

no apparent correlation of behavioral scores with the amount of inadvertent damage to a particular structure. For example, cases M2 and M3 had similar scores on acquisition yet ranked fourth and first among lesion subjects in extent of damage to the entorhinal cortex, respectively. In addition, based on prior work, we can be confident that the amount of damage estimated from T2-weighted scans, reported in [Supplementary](#)

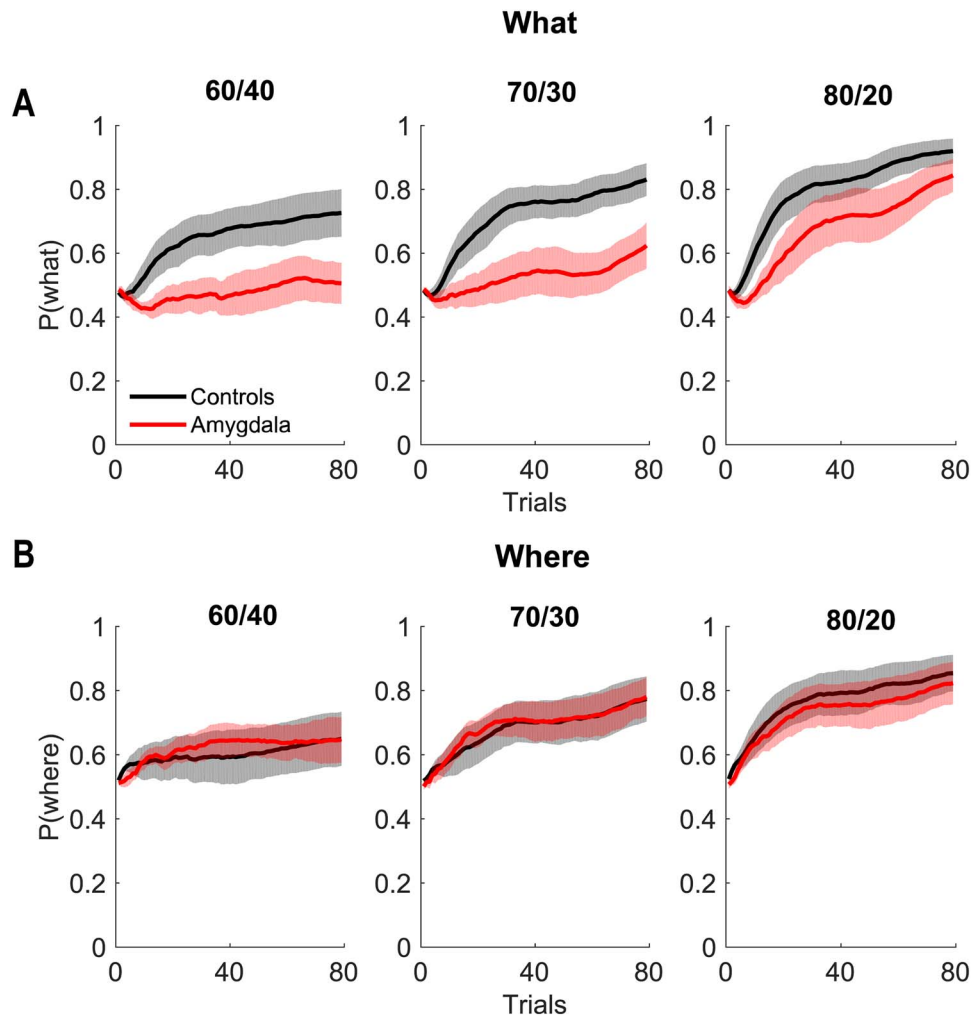


Figure 10. Posterior probability of the choice strategy used by the monkeys. (A) Probability that the monkeys were using a What strategy in What blocks. A What strategy implies that the monkeys are consistently picking one of the objects. (B) Probability that the animals were using a Where strategy in Where blocks. A Where strategy implies that the monkeys are consistently picking a location. Shaded region indicates ± 1 SEM, where the N = the number of animals in each group (4 lesion, 6 control).

Table 1, is an overestimate (Basile et al. 2017). These two factors militate against the possibility that extra-amygdala damage is responsible for the behavioral impairments we observed.

We also examined the reversal behavior in detail, to see whether the animals with amygdala lesions had specific deficits in reversing their choice-outcome preferences. For both block types, we found that the lesion group had higher entropy in their reversal distributions. This would be expected if the animals less consistently chose the better option in both the acquisition and reversal phase, as each time the animals choose the less preferred option, there is a small probability that they are reversing their choice preference. This is, therefore, consistent with the increased lose-switch probability of the lesioned group. When we compared the mean and absolute values of the estimated reversal trials, we found no average differences between the groups. Therefore, the monkeys with amygdala lesions reversed, on average, as effectively as the controls. In other work, we have found a correlate of the reversal inference in dorsal lateral prefrontal cortex (Bartolo and Averbeck 2020), which suggests it may be playing an important role in the reversal process, although we have not yet looked for such a correlate in the

amygdala. As stated earlier, learning in this task requires three processes. Monkeys have to infer the block type, they have to figure out the best choice within each block type, and they have to reverse this preference when the outcome mapping reverses. (It is possible that inferring the block type and figuring out the best option are done as one process.) Of these three processes, only the ability to consistently pick the best option was significantly impaired in monkeys with amygdala lesions, and this was primarily driven by more frequently switching after a negative outcome. The fact that our results are not statistically distinguishable across the different block types suggests that the amygdala plays a general role in forming associations between both objects and actions with rewards. Whether this is a deficit in representing the choices, the rewards, or in forming associations between them is not clear from the current results. Previous work would suggest the deficit, at least in part, is in forming the association (Cardinal et al. 2002; Chau et al. 2015; Costa et al. 2016). These data are also consistent with previous work, examining learning to reverse, which showed that monkeys with amygdala lesions learned to reverse faster (Jang et al. 2015). The current results suggest that the faster

reversals in those studies followed from weaker object outcome associations, not stronger prior probabilities on variability in the environment (Jang et al. 2015). Also relevant is the finding discerned from fMRI that, in intact monkeys, amygdala activity during both deterministic and probabilistic learning specifically predicts lose-shift behavior, and adaptive win-stay, lose-shift signals are evident in ventrolateral prefrontal cortex area 120 (Chau et al. 2015), a region necessary for probabilistic discrimination learning (Rudebeck, Ripple, et al. 2017). Future studies could address whether probabilistic learning like that examined here requires the functional interaction of the amygdala with the ventrolateral prefrontal cortex.

Learning systems and their anatomical substrates can be dissociated in various ways. For example, learning is often studied using Pavlovian or instrumental paradigms (Mackintosh 1994). Formation of Pavlovian CS-US associations is mediated, to some extent, by the amygdala (Cardinal et al. 2002; Braesicke et al. 2005). Formation of instrumental associations, on the other hand, is thought to be mediated by frontal-striatal systems (Balleine et al. 2009). Both forms of conditioning were developed from purely behavioral considerations and, therefore, they do not necessarily map cleanly onto separable neural systems. Furthermore, there is considerable interaction between these behavioral processes in tasks like Pavlovian Instrumental Transfer (Corbit and Balleine 2005) and conditioned reinforcement (Burns et al. 1993). The bandit tasks often used to study RL (O'Doherty et al. 2004; Daw et al. 2006; Costa et al. 2016), do not map cleanly onto Pavlovian or instrumental constructs. Actions are required to select options in bandit tasks. However, when the reward values of objects are being learned, the required action varies depending on the location of the object. Furthermore, when the reward values of actions are being learned, it is likely that learning the reward values of arm movements may engage different neural systems than learning the reward values of eye movements, given the differing neural systems engaged by each type of action (Alexander et al. 1986). It is also possible that learning deficits following amygdala lesions may depend on the type of motor response required to register choices. Additional work will be required to clarify this hypothesis.

From a psychological perspective, it is of interest that the amygdala is essential for both object-outcome and action-outcome associations as assessed with devaluations tasks (Málková et al. 1997; Rhodes and Murray 2013). Together with the present data, these findings show that the amygdala is important for learning about both objects and actions as they relate to reward probability (present study) and current reward value, including reward magnitude (Málková et al. 1997; Rhodes and Murray 2013; Rudebeck et al. 2013). An earlier study on object reversal learning found that, relative to unoperated controls, monkeys with amygdala lesions benefitted more from correct choices that follow an error in a deterministic setting (Rudebeck and Murray 2008). While this finding would seem to be at odds with the present findings, we note that the object reversal tasks differ in more ways than use of deterministic versus probabilistic outcomes. For example, the standard object reversal learning task employed by Rudebeck and Murray employed a small number (nine) of reversals, whereas in the present study, all monkeys had received extensive training in reversals. As a result, unlike in the present study, monkeys in the deterministic task experienced unexpected uncertainty, at least in early reversals. In addition, the present and earlier task differ in the type of response required (manual vs. eye-movement), in the location and type of reward (food reward

under object vs. fluid reward delivered to mouth), and in the number of trials administered per session (30 trials vs. massed trials). These task differences might account for the somewhat different picture gained from assessing amygdala contributions to the two kinds of reversal learning. Thus, the amygdala makes an essential contribution to reversal learning in probabilistic and even deterministic settings with massed trials in an automated apparatus (Costa et al. 2016) but not to reversal learning in deterministic settings with a small number of trials in a manual test apparatus (Izquierdo and Murray 2007).

The What versus Where task used in the current study was developed to separate neural circuits underlying learning rewards associated with objects whose locations vary, versus learning to associate rewards with actions independent of the object at the saccade target location. The hypothesis that such a dissociation should be possible follows from work in the visual and auditory systems (Ungerleider and Mishkin 1982; Romanski et al. 1999) based on the separable anatomical organization of visual cortex (Ungerleider and Mishkin 1982; Goodale and Milner 1992), as well as proposed frontal extensions of this circuitry (Averbeck and Murray 2020). Parietal cortex processes information about the spatial locations of objects in the environment, and the motor actions required to interact with these objects (Steinmetz et al. 1987; Mascaró et al. 2003; Caminiti et al. 2017). The ventral visual cortex, on the other hand, processes information about object features that allow for object identification and discrimination (Desimone et al. 1984; Yamins et al. 2014). This separable organization continues into prefrontal cortex (Barbas 1988; Webster et al. 1994; Averbeck and Seo 2008; Averbeck and Murray 2020), and correspondingly into the striatal circuitry (Haber et al. 2006; Averbeck et al. 2014; Neftci and Averbeck 2019). While there is evidence for anatomical segregation, neurophysiological recordings have shown integration of What and Where information in both prefrontal (Rao et al. 1997) and parietal (Serenó and Maunsell 1998) cortex. Thus, in neural circuits that are less proximal to sensory processes, it is not clear that these separate streams differentially process behaviorally relevant information, particularly in frontal-striatal systems.

The task was developed to separate learning about actions versus learning about objects. However, there are other differences between the conditions that may lead to behavioral effects. For example, in the What condition, the preferred object is present on the screen, whereas in the Where condition, although the response zones are indicated on the screen, the animals have to internally generate the action that will most likely lead to reward. Still, there is no reason to think monkeys with amygdala lesions are impaired in the ability to internally generate actions. An earlier study that examined the effects of amygdala lesions on conditional motor learning found no impairment in learning new conditional problems in a deterministic setting, even though the responses were internally generated (Murray and Wise 1996). It is also possible that the animals find one or the other condition to be more difficult. We have not systematically studied this, but in other work we have seen that some animals do better in either the Where relative to the What (Rothenhoefer et al. 2017) or the What relative to Where conditions (Bartolo and Averbeck 2020; Bartolo et al. 2020). Therefore, animals do not consistently show a clear preference for one or the other condition.

We have previously shown that learning oculomotor action sequences depends on a dorsal-lateral prefrontal, dorsal striatal circuit (Seo et al. 2012; Lee et al. 2015). The prefrontal and striatal

nodes in this circuit processed sequence related information (Seo et al. 2012), and local injections of dopamine antagonists into the DS led to deficits in performance during sequence learning (Lee et al. 2015). The DS also contains a stronger representation of action value than prefrontal cortex (Samejima et al. 2005; Lau and Glimcher 2008). This suggests that the dorsal circuit is important for action learning, when actions are eye movements. We have also found that lesions of the VS yield deficits specific to learning to select rewarding objects, without affecting learning to select rewarding actions, using the same What versus Where task used here (Rothenhoefer et al. 2017). In other work, we found that amygdala lesions affect learning to choose rewarding stimuli (Costa et al. 2016; Rudebeck, Ripple, et al. 2017). We have not, however, carried out a double dissociation experiment, using the same task with manipulations of either the DS or the VS.

Anatomy is often a guide to function. Anatomically, the basolateral amygdala is strongly interconnected with the ventral, visual object system (Turner et al. 1980; Amaral et al. 1992; Neftci and Averbeck 2019). It receives substantial projections from high level visual cortex (Amaral et al. 1992), and correspondingly projects to the VS (Friedman et al. 2002), and ventrolateral and orbital prefrontal cortex (Ghashghaei and Barbas 2002; Ghashghaei et al. 2007). Both prefrontal areas also receive input from temporal lobe, and not parietal lobe, visual areas (Webster et al. 1994). The amygdala also interacts with the medial portion of the mediodorsal thalamic nucleus (Russchen et al. 1987), which also projects to orbitofrontal cortex (Goldman-Rakic and Porrino 1985).

Given the anatomical connections of the amygdala with the ventral visual pathway, we had hypothesized that it would be mostly related to learning to choose objects and not actions. However, we found that lesions of the amygdala led to deficits in both learning to select actions and objects. Numerically the choice accuracy effect was larger for selection of objects. Although the amygdala has minimal connectivity with dorsal prefrontal areas that underlie oculomotor control (Ghashghaei et al. 2007) and no connections to the DS, it does have substantial back-projections across the visual hierarchy, including early visual cortex (Amaral et al. 1992). Given that early visual areas have minimal bilateral visual representation, these back-projections may affect lateralized spatial representations. Furthermore, neurophysiology studies have shown that amygdala neurons contain representations of the spatial locations of rewarded objects (Peck et al. 2013; Peck et al. 2014; Peck and Salzman 2014a, 2014b; Costa et al. 2019). These spatially selective responses are also in contrast to the VS, which contains no spatial information (Costa et al. 2019). Therefore, neurophysiology suggests a possible role for the amygdala in spatial-attentional processes, and the effects of these representations on behavior may be mediated by back-projections to early visual areas, or other pathways that connect these representations, polysynaptically, to areas that underlie eye movements. In related work, the amygdala is also involved in revaluing arm-motor responses in devaluation paradigms, where the value associated with a specific motor response changes following a selective satiation procedure (Rhodes and Murray 2013).

Amygdala interactions with orbitofrontal and ventrolateral prefrontal cortex are also likely important for the learning processes we have examined (Rudebeck, Saunders, et al. 2017; Murray and Rudebeck 2018). Recent work in rats has shown that ablation of amygdala neurons that project to the OFC impairs

reversal performance on a probabilistic spatial learning task (Groman et al. 2019). This deficit was due to the rats losing their ability to use positive outcomes to guide their choice behavior. In this same study, it was shown that ablation of OFC neurons projecting to amygdala enhanced reversal performance by destabilizing action values. Related to this, it has been shown that lesions of the OFC impair reversal behavior, but subsequent lesions of the amygdala in the same animals restore performance (Stalnaker et al. 2007). These results suggest that, under some circumstances, amygdala–OFC interactions may be detrimental to learning.

Our Bayesian model assumes that a state inference process underlies reversal learning. The model assumes the animals have a preference for one option (i.e., state A), which reverses at some point in the block (i.e., state B). The posterior distribution over reversals is the estimate of where the animals switch states in each block. Our data suggest that the amygdala is not involved in inferring state or state switches in our task. Although monkeys with amygdala lesions have deficits in consistently choosing the better option in both conditions, they reverse their choice preference as well as controls. Therefore, it is possible that state inference processes are being carried out by prefrontal cortical areas, as suggested by previous work (Durstewitz et al. 2010; Karlsson et al. 2012; Wilson et al. 2014; Schuck et al. 2016; Ebitz et al. 2018; Starkweather et al. 2018; Sarafyazd and Jazayeri 2019; Bartolo and Averbeck 2020). The interaction between cortical state-inference processes, and amygdala learning processes, may lead to deficits in some conditions, when there is a conflict between the best choices predicted by each process. Under these conditions, lesions of the OFC to amygdala pathway may improve performance. Future work can examine this possibility in more detail.

Conclusion

We found that lesions of the amygdala led to deficits in consistently choosing the more frequently rewarded options. We found these deficits in both the What condition, when animals had to learn to choose the best visual object, and in the Where condition, when the animals had to learn to choose the best action. The deficits in choice accuracy followed primarily from switching after a negative outcome, which led to decreased performance due to the stochastic schedules. We did not find deficits in reversal accuracy; thus, monkeys with amygdala lesions were able to reverse their choice–outcome mappings, in both conditions, as well as controls. Inferring reversals in choice–outcome mappings may, therefore, be more dependent on other brain areas, including cortex. Overall, this suggests that the amygdala is important for consistently choosing a rewarded option. Future work should focus on understanding how the network of areas that are important for learning orchestrates the multiple processes involved in learning in dynamic environments.

Supplementary Material

Supplementary material can be found at *Cerebral Cortex* online.

Notes

Conflict of Interest. None declared.

Funding

The National Institute of Mental Health at the National Institutes of Health (grant number ZIA MH002928).

References

- Alexander GE, DeLong MR, Strick PL. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci.* 9:357–381.
- Amaral DG, Price JL. 1984. Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J Comp Neurol.* 230:465–496.
- Amaral DG, Price JL, Pitkanen A, Carmichael ST. 1992. Anatomical organization of the primate amygdaloid complex. In: Aggleton JP, editor. *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction*. New York: Wiley-Liss. pp. 1–66.
- Asaad WF, Eskandar EN. 2008. A flexible software tool for temporally-precise behavioral control in Matlab. *J Neurosci Methods.* 174:245–258.
- Averbeck BB, Costa VD. 2017. Motivational neural circuits underlying reinforcement learning. *Nature Neurosci.* 20:505–512.
- Averbeck BB, Lehman J, Jacobson M, Haber SN. 2014. Estimates of projection overlap and zones of convergence within frontal-striatal circuits. *J Neurosci.* 34:9497–9505.
- Averbeck BB, Murray EA. 2020. Hypothalamic interactions with large-scale neural circuits underlying reinforcement learning and motivated behavior. *Trends Neurosci.* 43:681–694.
- Averbeck BB, Seo M. 2008. The statistical neuroanatomy of frontal networks in the macaque. *PLoS Comput Biol.* 4:e1000050.
- Balleine BW, Liljeholm M, Ostlund SB. 2009. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res.* 199:43–52.
- Barbas H. 1988. Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J Comp Neurol.* 276:313–342.
- Bartolo R, Averbeck BB. 2020. Prefrontal cortex predicts state switches during reversal learning. *Neuron.* 106:1044–1054.
- Bartolo R, Saunders RC, Mitz AR, Averbeck BB. 2020. Information-limiting correlations in large neural populations. *J Neurosci.* 40:1668.
- Basile BM, Karaskiewicz CL, Fiuzat EC, Malkova L, Murray EA. 2017. MRI overestimates excitotoxic amygdala lesion damage in rhesus monkeys. *Front Integr Neurosci.* 11:12.
- Baxter MG, Murray EA. 2002. The amygdala and reward. *Nat Rev Neurosci.* 3:563–573.
- Belova MA, Paton JJ, Salzman CD. 2008. Moment-to-moment tracking of state value in the amygdala. *J Neurosci.* 28:10023–10030.
- Braesicke K, Parkinson JA, Reekie Y, Man MS, Hopewell L, Pears A, Crofts H, Schnell CR, Roberts AC. 2005. Autonomic arousal in an appetitive context in primates: a behavioural and neural analysis. *Eur J Neurosci.* 21:1733–1740.
- Burns LH, Robbins TW, Everitt BJ. 1993. Differential effects of excitotoxic lesions of the basolateral amygdala, ventral subiculum and medial prefrontal cortex on responding with conditioned reinforcement and locomotor activity potentiated by intra-accumbens infusions of D-amphetamine. *Behav Brain Res.* 55:167–183.
- Bussey TJ, Wise SP, Murray EA. 2002. Interaction of ventral and orbital prefrontal cortex with inferotemporal cortex in conditional visuomotor learning. *Behav Neurosci.* 116:703–715.
- Caminiti R, Borra E, Visco-Comandini F, Battaglia-Mayer A, Averbeck BB, Luppino G. 2017. Computational architecture of the Parieto-frontal network underlying cognitive-motor control in monkeys. *eNeuro.* 4:ENEURO.0306-16.2017.
- Cardinal RN, Parkinson JA, Hall J, Everitt BJ. 2002. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev.* 26:321–352.
- Chau BK, Sallet J, Papageorgiou GK, Noonan MP, Bell AH, Walton ME, Rushworth MF. 2015. Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron.* 87:1106–1118.
- Corbit LH, Balleine BW. 2005. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci.* 25:962–970.
- Costa VD, Dal Monte O, Lucas DR, Murray EA, Averbeck BB. 2016. Amygdala and ventral striatum make distinct contributions to reinforcement learning. *Neuron.* 92:505–517.
- Costa VD, Mitz AR, Averbeck BB. 2019. Subcortical substrates of explore-exploit decisions in primates. *Neuron.* 103:53–545 e535.
- Costa VD, Tran VL, Turchi J, Averbeck BB. 2015. Reversal learning and dopamine: a bayesian perspective. *J Neurosci.* 35:2407–2416.
- Cox RW. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 29:162–173.
- Daw ND, O’Doherty JP, Dayan P, Seymour B, Dolan RJ. 2006. Cortical substrates for exploratory decisions in humans. *Nature.* 441:876–879.
- Desimone R, Albright TD, Gross CG, Bruce C. 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci.* 4:2051–2062.
- Durstewitz D, Vitoz NM, Floresco SB, Seamans JK. 2010. Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron.* 66:438–448.
- Ebitz RB, Albarran E, Moore T. 2018. Exploration disrupts choice-predictive signals and alters dynamics in prefrontal cortex. *Neuron.* 97:450–461 e459.
- Friedman DP, Aggleton JP, Saunders RC. 2002. Comparison of hippocampal, amygdala, and perirhinal projections to the nucleus accumbens: combined anterograde and retrograde tracing study in the macaque brain. *J Comp Neurol.* 450:345–365.
- Gershman SJ, Pesaran B, Daw ND. 2009. Human reinforcement learning subdivides structured action spaces by learning effector-specific values. *J Neurosci.* 29:13524–13531.
- Ghashghaei HT, Barbas H. 2002. Pathways for emotion: interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience.* 115:1261–1279.
- Ghashghaei HT, Hilgetag CC, Barbas H. 2007. Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage.* 34:905–923.
- Goldman-Rakic PS, Porrino LJ. 1985. The primate mediodorsal (MD) nucleus and its projection to the frontal lobe. *J Comp Neurol.* 242:535–560.
- Goodale MA, Milner AD. 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15:20–25.

- Groman SM, Keistler C, Keip AJ, Hammarlund E, DiLeone RJ, Pittenger C, Lee D, Taylor JR. 2019. Orbitofrontal circuits control multiple reinforcement-learning processes. *Neuron*. 103:734–746 e733.
- Haber SN, Kim KS, Maily P, Calzavara R. 2006. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J Neurosci*. 26:8368–8376.
- Hampton AN, Adolphs R, Tyszka MJ, O'Doherty JP. 2007. Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron*. 55:545–555.
- Izquierdo A, Murray EA. 2007. Selective bilateral amygdala lesions in rhesus monkeys fail to disrupt object reversal learning. *J Neurosci*. 27:1054–1062.
- Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, Averbeck BB. 2015. The role of frontal cortical and medial-temporal lobe brain areas in learning a Bayesian prior belief on reversals. *J Neurosci*. 35:11751–11760.
- Karlsson MP, Tervo DG, Karpova AY. 2012. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science*. 338:135–139.
- Lau B, Glimcher PW. 2005. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav*. 84:555–579.
- Lau B, Glimcher PW. 2008. Value representations in the primate striatum during matching behavior. *Neuron*. 58:451–463.
- Lee E, Seo M, Dal Monte O, Averbeck BB. 2015. Injection of a dopamine type 2 receptor antagonist into the dorsal striatum disrupts choices driven by previous outcomes, but not perceptual inference. *J Neurosci*. 35:6298–6306.
- Mackintosh NJ, editor. 1994. *Animal learning and cognition*. New York: Academic Press, Handbook of Perception and Cognition.
- Málková L E, Gaffan D, Murray EA. 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.
- Mascaro M, Battaglia-Mayer A, Nasi L, Amit DJ, Caminiti R. 2003. The eye and the hand: neural mechanisms and network models for oculomanual coordination in parietal cortex. *Cereb Cortex*. 13:1276–1286.
- Morecraft RJ, McNeal DW, Stilwell-Morecraft KS, Gedney M, Ge J, Schroeder CM, van Hoesen GW. 2007. Amygdala interconnections with the cingulate motor cortex in the rhesus monkey. *J Comp Neurol*. 500:134–165.
- Murray EA, Rudebeck PH. 2018. Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. *Nature Rev Neurosci*. 19:404–417.
- Murray EA, Wise SP. 1996. Role of the hippocampus plus subjacent cortex but not amygdala in visuomotor conditional learning in rhesus monkeys. *Behav Neurosci*. 110:1261–1270.
- Neftci EO, Averbeck BB. 2019. Reinforcement learning in artificial and biological systems. *Nat Mach Intell*. 1:133–143.
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*. 304:452–454.
- Olejnik S, Algina J. 2000. Measures of effect size for comparative studies: applications, interpretations, and limitations. *Contemp Educ Psychol*. 25:241–286.
- Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, Witten IB. 2016. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neurosci*. 19:845–854.
- Paton JJ, Belova MA, Morrison SE, Salzman CD. 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*. 439:865–870.
- Peck CJ, Lau B, Salzman CD. 2013. The primate amygdala combines information about space and value. *Nature Neurosci*. 16:340–348.
- Peck CJ, Salzman CD. 2014a. The amygdala and basal forebrain as a pathway for motivationally guided attention. *J Neurosci*. 34:13757–13767.
- Peck CJ, Salzman CD. 2014b. Amygdala neural activity reflects spatial attention towards stimuli promising reward or threatening punishment. *eLife*. 3:e04478.
- Peck EL, Peck CJ, Salzman CD. 2014. Task-dependent spatial selectivity in the primate amygdala. *J Neurosci*. 34:16220–16233.
- Rao SC, Rainer G, Miller EK. 1997. Integration of what and where in the primate prefrontal cortex. *Science*. 276:821–824.
- Rhodes SE, Murray EA. 2013. Differential effects of amygdala, orbital prefrontal cortex, and prelimbic cortex lesions on goal-directed behavior in rhesus macaques. *J Neurosci*. 33:3380–3389.
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci*. 2:1131–1136.
- Rothenhoefer KM, Costa VD, Bartolo R, Vicario-Feliciano R, Murray EA, Averbeck BB. 2017. Effects of ventral striatum lesions on stimulus versus action based reinforcement learning. *J Neurosci*. 37:6920–6914.
- Rudebeck PH, Mitz AR, Chacko RV, Murray EA. 2013. Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. *Neuron*. 80:1519–1531.
- Rudebeck PH, Murray EA. 2008. Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *J Neurosci*. 28:8338–8343.
- Rudebeck PH, Ripple JA, Mitz AR, Averbeck BB, Murray EA. 2017. Amygdala contributions to stimulus-reward encoding in the macaque medial and orbital frontal cortex during learning. *J Neurosci*. 37:2186–2202.
- Rudebeck PH, Saunders RC, Lundgren DA, Murray EA. 2017. Specialized representations of value in the orbital and ventrolateral prefrontal cortex: desirability versus availability of outcomes. *Neuron*. 95:1208–1220 e1205.
- Russchen FT, Amaral DG, Price JL. 1987. The afferent input to the magnocellular division of the mediodorsal thalamic nucleus in the monkey, *Macaca fascicularis*. *J Comp Neurol*. 256:175–210.
- Salzman CD, Fusi S. 2010. Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annu Rev Neurosci*. 33:173–202.
- Samejima K, Ueda Y, Doya K, Kimura M. 2005. Representation of action-specific reward values in the striatum. *Science*. 310:1337–1340.
- Sarafyazd M, Jazayeri M. 2019. Hierarchical reasoning by neural circuits in the frontal cortex. *Science*. 364:eaav8911.
- Schuck NW, Cai MB, Wilson RC, Niv Y. 2016. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*. 91:1402–1412.
- Seidlitz J, Sponheim C, Glen D, Ye FQ, Saleem KS, Leopold DA, Ungerleider L, Messinger A. 2018. A population MRI brain template and analysis tools for the macaque. *Neuroimage*. 170:121–131.

- Seo M, Lee E, Averbeck BB. 2012. Action selection and action value in frontal-striatal circuits. *Neuron*. 74:947–960.
- Sereno AB, Maunsell JH. 1998. Shape selectivity in primate lateral intraparietal cortex. *Nature*. 395:500–503.
- Stalnaker TA, Franz TM, Singh T, Schoenbaum G. 2007. Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron*. 54:51–58.
- Starkweather CK, Gershman SJ, Uchida N. 2018. The medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. *Neuron*. 98:616–629 e616.
- Steinmetz MA, Motter BC, Duffy CJ, Mountcastle VB. 1987. Functional properties of parietal visual neurons: radial organization of directionalities within the visual field. *J Neurosci*. 7:177–191.
- Taswell CA, Costa VD, Murray EA, Averbeck BB. 2018. Ventral striatum's role in learning from gains and losses. *Proc Natl Acad Sci USA*. 115:E12398–E12406.
- Turner BH, Mishkin M, Knapp M. 1980. Organization of the amygdalopetal projections from modality-specific cortical association areas in the monkey. *J Comp Neurol*. 191:515–543.
- Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. *Analysis of visual behavior*. Cambridge: MIT Press, pp. 549–587.
- Webster MJ, Bachevalier J, Ungerleider LG. 1994. Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb Cortex*. 4:470–483.
- Willenbockel V, Sadr J, Fiset D, Horne GO, Gosselin F, Tanaka JW. 2010. Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods*. 42:671–684.
- Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014. Orbitofrontal cortex as a cognitive map of task space. *Neuron*. 81:267–279.
- Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. 2014. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci USA*. 111:8619–8624.
- Zar JH. 1999. *Biostatistical analysis*. Prentice Hall: Upper Saddle River.