

Reward Modulates Attention Independently of Action Value in Posterior Parietal Cortex

Christopher J. Peck,^{1*} David C. Jangraw,^{1*} Mototaka Suzuki,¹ Richard Efem,¹ and Jacqueline Gottlieb^{1,2}

Departments of ¹Neuroscience and ²Psychiatry, Columbia University, New York, New York 10032

While numerous studies have explored the mechanisms of reward-based decisions (the choice of action based on expected gain), few have asked how reward influences attention (the selection of information relevant for a decision). Here we show that a powerful determinant of attentional priority is the association between a stimulus and an appetitive reward. A peripheral cue heralded the delivery of reward or no reward (these cues are termed herein RC+ and RC−, respectively); to experience the predicted outcome, monkeys made a saccade to a target that appeared unpredictably at the same or opposite location relative to the cue. Although the RC had no operant associations (did not specify the required saccade), they automatically biased attention, such that an RC+ attracted attention and an RC− repelled attention from its location. Neurons in the lateral intraparietal area (LIP) encoded these attentional biases, maintaining sustained excitation at the location of an RC+ and inhibition at the location of an RC−. Contrary to the hypothesis that LIP encodes action value, neurons did not encode the expected reward of the saccade. Moreover, at odds with an adaptive decision process, the cue-evoked biases interfered with the required saccade, and these biases increased rather than abating with training. After prolonged training, valence selectivity appeared at shorter latencies and automatically transferred to a novel task context, suggesting that training produced visual plasticity. The results suggest that reward predictors gain automatic attentional priority regardless of their operant associations, and this valence-specific priority is encoded in LIP independently of the expected reward of an action.

Introduction

A central question in neuroscience concerns the mechanisms by which animals make reward-based decisions (Sutton and Barto, 1998; Sugrue et al., 2005; Bogacz, 2007). A system of choice for the study of decision making has been the oculomotor system, in particular the mechanisms guiding rapid eye movements (saccades). Many experiments have focused on the lateral intraparietal area (LIP), a cortical area that has a spatiotopic visual representation and is implicated in attention and saccade planning. LIP neurons encode the direction of an upcoming saccade, and their presaccadic responses are scaled by expected reward, suggesting that LIP encodes a representation of action value that specifies the metrics and expected gain of a potential saccade (Platt and Glimcher, 1999; Sugrue et al., 2004).

However, a question left open by prior studies is whether reward information in LIP modulates activity related to attention or saccade decisions. This distinction is significant because atten-

tion is important for monitoring informative or salient objects, even if these objects do not specify a decision alternative. LIP neurons respond robustly to objects that are not action targets but are covertly attended by virtue of their intrinsic salience (Bisley and Goldberg, 2003; Balan and Gottlieb, 2006; Ipata et al., 2006) or relevance to a task (Oristaglio et al., 2006; Balan et al., 2008). Thus, neurons may carry a reward-modulated signal of attention that is distinct from a representation of action value. Prior studies did not examine this possibility, because they used tasks in which all the visual stimuli in the display were part of the decision set (each stimulus represented a decision alternative), confounding an attentional and decisional interpretation (Platt and Glimcher, 1999; Sugrue et al., 2004).

To address this question, we used a novel task in which a peripheral visual cue (RC) predicted the trial's outcome but, to experience the expected outcome, monkeys made a saccade to a separate target whose location was independent of the cue. We report that, although the RCs had no operant significance, they automatically biased attention in valence-specific manner, and these biases were encoded in LIP. Cues predicting reward attracted attention to their location and evoked sustained excitation in LIP, whereas cues predicting no reward repulsed attention and evoked sustained inhibition in LIP. These biases were maladaptive, as they interfered with the required (optimal) saccade to the target. Yet, strikingly at odds with an adaptive decision process, the biases grew rather than abating after prolonged training. The results suggest that LIP encodes the power of reward predictors to bias attention in valence-specific manner, whether or not the attentional biases reflect the expected reward of an action.

Received April 22, 2009; revised July 16, 2009; accepted July 29, 2009.

This work was supported by The National Eye Institute (R01 EY014697-01 and R24 EY015634) and a Swiss National Science Foundation Fellowship to M.S. We are indebted to members of the laboratory, most especially Puiu Balan, for insightful discussions and technical help. We thank M. Osman and G. Asfaw for veterinary care, the Columbia University MRI Research Center for MRI scans, and Latoya Palmer for administrative assistance. C.J.P. and D.C.J. conducted the experiment and analyzed the data. M.S. contributed to data collection and analysis. R.E. performed initial behavioral training. J.G. conceived the experiments and wrote this manuscript.

*C.J.P. and D.C.J. contributed equally to this report.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Jacqueline Gottlieb, Department of Neuroscience, Columbia University, 1051 Riverside Drive, Unit 87, New York, NY 10032. E-mail: jg2141@columbia.edu.

DOI:10.1523/JNEUROSCI.1929-09.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/2911182-10\$15.00/0

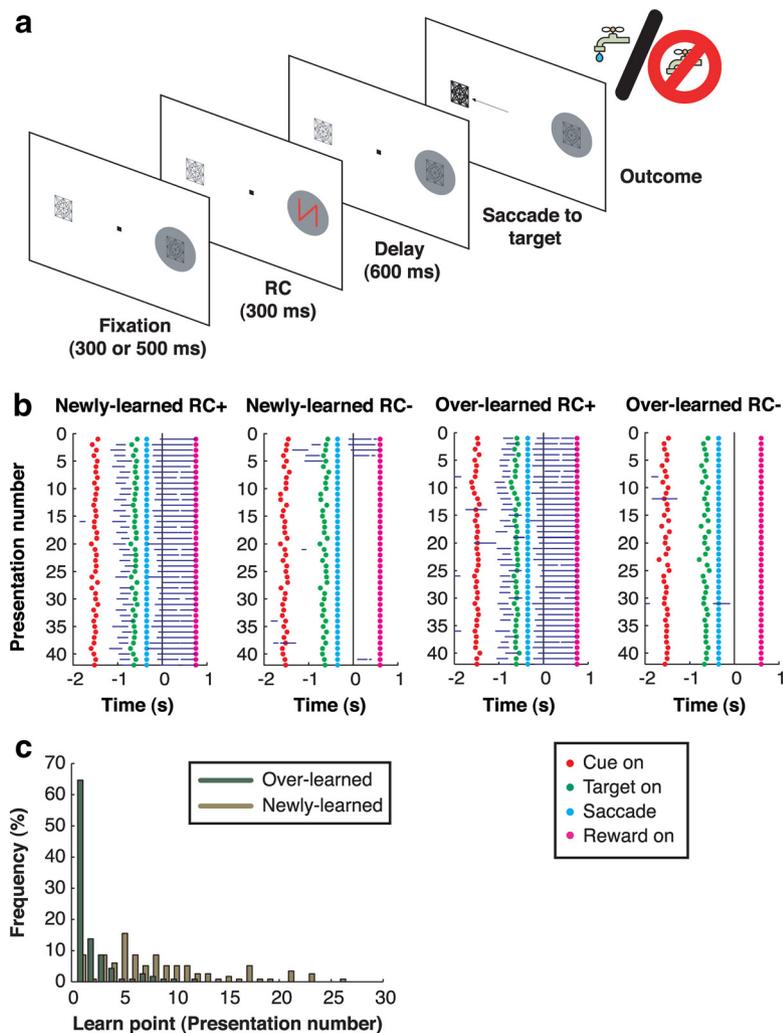


Figure 1. Behavioral task and licking behavior. *a*, Task sequence. A stable display with two placeholders remains visible during the intertrial interval. A trial begins when the monkey achieves central fixation, bringing one of the placeholders into the RF (gray circle). After a fixation period, an RC appears, followed by a delay period and illumination of one of the placeholders. The monkey is required to make a saccade to the illuminated placeholder to receive the outcome predicted by the RC. If applicable (i.e., on a correct RC+ trial), the reward is given 350 ms after the end of the saccade; otherwise, no reward is given. *b*, Licking behavior during an example session. Trials are sorted off-line by RC type and plotted in chronological order, with the first trial on top. Blue horizontal lines indicate the times at which the monkey was licking during each trial. Rasters are aligned at the time of reward delivery on RC+ trials and at the corresponding time point (350 ms after saccade end) on RC- trials. The dots mark trial events as indicated in the legend. *c*, Frequency distribution of behavioral learn points for individual RC- during all 58 recording sessions.

Materials and Methods

General methods. Data were collected from two adult male rhesus monkeys (*Macaca mulatta*) by use of standard behavioral and neurophysiological techniques as described previously (Oristaglio et al., 2006). Visual stimuli were presented on a Sony GDM-FW9000 Trinitron monitor (30.8 by 48.2 cm viewing area) located 57 cm in front of the monkey. The precise timing of stimulus presentation was measured accurately using a diode fixed to the top left corner of the monitor to detect the onset of a refresh cycle. Licking was measured by means of an infrared beam that was projected between the monkey’s mouth and the reward spout and produced a transistor–transistor logic pulse each time it was interrupted by protrusions of the monkey’s tongue. Eye position was recorded using an eye coil system and digitized at 500 Hz. All methods were approved by the Animal Care and Use Committees of Columbia University and New York State Psychiatric Institute as complying with the guidelines within the Public Health Service *Guide for the Care and Use of Laboratory Animals*.

Behavioral task. During the task, two placeholders were continuously present, positioned so they fell in the receptive field (RF) and at the

opposite location when the monkey achieved central fixation (Fig. 1*a*). After the monkey achieved fixation, an RC was then presented for 300 ms either at the RF placeholder or at the opposite location. Some RCs indicated that the trial would end in juice reward (designated as RC+), while others indicated that the trial, even if correctly performed, would end in no reward (designated as RC-). The RC was followed by a 600 ms delay period during which monkeys had to maintain fixation. At the end of the delay, the fixation point was removed and one of the two placeholders (randomly selected) brightened, indicating the saccade target; monkeys had to make a saccade to this target to complete the trial. On RC+ trials, a reward of constant size (250 ms solenoid open time) was delivered at 350 ms after the end of a correct saccade. On an unrewarded (RC-) trial, there was no juice reward, but a 600 ms postsaccade delay was applied to equate the total trial length across RC conditions. Error trials (premature or late saccades, or saccades away from the target) were aborted without reward and immediately repeated until correctly performed.

RCs were abstract computer-generated wireframe figures of distinct shape and color, approximately equated for size and luminance. Stimuli were scaled with retinal eccentricity to range from 1.5 to 3.0° in height and from 1.0 to 2.0° in width. The fixation point was a 0.5 × 0.5° square, and fixation was enforced within 2.5° of the fixation point and 3° of the saccade target. The fixation and saccade windows were constant across trials, so that accuracy requirements did not differ according to reward condition. Eight RCs (two of each of the following: overlearned RC+, overlearned RC-, newly learned RC+, and newly learned RC-) were presented in random order for a total of 32 correct trials per RC or 256 trials per block.

Neural recordings. Electrode tracks were aimed to the lateral bank of the intraparietal sulcus based on stereotactic coordinates and structural MRI. Neurons were tested on the task if they had spatially tuned visual, delay, or presaccadic activity on a standard memory-guided saccade task. A total of 58 neurons (21 from monkey C and 37 from monkey S) provided a full data set and are included in the analysis.

Statistical analysis. All analyses were initially performed for each monkey and, since there were no significant differences between monkeys, the pooled data are presented here. Values in the text represent mean ± SE unless otherwise noted. Statistical comparisons were performed with nonparametric two-sample tests (paired or unpaired Wilcoxon test) or with two-way ANOVA and evaluated at *p* = 0.05 unless otherwise indicated.

Analysis of behavioral data. All analyses are based on correct trials, with the exception of the analyses of saccade accuracy (Fig. 2*a,b*), which included correct and error trials. Although error trials were immediately repeated, sometimes multiple times, we included only trials which the monkey successfully completed on the first repetition in our analysis to avoid biasing the results by long runs of repeated trials. Saccade onset was detected off-line using velocity and acceleration criteria; saccade latency was measured from target onset to saccade onset. Saccade accuracy was defined as (180 - *d*)/180, where *d* is the absolute angular distance, in degrees, between the vectors representing the target and the saccade endpoint relative to the fixation position.

We measured anticipatory licking in a window extending from 20 ms before to 50 ms after reward delivery for RC+ trials and between 300 ms

before and 350 ms after the time when juice would be delivered on a rewarded trial for RC– trials. The more generous time window was used to allow for the possibility that monkeys may be inaccurate at estimating the expected time of reward. Because monkeys licked by default for all newly learned RCs, we defined the learn point by examining the extinction of licking on RC– trials. A trial was considered a “nonlick” trial if the monkey was licking for <20% of the time during the measurement window, and the learn point was defined as the first of three consecutive nonlick RC– presentations.

Analysis of neural data. We calculated receiver operating characteristic (ROC) indices (Green and Swets, 1968; Oristaglio et al., 2006) in five separate analyses as follows: (1) to measure neural reward selectivity, by comparing firing rates on RC+ and RC– trials, pooled across individual RCs (see Fig. 4); (2) to measure visual response latency, by comparing post-RC with pre-RC firing rates, pooled across all RCs; (3) to measure the post-RC spatial bias, by comparing firing rates after RC inside and opposite the RF, 600–900 ms after RC onset (see Fig. 5); (4) to measure selectivity for saccade direction, by comparing activity before saccades toward and opposite the RF, 100–200 ms after target onset, excluding all spikes that occurred after saccade initiation (see Fig. 6); and (5) to compare selectivity for probe valence, by comparing activity evoked by RC+ and RC– probes. To measure reward selectivity (see Fig. 4*b–d*), we calculated ROC values in consecutive nonoverlapping 10 ms bins aligned on RC onset. Raw trial-by-trial spike trains were smoothed using a half-Gaussian filter (half-Gaussian SD = 20 ms), which smeared the signal only forward in time, thus avoiding an underestimation of the latency. For each bin, the firing rate distribution evoked by the two RC+ was compared with that evoked by the two RC–, so that ROC values represent preference for RC+ or RC– regardless of feature selectivity. The statistical significance of each value was assessed using a permutation test ($n = 1000$). A neuron was deemed valence selective if it showed significant ROC indices ($p < 0.001$) for 12 consecutive bins, and the latency of reward selectivity was marked at the beginning of the first of these 12 bins. The same criterion was applied to standard trials (see Fig. 4) and probe trials. To calculate the latency of the visual response, we calculated ROC values comparing RC-evoked firing rates (pooled across all RC types for inside RF presentations) with baseline firing rates 0–30 ms after RC onset. The visual latency was defined as the first of 12 consecutive significant time steps ($p < 0.001$; 1 ms bin width).

To confirm that differences in presaccadic activity (see Fig. 6) between congruent and incongruent trials were not simply a correlate of different saccade metrics for each RC, we repeated the analysis on a subset of trials matched for reaction times. For each RC group (newly learned RC+, overlearned RC+, newly learned RC–, overlearned RC–), we selected a subset (90% of the trials in each original data set) of congruent and incongruent trials that did not differ significantly ($p > 0.5$) in their distributions of reaction times.

Results

Behavioral task

To examine the impact of reward predictors on spatial attention, we used a method borrowed from visual cueing tasks. In these

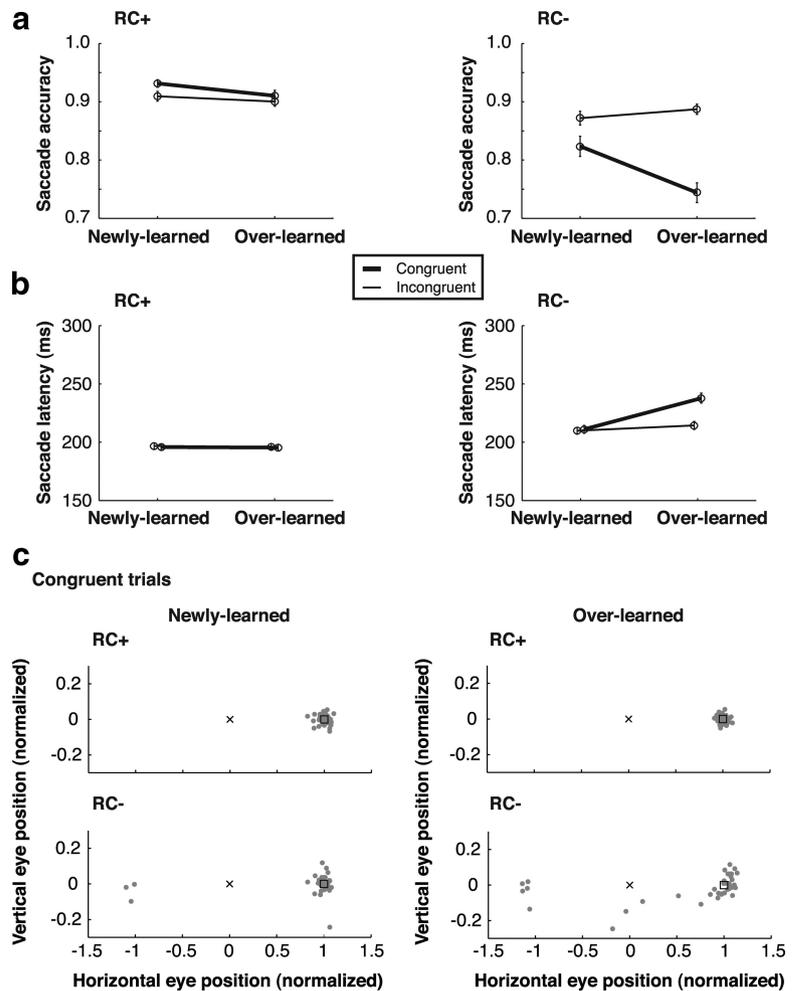


Figure 2. The RCs exert spatial biases on saccade accuracy and RT. *a*, Angular saccade accuracy during all recording sessions ($n = 58$, mean \pm SEM) as a function of RC type and spatial congruence between RC and target (congruent: thick lines; incongruent: thin lines). *b*, Saccade reaction times over all recording sessions in the same format as *a*. *c*, Endpoints of individual saccades for congruent trials in an example session. All saccades are included, regardless of whether they were scored as correct or errant. Saccade coordinates are normalized so that the target (indicated by the open square) was always mapped onto the point (1,0).

tasks, a peripheral visual cue is first presented and is followed, after a delay period, by the appearance of a saccade target either at the cued location or at the opposite location. Although cues are not informative regarding saccade direction, they automatically bias attention by virtue of their bottom-up salience. Attentional biases are measured by comparing saccades directed toward and opposite the cue. At short cue target onset asynchronies, salient cues facilitate same-direction (congruent) relative to opposite-direction (incongruent) saccades, suggesting that they capture attention; at longer asynchronies, the cues impair congruent relative to incongruent saccades, suggesting that they repel attention from their location (Klein, 2000; Fecteau et al., 2004; Fecteau and Munoz, 2005).

In the present task, the peripheral cue did not indicate saccade direction, but it validly signaled the expected reward of the trial (Fig. 1*a*). An RC appeared for 300 ms in the neurons' RF or at the opposite location, and was followed, after a 600 ms delay period, by presentation of a saccade target either at the same or at the opposite location. Positive cues (RC+) signaled that a correct saccade will result in reward, while negative cues (RC–) signaled that, even if correct, the saccade will result in no reward. Because error trials were unrewarded and immediately repeated until cor-

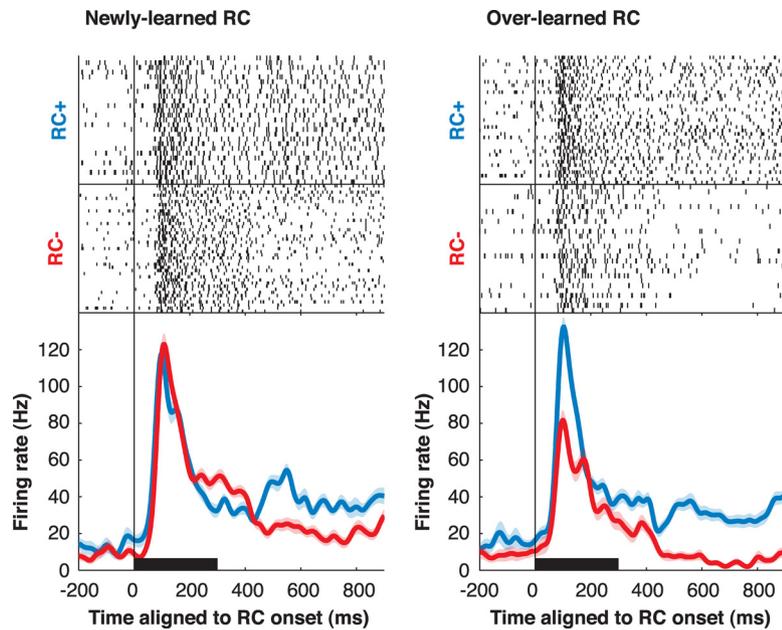


Figure 3. Response of a representative neuron to newly learned and overlearned RCs in its RF. In the raster displays, each tick represents an action potential and each row a single correct trial. Trials are aligned on RC onset and truncated at the end of the delay period, and they are shown in chronological order with the first presentation at the top. Trials with distinct RCs within a category are intermingled. The spike density traces (bottom) show the average firing rates for RC+ (blue) and RC- (red) trials, considering only trials after the learn point for each RC. Shading shows SEM. The black horizontal bar denotes RC duration.

rectly completed, the optimal strategy was to make a saccade to the target regardless of the location or valence of the RC.

Each neuron was tested with a set of newly learned RC that were introduced and trained for one session only. In addition, neurons were tested with a set of familiar (overlearned) RCs, which had been associated with a constant outcome (reward or no reward) for at least 13 sessions before neural recordings began (at least 2600 correct trials per stimulus). Two distinct stimuli were assigned to each RC type (RC+, RC-, newly learned, and overlearned) to control for stimulus-specific effects. All conditions (RC type, RC location, and target location) were counter-balanced and randomly interleaved within a session.

Behavioral performance

We used three measures of performance. First, to ascertain that monkeys learned the reward valence of the cue, we measured anticipatory licking, a conditioned response that is a reliable measure of appetitive learning (Schultz, 2006). For overlearned RCs, monkeys licked selectively for RC+ but not RC- from the very first cue presentation, indicating their familiarity with the stimuli (Fig. 1*b*, two right panels). For novel RC, monkeys began by licking for all stimuli but, within the first few presentations, ceased licking for RC- (two left panels). The average learn point (the first presentation of an RC- at which licking reliably extinguished) (Fig. 1*c*) was 8.41 ± 0.54 presentations for newly learned RCs (mode: five presentations; $n = 58$) and 2.03 ± 0.19 presentations for overlearned RCs (mode: 1 presentation; $n = 58$). Trials before the learn point were excluded from subsequent analysis. Thus, in comparing newly learned and overlearned RCs, we capture the differential effects of stimuli that had known reward valence, but on which monkeys had been trained for short or long periods of time. (We use the term “overlearned” where others may use “well learned” or “long learned” to indicate that training proceeded long past the behavioral learn point as defined by anticipatory licking.)

Second, to assess how reward impacted motivation, we compared saccade metrics on rewarded and unrewarded trials. Reward expectation improved saccade performance, as shown by a higher fraction of correctly completed trials (91 vs 76%), higher saccade accuracy (Fig. 2*a*, compare left and right panels) ($p < 10^{-4}$; Wilcoxon signed-rank test), and shorter reaction times (Fig. 2*b*, compare left and right panels) ($p < 10^{-79}$) on RC+ relative to RC- trials. This replicates the well known effects of motivation on operant behavior (Watanabe et al., 2001; Bendiksbj and Platt, 2006; Kobayashi et al., 2006; Roesch and Olson, 2007).

Our third and primary measure determined whether the RCs biased attention in a spatially specific manner. To answer this question, we examined whether the RCs differentially affected saccades that happened to be directed toward or opposite to the RC location. We found that the RC spatially biased saccades in a valence-specific manner, such that the RC+ slightly attracted (facilitated) saccades toward its location, whereas the RC- strongly repulsed (impaired) saccades

away from its location. Figure 2 compares congruent and incongruent saccades with respect to accuracy (Fig. 2*a*) and response time (RT) (Fig. 2*b*). On RC+ trials, saccade accuracy was slightly higher on congruent than on incongruent trials (Fig. 2*a*, left), revealing a slight attractive effect of the RC+ (two-way ANOVA; $p = 0.049$ for effect of congruence, $p > 0.05$ for effect training length, and $p = 0.455$ for interaction). On RC- trials, in contrast, accuracy was markedly lower (Fig. 2*a*, right) and RT was higher (Fig. 2*b*, right) on congruent relative to incongruent trials, suggesting a strong repulsive effect. Figure 2*c* illustrates the endpoints of saccades on congruent trials in a representative session. Whereas on RC+ trials saccade endpoints were tightly clustered around the target (top), on RC- trials saccades had large endpoint scatter, and many fell outside the target window. This was not simply an effect of motivation, as this large scatter was seen only for congruent saccades (Fig. 2*a*, right). Moreover, the effect was maladaptive, as it caused many erroneous saccades. Nevertheless, the effect grew with training, becoming worse for the overlearned relative to newly learned RCs. A two-way ANOVA across the data set revealed, for RC- trials, significant effects of congruence, training length, and interaction for both accuracy ($p < 10^{-9}$, $p = 0.025$, and $p = 0.001$, respectively) (Fig. 2*a*, right) and RT ($p = 0.0001$, 0.002 , 0.005) (Fig. 2*b*, right). Thus, the RC- generated a spatial repulsion that increased with training even though it lowered the rate of reward.

LIP neurons show valence selectivity that increases with long-term training

To see whether LIP neurons encoded the reward effects in this task, we tested their responses to the RC and saccade target. In the following, we first compare neuronal valence selectivity for newly learned and overlearned RCs appearing in the RF. We next examine the spatial specificity of RC responses by comparing trials in which the RC appeared inside and opposite the RF. Finally, we

analyze the impact of reward on responses to the saccade target.

When cues appeared in the RF, neurons had a fast transient response to cue onset, which soon differentiated according to cue valence, becoming stronger after RC+ relative to RC-. Responses of a representative neuron are shown in Figure 3, and population responses ($n = 58$ neurons) are shown in Figure 4*a*. Valence selectivity did not simply reflect neural preference for RC shape or color, as it clearly persisted when distinct RCs were averaged within each reward category (Figs. 3, 4*a*, averaged histograms). Of the neurons showing significant modulation by reward, some had additional stimulus selectivity or interaction between stimulus and reward (14/25 for newly learned RC, 24/41 for overlearned RC; two-way ANOVA; 100–500 ms after RC onset). However, among neurons that were modulated either by reward or by stimulus identity, the former were significantly more prevalent (25/27 for overlearned RC, 14/26 for newly learned RC; χ^2 test; $p = 0.01$). Thus, although shape/color preference was present in a subset of cells (Serenio and Maunsell, 1998), it could not account for the stronger and more prevalent sensitivity to reward.

As shown in Figures 3 and 4, reward effects were larger for overlearned relative to newly learned cues. Because monkeys understood the reward valence of all RC (as shown by their anticipatory licking), these differences represent a long-term effect that develops with a time course longer than the fast acquisition of the conditioned response. We note, however, that neurons also showed a short-term learning effect, acquiring valence selectivity for newly learned RCs within the first few trials. A population analysis of this fast-learning component is shown in supplemental Figure S1, available at www.jneurosci.org as supplemental material.

We quantitatively measured valence selectivity using an ROC analysis (Green and Swets, 1968) comparing the distribution of firing rates evoked by RC+ and RC-. This yielded an ROC index ranging between 0 and 1, where values of 0.5 indicate no reward selectivity and values below and above 0.5 indicate preference for RC- and RC+, respectively. Significant valence selectivity, predominantly preference for RC+, was found in a vast majority of neurons (Fig. 4*b*). Long-term training increased the fraction of selective neurons from 85% for newly learned to 95% for overlearned RC ($p = 0.023$; χ^2 test) and significantly strengthened overall preference for the RC+ (Fig. 4*d*). Training also decreased the latency of reward selectivity. Median latencies of the reward effect were, across individual neurons (Fig. 4*c*), 155 ms for overlearned RCs versus 245 ms for newly learned RCs ($p = 0.0018$), and in the population response, 95 versus 235 ms (Fig. 4*d*). To

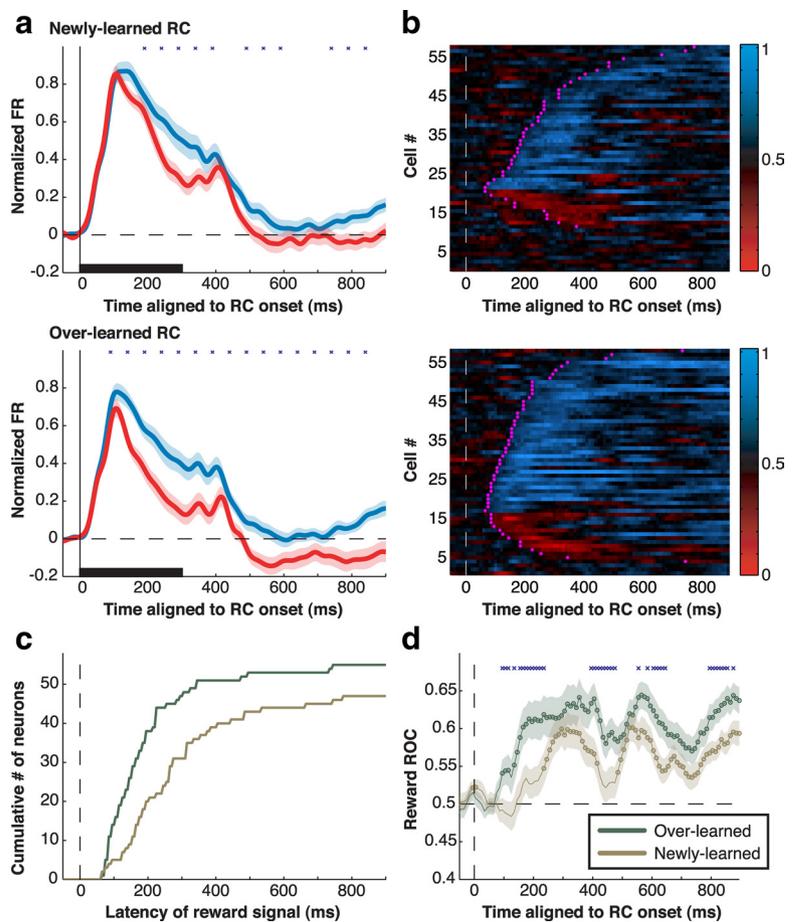


Figure 4. Population analysis of reward modulation. *a*, Average normalized firing rates for the sample of neurons ($n = 58$) in response to RC- (red) and RC+ (blue) stimuli in the RF for newly learned (top) and overlearned (bottom) stimuli. Responses are aligned on RC onset (time 0) and are truncated at the end of the delay period. Shading indicates SE. Firing rates were normalized for each neuron by subtracting the baseline rate (50 ms before RC onset) and dividing by the peak response across all RC types. Stars denote 100-ms-firing-rate bins (beginning at 40 ms and shifted by 50 ms) with statistically significant differences between RC+ and RC- (Wilcoxon signed-rank test; $p < 0.05$). An apparent minor response peak at ~ 400 ms is attributable to a subset of neurons that showed an off response to the disappearance of the RC. *b*, ROC analysis of reward modulation. Each row represents an individual neuron, and each pixel an ROC index in a 10 ms time bin. The white dashed line shows RC onset (time 0). ROC values of >0.5 (blue) signify preference for RC+, values of <0.5 (red) signify preference for RC-, and values close to 0.5 (black) signify no preference. Magenta crosses mark the latency at which each cell met the significance criterion for a reward effect. *c*, Cumulative distribution of the reward latencies for overlearned (green) and newly learned (goldenrod) RCs. *d*, Population average of the ROC values in *b* for overlearned and newly learned RCs. Circles denote time bins in which the value is significantly different from 0.5 ($p < 0.05$), and blue crosses denote bins with significant differences between overlearned and newly learned stimuli. Shading shows SEM.

directly examine whether valence selectivity was present in the early visual response, we calculated selectivity in a 70 ms window aligned on each neuron's visual response latency (see Materials and Methods). The fraction of neurons with significant selectivity in this early time window increased from 15% for newly learned RCs to 41% for overlearned RCs ($p = 0.0020$; χ^2 test). Thus, long-term training increased the prevalence and magnitude and decreased the latency of valence coding in LIP.

Measurement of firing rates showed that the training-related increase in selectivity was associated with a decline in visually evoked responses. Visual responses (100–500 ms after RC onset) declined by 4.90 ± 0.86 spikes per second (sp/s) for overlearned versus newly learned RC+, but showed a larger decline for RC-, of 8.32 ± 1.37 sp/s (in normalized units, differences were 0.07 ± 0.02 vs 0.14 ± 0.03 ; all $p < 10^{-5}$ relative to 0 and $p < 10^{-7}$ for RC+ versus RC-). Thus, prolonged training produced a global decline in firing rates for both RC+ and RC-, which may have indicated an effect of stimulus familiarity, and an additional a

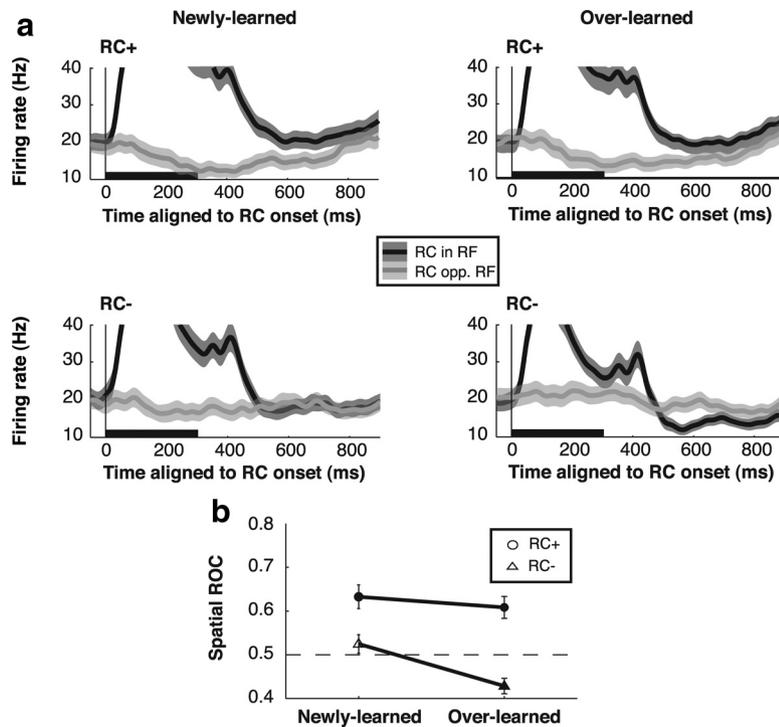


Figure 5. Spatial effects of RC during the delay period. *a*, Population firing rates on trials in which the RC appeared in the RF (black) or opposite (opp.) the RF (gray) for each RC type. Shading shows SEM. The vertical scale is expanded and truncates the visual response to highlight delay period activity. *b*, Population ROC values for RC location in the last 300 ms of the delay period. All symbols show mean \pm SEM. Circles indicate RC+ trials, and triangles, RC- trials. Filled symbols show values that are significantly different from 0.5 (*t* test; $p < 0.05$).

Table 1. Firing rates (sp/s, mean \pm SEM), according to RC training history and location, in the visual epoch (100–500 ms relative to RC onset) and at the end of the delay period (600–900 ms after RC onset)

Cue location	Newly learned		Overlearned	
	Visual	Delay	Visual	Delay
RC+				
In RF	45.96 \pm 3.72**	21.96 \pm 2.38**	41.06 \pm 3.39**	21.16 \pm 2.26*
Opposite RF	14.75 \pm 2.20**	18.16 \pm 2.36	15.10 \pm 2.24*	17.43 \pm 2.28
RC-				
In RF	41.34 \pm 3.58**	18.34 \pm 2.15	33.02 \pm 2.96**	14.04 \pm 1.47*
Opposite RF	17.96 \pm 2.34	18.54 \pm 2.17	21.37 \pm 2.44	18.53 \pm 1.94

* $p < 0.05$, ** $p < 0.001$ relative to baseline. Baseline (50 ms before RC onset) = 19.12 \pm 2.30 sp/s.

valence-specific effect—an especially pronounced decline for stimuli predicting no reward.

Reward cues produce spatial biases in sustained activity

The observation that neurons maintained valence selectivity throughout the delay period (Fig. 4) was surprising given that the RC were uninformative regarding saccade direction. One possibility is that these sustained responses reflect global effects of motivation—a general increase in activity in rewarded relative to unrewarded trials regardless of RC location—as reported for the frontal lobe (Roesch and Olson, 2004; Kobayashi et al., 2006) and LIP itself (Bendiksbj and Platt, 2006). Alternatively, the cue-evoked responses may be spatially specific, evoking relatively higher or lower activity at the cue location, consistent with a spatial attentional bias toward or away from the cue (Bisley and Goldberg, 2003). To distinguish these possibilities, we compared sustained activity evoked by the RC when these appeared inside and opposite the RF (Fig. 5*a*, black vs gray traces).

Sustained RC-evoked responses were spatially specific, appearing only if the RC were presented in the RF. After presentation of an RC+ in the RF (Fig. 5*a*, top), neurons generated sustained excitation. However, if the RC+ appeared opposite the RF, there was no response during the delay period, although there was a transient decline in firing during the visual epoch, suggestive of a push-pull mechanism. Table 1 provides detailed comparisons of visual and delay firing rates with the pre-RC baseline. If an overlearned RC- appeared in the RF (Table 1, bottom of rightmost column) the transient visual response was followed by sustained inhibition during the delay period; again, there was no response if the RC- appeared at the opposite location (bottom right). We measured these spatial biases using ROC analysis, comparing delay period firing rates on inside-RF and opposite-RF trials (Fig. 5*b*; note that these ROC indices reflect neural selectivity for RC location not valence). For RC+, spatial ROC indices were >0.5 , indicating an attractive bias toward the cue’s location (0.63 \pm 0.03 for newly learned RC, 0.61 \pm 0.02 for overlearned RC; both $p < 10^{-4}$ relative to 0.5). For newly learned RC-, there was no significant bias (0.53 \pm 0.02; $p = 0.25$), but for overlearned RC-, indices were lower than 0.5, indicating a repulsive bias away from the RC- location (0.43 \pm 0.02; $p < 10^{-3}$). Thus, the RC set up a spatial bias across the topographical representation in LIP which was attractive toward the location of an RC+ and repulsive away from an overlearned RC-, consistent with the spatial biases exerted by the RC on saccades (compare with Fig. 2).

In principle, it is possible that neurons encode RC valence during the delay period and encode expected reward only after saccade direction is specified—i.e., after presentation of the saccade target (Platt and Glimcher, 1999; Sugrue et al., 2004). Therefore, we analyzed activity during the saccade reaction time, between target presentation and saccade onset. We computed saccade direction selectivity as the difference in firing rates before saccades toward and opposite the RF (Fig. 6*a*) and quantitatively measured this selectivity using ROC analysis (comparing firing rates before saccades into and opposite the RF, 100–200 ms after target onset) (Fig. 6*b*). Neurons reliably encoded saccade direction in all trials, as shown by the positive response differences in Figure 6*a* and ROC values of >0.5 in Figure 6*b*. However, directional responses were not affected by expected reward. A Wilcoxon signed-rank test showed no significant differences in ROC indices between rewarded and unrewarded trials ($p = 0.132$ for congruent saccades and $p = 0.105$ for incongruent saccades) (Fig. 6*b*, compare right and left panels). On the other hand, presaccadic responses did reflect congruence with the RC on unrewarded trials, being stronger for incongruent relative to congruent saccades (a two-way ANOVA on ROC values revealed, for RC-, $p = 0.026$ for effect of congruence and $p > 0.7$ for effect of training and interaction; for RC+, $p > 0.3$ for all effects). Thus, neural activity continued to reflect the RC-induced bias but was not modulated by expected reward even while monkeys were planning the saccade.

A potential concern raised by these observations is that reward may have affected saccade metrics by a mechanism outside of LIP, and LIP neurons merely reflected this change in saccade metrics. However, this concern can be ruled out with respect to saccade accuracy, because, while the analysis of saccade accuracy in Figure 2 included correct and error saccades, the neuronal analyses in Figures 5 and 6 were based only on correct trials. In correct trials, saccade accuracies were high and statistically equivalent for all RC and congruence classes, so that neuronal valence effects could not be a secondary effect of saccade endpoint scatter. To rule out that neural effects were attributed to differences in RT (which were significant even for correct trials) (Fig. 2*b*, right), we recalculated saccade ROC indices in a subset of trials that were equated for RT (see Materials and Methods). Significant effects of congruence persisted in this subset (RC−: $p = 0.0463$ for effect of congruence, $p > 0.52$ for training and interactions; RC+: $p > 0.33$ for all; two-way ANOVA). Thus, neurons encoded RC valence and not a spurious effect of saccade accuracy or RT.

Learned salience and visual plasticity

As shown in Figure 4, prolonged training increased the strength and reduced the latency of reward selectivity in LIP. The short latency of reward effects for over-trained RC has two potential explanations. First, training may have increased the speed of reward evaluation and of the reward-related feedback being relayed to LIP. Alternatively, training may have induced a hardwired (plastic) change in the bottom-up visual response to the RC. We reasoned that, if training produced visual plasticity, valence selectivity should automatically transfer to a novel context and persist even if the RC no longer predicts reward.

We tested a subset of neurons in a new behavioral context, probe trials, which were presented after a block of standard trials. Probe trials were identical with standard trials in that they started with a valid RC followed by a delay period and by presentation of a saccade target. However, in contrast with standard trials, the RC and saccade target always appeared opposite the RF, distracting attention as much as possible away from the RF (Fig. 7*a*). Simultaneous with illumination the saccade target opposite the RF, a task-irrelevant probe was briefly flashed in the RF. The RF probe was chosen from the set of RCs trained on the previous block of standard trials, but its valence was unpredictable of reward.

In environments containing multiple potential predictors, animals use the earliest and most reliable predictor to infer expected reward (Fanselow and Poulos, 2005). Thus, we expected that the first, informative, RC would block the monkeys' interpretation of probe valence. This was confirmed by measurements of anticipatory licking (Fig. 7*b*). Although reward was delivered 500–600 ms after probe onset (350 ms after the end of the saccade), allowing ample time to generate anticipatory licking in response to the

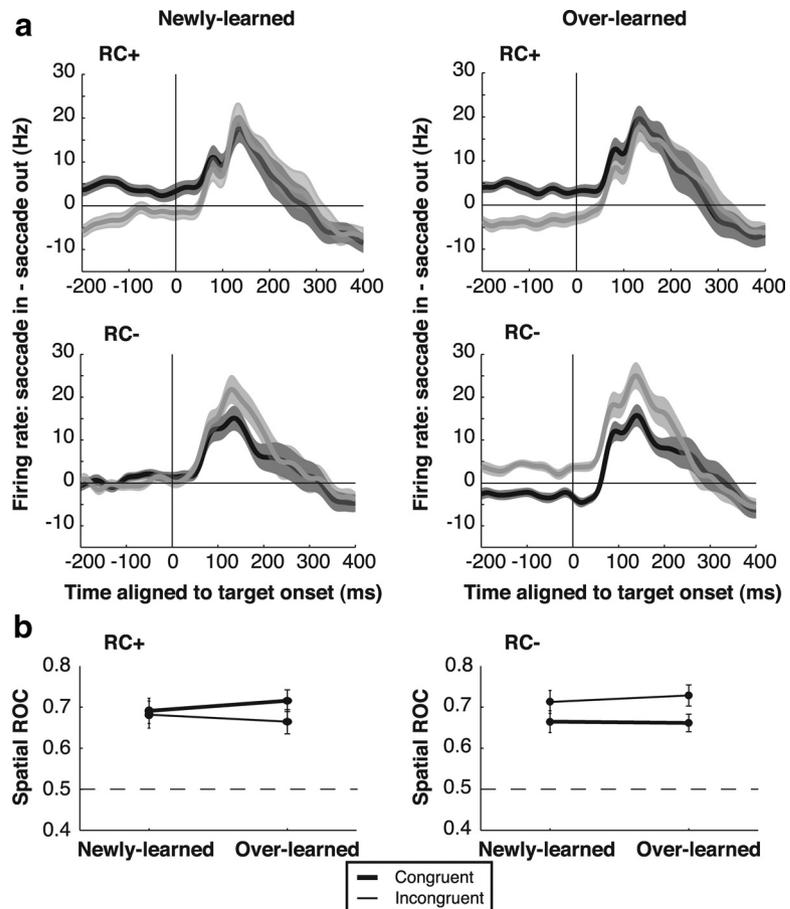


Figure 6. Spatial effects of the RC on presaccadic activity. *a*, Difference traces between activity on saccade into the RF and saccade of the RF trials. Trials are sorted by congruence (congruent: black; incongruent: gray) for each RC category and are aligned on the onset of the saccade target. Shading shows SEM. *b*, ROC for saccade direction 100–200 ms after target onset. All values are significantly more than 0.5 and are thus shown with filled symbols.

probe, licking depended solely on the informative cue and was entirely unaffected by the probe ($p < 10^{-7}$ for main effect of first RC valence; $p > 0.89$ for main effect of probe valence and interaction; two-way ANOVA). This suggests that monkeys actively evaluated the reward valence of the first cue but not that of the probe.

Despite their lack of relevance for reward, the probes were still salient visual stimuli and were expected to elicit visual responses and bottom-up shifts of attention. If the visual responses and/or the attentional weight of the stimuli were permanently modified by reward training, these bottom-up responses may be valence-specific.

Examination of neural responses and saccade latencies confirmed this result with respect to overtrained probes (Fig. 8). For overlearned probes, significant valence selectivity was present across the population (Fig. 8*a*, top) (ROC analysis; 130–230 ms after probe onset; $p = 0.0064$; $n = 34$ neurons) and individually in 15 of the 34 neurons tested (40%) (Fig. 8*b*, main panel). Despite the differences in task conditions, selectivity on probe trials was positively correlated with that on standard trials ($r = 0.44$; $p = 0.0099$; $n = 34$). There was no interaction between probe responses and reward expectation (the valence of the informative RC) (see supplemental Fig. S2, available at www.jneurosci.org as supplemental material) (all $p > 0.71$; two-way ANOVA), showing that these effects could not merely reflect reward expectation. Because visual responses in LIP correlate

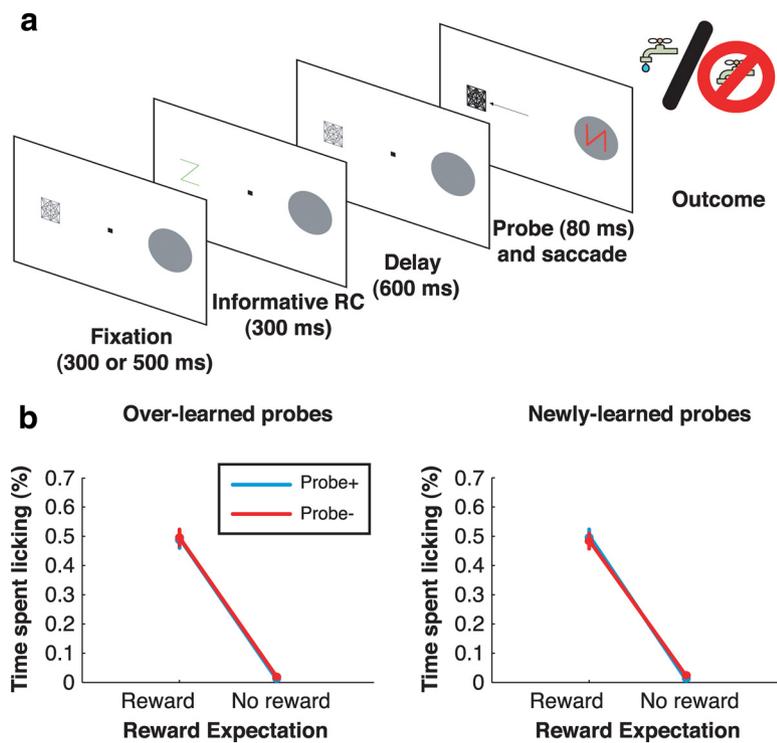


Figure 7. Probe task structure and licking behavior. *a*, Structure of probe trials. A single placeholder was present at the location opposite the RF during the intertrial interval, marking the location of the informative RC and the saccade target. A trial began as before with presentation of an informative RC followed by a 600 ms delay period. Simultaneous with target onset, a behaviorally irrelevant probe was flashed inside the RF. The probe was flashed for 80 ms and extinguished before the onset of the saccade. *b*, Licking behavior during the probe task. Percentage of time spent licking (mean \pm SEM over all probe sessions; $n = 34$) immediately before juice delivery (or the equivalent time on RC– trials) as a function of the expected reward predicted by the informative RC (x-axis) and the valence of the probe (blue: probe RC+; red: probe RC–).

with the distracting power of a task-irrelevant stimulus (Balan and Gottlieb, 2006; Ipata et al., 2006), this suggests that an RC+ probe would produce stronger interference with the saccade relative to an RC– probe. Consistent with this, saccade latencies were longer if the saccade was performed in the presence of an RC+ relative to an RC– (Fig. 8c, left, blue vs red traces; $p = 0.058$ for rewarded trials and 0.0023 for unrewarded trials). Note that probes did not affect motivation, as this effect would have had the opposite sign (i.e., shorter RT in the presence of an RC+ relative to an RC–).

In contrast to overlearned probes, the visual responses evoked by newly learned probe were not valence selective (Fig. 8a, bottom; ROC analysis; $p > 0.71$), even though the neurons showed clear selectivity in the corresponding time window on standard trials (inset). Significant selectivity was present in only 7 of 34 individual neurons, was not correlated between probe and standard trials ($r = 0.14$; $p = 0.4442$), and was absent even in the subset ($n = 15$) that was selective for overlearned probes (Fig. 8c, inset). Consistent with these neural results, newly learned probes did not differentially affect saccade RT (Fig. 8c, right, red vs blue lines; $p = 0.65$ and 0.13). Thus, the valence-specific attentional effects of an overtrained RC, but not the effects of a newly learned RC, automatically transfer to a novel context in which monkeys do not actively evaluate RC valence.

Discussion

While multiple studies have explored the factors that govern saccade decisions, much less is known about how the brain determines the attentional priority or salience of informative stimuli.

Here we show that a powerful determinant of attentional priority is the learned association between a stimulus and an appetitive reward. Stimuli associated with reward gain an enhanced representation in LIP and attract attention to their location; stimuli associated with no reward evoke lower or inhibitory responses in LIP and repel attention from their location. This valence-dependent priority is assigned automatically even when it is objectively nonoptimal—when a reward predictor is spatially separate from, and interferes with, a required action. The results suggest that associations between a stimulus and an appetitive reward, even when established independently of an operant association, are important determinants of the power of the stimulus to attract attention.

Multiple effects of reward

The estimation of potential gains or losses is critical for survival, and it is not surprising that reward computations produce diverse behavioral effects mediated by multiple neural mechanisms. In the present task, we identified three distinct behavioral effects of reward; however, only one of these effects was encoded in LIP. First, reward expectation engendered a conditioned response, anticipatory licking. The properties of anticipatory licking were dissociated from firing rates in LIP.

While monkeys acquired discriminatory licking after the first few stimulus presentations, valence effects in LIP continued to grow on much longer time scales, becoming larger for overlearned relative to newly learned RC. In addition, on probe trials neurons encoded the valence of task-irrelevant probe stimuli, even though these stimuli did not affect licking. These dissociations suggest that anticipatory licking did not depend on LIP, in line with a wealth of evidence showing that conditioned behaviors depend on subcortical, not cortical, mechanisms (Fanselow and Poulos, 2005). A second effect of reward was on motivation, resulting in superior saccade performance on trials that culminated in reward. LIP neurons did not encode motivation either, as firing rates did not differ between rewarded and unrewarded trials unless the RC was in the RF. Motivation may have been mediated by prefrontal and premotor areas that are sensitive to expected outcome (Kobayashi et al., 2006; Roesch and Olson, 2007). A final, and previously uninvestigated, effect was the ability of the RC to bias spatial attention. Positive reward predictors attracted attention, whereas negative predictors repelled attention from their location. These spatial biases were encoded in LIP through sustained excitatory or inhibitory responses specific to the RC location. Thus, LIP neurons did not reliably encode nonspatial aspects of reward computations indexed by conditioned behaviors or motivation but encoded only a very specific effect: the valence-specific attentional weight of a stimulus associated with reward.

The specificity of these responses speaks to a longstanding question regarding the role of LIP in reward-based behaviors (Maunsell, 2004). Our findings suggest that LIP is not a critical

part of the neural network that evaluates potential gains or losses (Schultz, 2006), but is important for expressing the results of this evaluation in spatial behaviors. Consistent with this, we recently showed that reversible inactivation of LIP does not affect reward evaluation processes themselves but affects the ability to use reward (or other sources of information) in a spatially unbiased manner (Balan and Gottlieb, 2009). Thus, LIP provides a visuospatial map that reads out the outcome of reward computations in spatial terms for the purpose of guiding attention.

Distinguishing between the reward effects on attention and action

Although the proposal that LIP represents a pragmatic reward-modulated spatial representation has been advanced before (Sugrue et al., 2004), previous studies concluded that this representation encodes the expected reward of an action (a saccade) (Platt and Glimcher, 1999; Dorris and Glimcher, 2004; Sugrue et al., 2004). However, the present results suggest that this interpretation may not provide a general description of expected reward influence on LIP responses. In the present task, neurons encoded the reward valence of a stimulus even when this valence was not aligned with the reward of the saccade.

Four central properties distinguish the reward modulations in our task from a code of action value. First, the action value hypothesis predicts that after the initial visual response, neurons should weight each potential target position equally (i.e., have equal delay period firing rates regardless of RC location) given the equal probability that the target will appear there. However, we find that neural responses (Fig. 5) and subsequent saccades (Fig. 2) remained biased by RC valence and location, even after a delay following extinction of the RC. Second, the action value hypothesis predicts that neural responses should differ according to RC valence when the target appears in the neuron's RF and the "action-value" contingency has been established. Contrary to this prediction, presaccadic responses in LIP were not modulated by expected reward but reflected only the biases exerted by the RC (Fig. 6). Third, a reward-enhanced signal of action value is expected to be adaptive, facilitating the choice of action that harvests the higher reward. In contrast, the reward-enhanced responses in our task were maladaptive, interfering with the required saccade. Finally, an account rooted in decision processes implies that reward prediction will improve with learning, becoming more closely aligned with an optimal (reward-maximizing) strategy (Sugrue et al., 2004). In contrast, learning in our task impaired performance, exacerbating the maladaptive effects exerted by the RC- (Figs. 2–6). It may be argued that monkeys did not detect the detrimental effect of an RC- error, because the trial was unrewarded anyway and an error reduced reward only over longer time scales (delayed the opportunity to progress to a rewarded trial). However, this is highly unlikely. Had monkeys only used short-term reward evaluation, they would have immediately aborted each RC- trial; in addition,

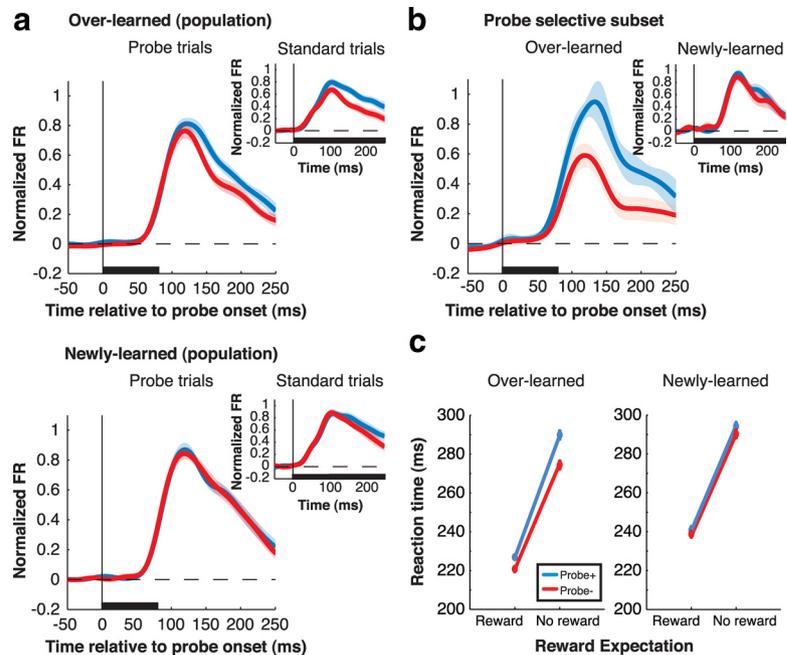


Figure 8. Neural responses and saccade reaction times during the probe task. *a*, Population neural response ($n = 34$) during the probe task (main panels) and standard task (same 34 neurons; insets). Trials were sorted by learning history (overlearned: top; newly learned: bottom) and the valence of the stimulus in the RF (blue: RC +; red: RC -). Responses to the probes are truncated at 250 ms—the average time at which the saccade moved the neuron's RF away from the probe location. Responses during the standard task are truncated at 250 ms as well for the sake of comparison. Shading is \pm SEM. *b*, Average response of neurons with significant selectivity for the valence of overlearned probes ($n = 15$). Inset shows the responses of the same neurons for newly learned probes. *c*, Saccade reaction times during the probe task. Mean reaction times (\pm SEM) are plotted as a function of expected reward (x -axis) and probe valence (blue: probe +; red: probe -), separately for overlearned and newly learned probes.

because an error trial was immediately repeated, its highest impact was on reward rate immediately after an error, precisely the time scale that monkeys seem to rely on when estimating expected reward (Sugrue et al., 2004). Thus, the RC effects in our task were distinguished from action value in that they acted automatically, regardless of the operant significance of the reward predictors and persisted even when they interfere with optimal behavior.

Our findings do not exclude the possibility that in a restricted set of circumstances LIP may provide a de facto signal of action value, as concluded in previous studies. Indeed, in the conditions used in these studies, all the stimuli presented to the monkey were part of the decision set (each stimulus represented a decision alternative); in these circumstances, the expected reward of a stimulus is equivalent to the expected reward of the saccade, and LIP activity accurately reflects action value (Platt and Glimcher, 1999; Dorris and Glimcher, 2004; Sugrue et al., 2004). However, in natural behavior, decisions are forged in the presence of multiple stimuli that may have reward valence but do not specify immediate action. In these conditions, LIP is likely to mediate attentional biases produced by reward predictors even when they come in conflict with the objective gain of the action. Thus, LIP reflects reward evaluation at the level of stimulus selection. In contrast, critical aspects of action evaluation [like critical aspects of a saccade decision itself (Gottlieb and Goldberg, 1999)] seem to be computed in separate structures, possibly downstream from LIP.

Learning of attentional priority

The tenacity and automaticity of the effects we describe suggests that these effects may be rooted not so much in cognitive decision processes but in more automatic forms of learning perhaps related to

emotional learning. In human subjects, stimuli with intrinsic emotional significance (e.g., fearful faces) automatically attract attention and modify visually evoked responses and psychophysical performance even when they are irrelevant to a task (Vuilleumier, 2005; Phelps et al., 2006; Padmala and Pessoa, 2008). Our results extend these findings by demonstrating that attentional biases are also produced by predictors of appetitive, not aversive, outcomes, possibly representing a mechanism for automatic monitoring of information regarding appetitive reward.

A final important finding we report is that prolonged reward learning produced profound changes in LIP responses, consistent with visual plasticity. Prior studies have documented learning in the visual and oculomotor system in operant tasks requiring conditional visuomotor learning (Chen and Wise, 1996; Asaad et al., 1998), categorization (Sigala and Logothetis, 2002; Freedman and Assad, 2006), decision making (Sugrue et al., 2004), or target selection (Bichot et al., 1996; Mruczek and Sheinberg, 2005, 2007). Our results extend these findings in two ways. First, we show that learning occurs merely through training of stimulus-response associations in the absence of operant associations. Second, we provide explicit evidence that training produced plasticity in the bottom-up visual response. We showed that, for overtrained RC, the valence-specific effects automatically transferred to a novel context in which the stimuli did not govern reward expectation. This suggests that long-term training conferred on the RC an intrinsic salience that was akin to the bottom-up salience of a conspicuous object. This form of learning may underlie the special salience of highly familiar stimuli such as the letters of the alphabet or a friend's face, which automatically pop out from a crowded visual scene. Thus, our results support the idea that operant learning involves, at least in part, changes in the salience of task-relevant objects driven by learned associations between these objects and reward (Ferrera and Grinband, 2006).

References

- Asaad WF, Rainer G, Miller EK (1998) Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21:1399–1407.
- Balan PF, Gottlieb J (2006) Integration of exogenous input into a dynamic salience map revealed by perturbing attention. *J Neurosci* 26:9239–9249.
- Balan PF, Gottlieb J (2009) Functional significance of nonspatial information in monkey lateral intraparietal area. *J Neurosci* 29:8166–8176.
- Balan PF, Oristaglio J, Schneider DM, Gottlieb J (2008) Neuronal correlates of the set-size effect in monkey lateral intraparietal area. *PLoS Biol* 6:e158.
- Bendiksy MS, Platt ML (2006) Neural correlates of reward and attention in macaque area LIP. *Neuropsychologia* 44:2411–2420.
- Bichot NP, Schall JD, Thompson KG (1996) Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature* 381:697–699.
- Bisley JW, Goldberg ME (2003) Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299:81–86.
- Bogacz R (2007) Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn Sci* 11:118–125.
- Chen LL, Wise SP (1996) Evolution of directional preferences in the supplementary eye field during acquisition of conditional oculomotor associations. *J Neurosci* 16:3067–3081.
- Dorris MC, Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44:365–378.
- Fanselow MS, Poulos AM (2005) The neuroscience of mammalian associative learning. *Annu Rev Psychol* 56:207–234.
- Fecteau JH, Munoz DP (2005) Correlates of capture of attention and inhibition of return across stages of visual processing. *J Cogn Neurosci* 17:1714–1727.
- Fecteau JH, Bell AH, Munoz DP (2004) Neural correlates of the automatic and goal-driven biases in orienting spatial attention. *J Neurophysiol* 92:1728–1737.
- Ferrera VP, Grinband J (2006) Walk the line: parietal neurons respect category boundaries. *Nat Neurosci* 9:1207–1208.
- Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. *Nature* 443:85–88.
- Gottlieb J, Goldberg ME (1999) Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. *Nat Neurosci* 2:906–912.
- Green DM, Swets JA (1968) Signal detection theory and psychophysics. New York: Wiley.
- Ipata AE, Gee AL, Gottlieb J, Bisley JW, Goldberg ME (2006) LIP responses to a popout stimulus are reduced if it is overtly ignored. *Nat Neurosci* 9:1071–1076.
- Klein RM (2000) Inhibition of return. *Trends Cogn Sci* 4:138–147.
- Kobayashi S, Nomoto K, Watanabe M, Hikosaka O, Schultz W, Sakagami M (2006) Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. *Neuron* 51:861–870.
- Maunsell JH (2004) Neuronal representations of cognitive state: reward or attention? *Trends Cogn Sci* 8:261–265.
- Mruczek RE, Sheinberg DL (2005) Distractor familiarity leads to more efficient visual search for complex stimuli. *Percept Psychophys* 67:1016–1031.
- Mruczek RE, Sheinberg DL (2007) Context familiarity enhances target processing by inferior temporal cortex neurons. *J Neurosci* 27:8533–8545.
- Oristaglio J, Schneider DM, Balan PF, Gottlieb J (2006) Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *J Neurosci* 26:8310–8319.
- Padmala S, Pessoa L (2008) Affective learning enhances visual detection and responses in primary visual cortex. *J Neurosci* 28:6202–6210.
- Phelps EA, Ling S, Carrasco M (2006) Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychol Sci* 17:292–299.
- Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
- Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304:307–310.
- Roesch MR, Olson CR (2007) Neuronal activity related to anticipated reward in frontal cortex: does it represent value or reflect motivation? *Ann N Y Acad Sci* 1121:431–446.
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
- Sereno AB, Maunsell JH (1998) Shape selectivity in primate lateral intraparietal cortex. *Nature* 395:500–503.
- Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415:318–320.
- Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787.
- Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363–375.
- Sutton RS, Barto AG (1998) Reinforcement learning. Cambridge, MA: MIT.
- Vuilleumier P (2005) How brains beware: neural mechanisms of emotional attention. *Trends Cogn Sci* 9:585–594.
- Watanabe M, Cromwell HC, Tremblay L, Hollerman JR, Hikosaka K, Schultz W (2001) Behavioral reactions reflecting differential reward expectations in monkeys. *Exp Brain Res* 140:511–518.

SUPPLEMENTARY MATERIAL

The fast and slow components of learning (Fig. S1)

Since in our task the RC appeared either inside or opposite the RF, we only observed visually-evoked neural responses on half of the trials in which a RC was presented. This feature, combined with fast behavioral learning of newly-learned RC valence precluded detailed analysis of the timecourse of rapid learning in individual neurons. However, we analyzed the fast component of reward selectivity in the population of RC+ preferring cells by pooling data so that, for each consecutive presentation of a RC, we included only those neurons for which that presentation fell inside the RF. Using the same subset of trials, we calculated the latency of reward selectivity for each presentation using ROC analysis as described in the main text (Fig. 4). The results of this analysis are shown in Figure S1 A and B. The neural population developed significant selectivity for newly-learned RC within 6 presentations, a timecourse comparable with the behavioral learn point (8.41 ± 0.54 ; Fig. S1C). After it appeared, selectivity for newly-learned RC remained weaker and longer-latency relative to the selectivity for over-learned RC+. The latency of the reward signal for newly-learned RC (presentations 6-30) was 251 ± 10.86 ms, significantly longer than that for over-learned RC (presentations 1-30 152 ± 4.16 ms; $p < 10^{-9}$).

Independent processing of cues and probes (Fig. S2)

In the probe task (main Fig. 7), monkeys saw two RC – a cue conveying trial outcome opposite the RF, and an uninformative probe inside the RF. We examined whether the two stimuli had interacting effects in neural activity by examining all possible combinations of RC+ or RC- with probe RC+ or RC-. Figure S2 plots the interactions for neural responses for each of the 4 combinations of cue and probe valence. Responses to over-learned probes (right) were significantly stronger for RC+ than for RC- probes, but had no effect of cue valence nor cue/probe interaction. Responses to newly-learned probes were unaffected by both cue and probe valence.

Supplementary figure legends

Fig. S1. The fast component of learning **(A)** Normalized population firing rate (100-900 ms after RC onset) for consecutive RC presentations, for the population of cells with RC+ preference (n = 42 for over-learned, n = 38 for newly-learned). Each data point is based on approximately 75% of the neurons for which a given presentation was in the RF. The data were binned with a moving window of 5 presentations and step size of 1 presentation. Blue stars denote a significant difference between firing rates between RC+ and RC-. **(B)** Latency of neural reward selectivity as a function of presentation number. Each row shows the development of reward selectivity for one RC presentation (binned as in panel A). Magenta stars show the latency relative to RC onset (determined as the first of 4 consecutive bins for which the ROC was significantly above 0.5 ($p < 0.01$, permutation test)). **(C)** Frequency distribution of behavioral learn points for over-learned

(green) and newly-learned (goldenrod) RC. For over-learned trials, the vast majority of learn points were on the first presentation, and the average learn point of 2 is due to a small number of sessions in which discriminatory licking appeared at slightly later time points.

Figure S2 Dependence of neural activity on the valence of informative and non-informative RC on probe trials. The mean and SEM of neural responses to the probes (measured in the 300 ms following the visual latency of each neuron) are plotted. Two-way ANOVAs with RC valence and probe valence as main factors were used to calculate p-values.

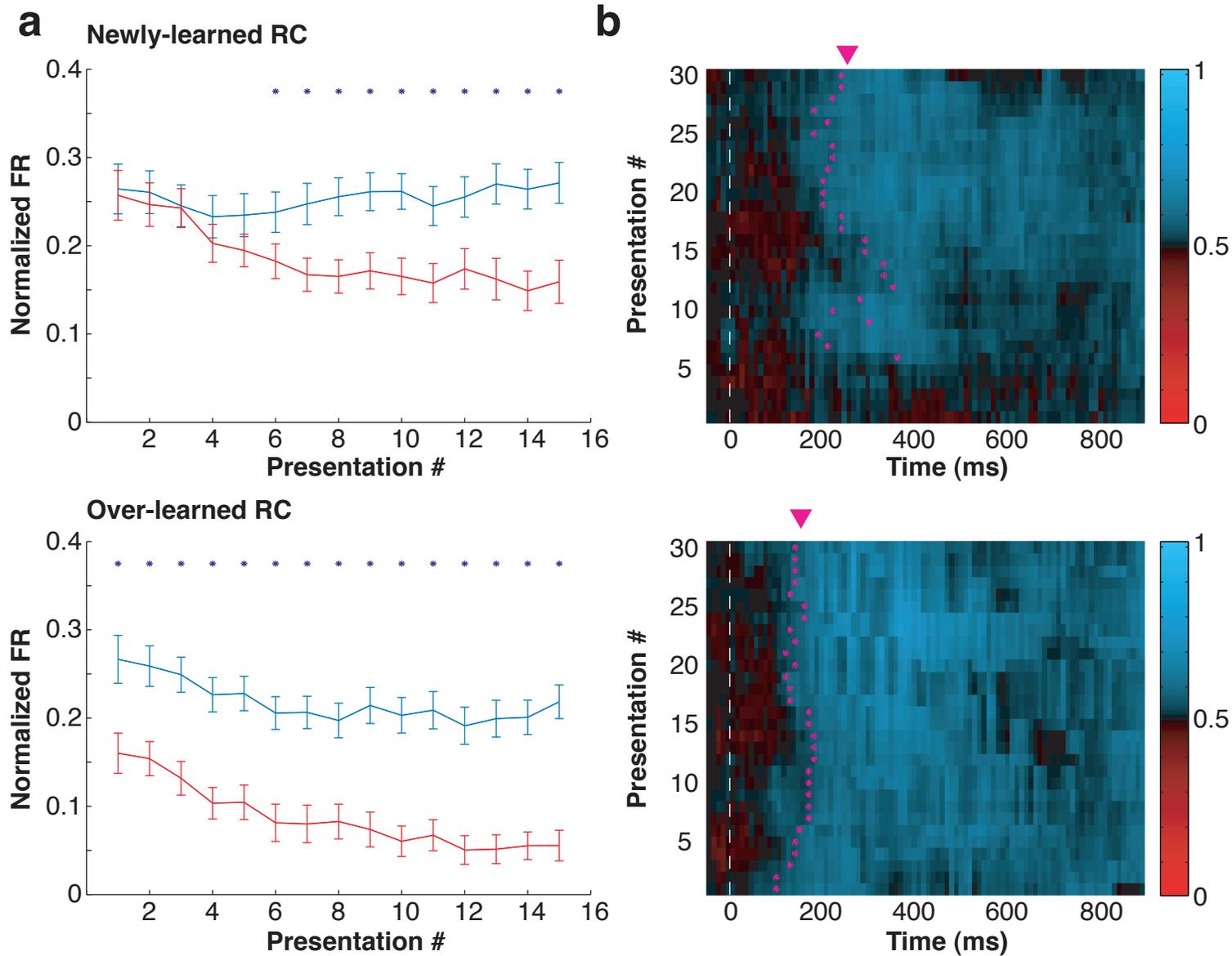
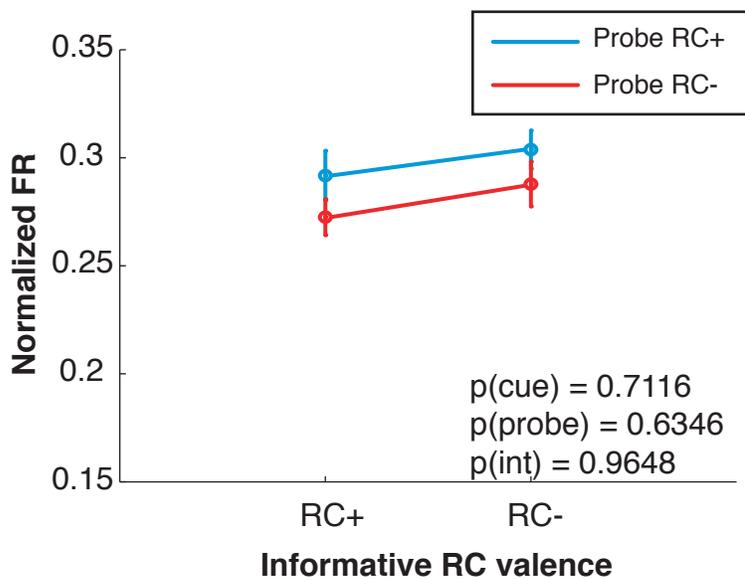


Figure S1, Peck et al.

Newly-learned Probes



Over-learned Probes

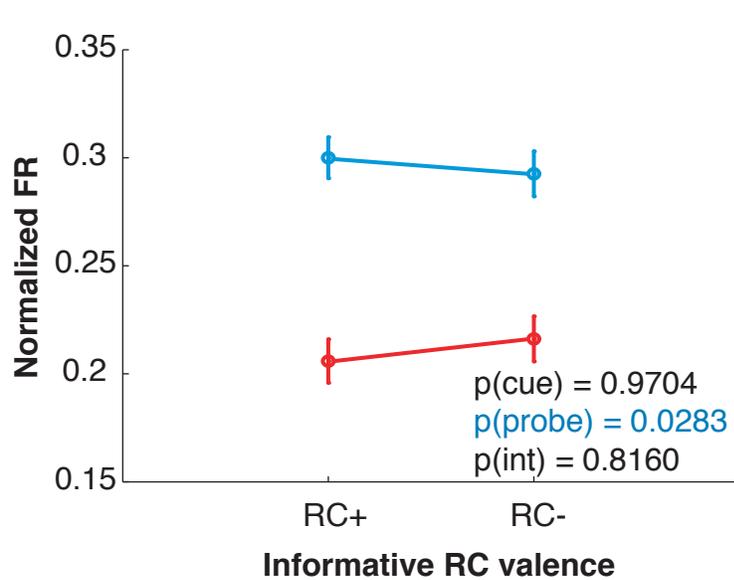


Figure S2, Peck et al.