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.

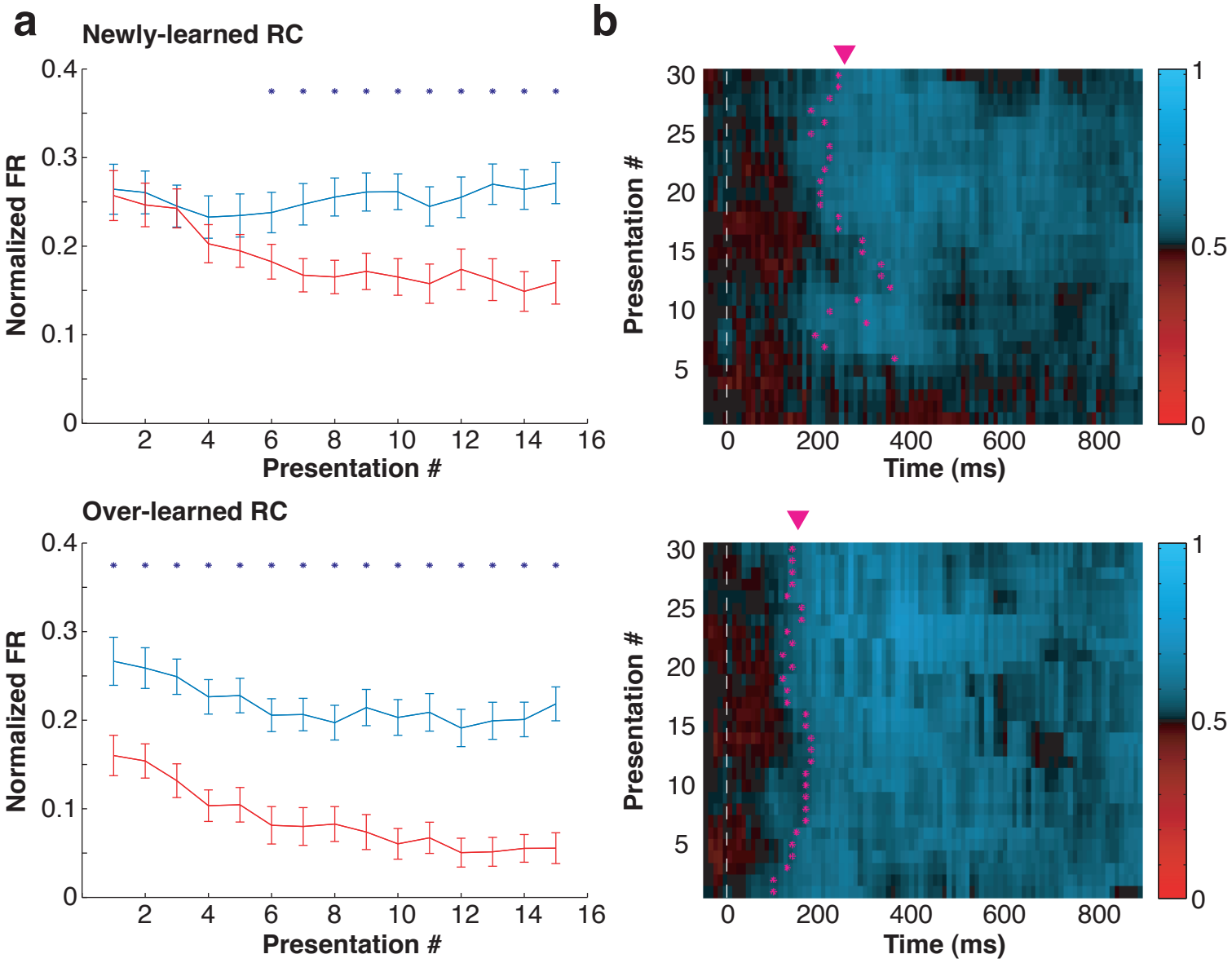

Figure S1, Peck et al.

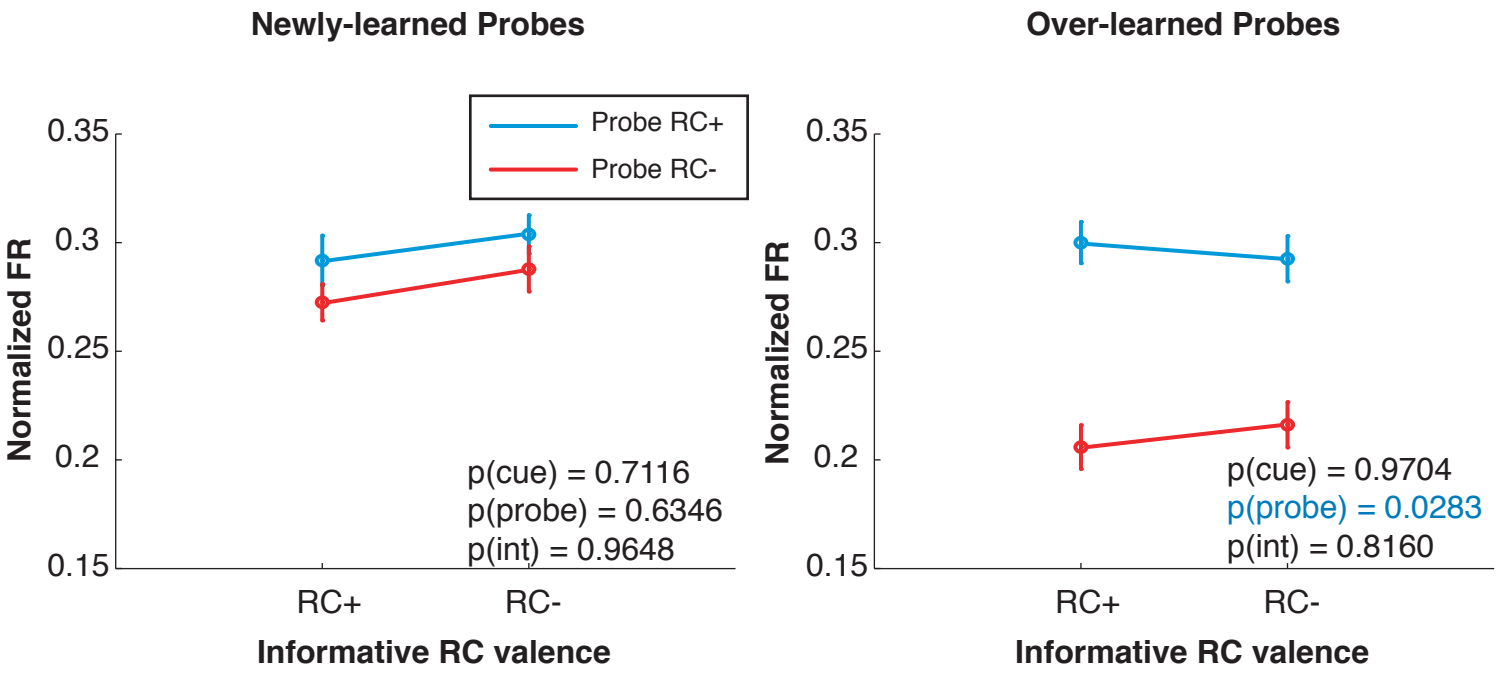

Figure S2, Peck et al.