Roles of the lateral habenula and anterior cingulate cortex in negative outcome monitoring and behavioral adjustment in nonhuman primates

Takashi Kawai [1, 2, 3, 5], Hiroshi Yamada [3, 4], Nobuya Sato [2], Masahiko Takada [1], Masayuki Matsumoto [1, 3, 4]

[1] *Systems Neuroscience Section, Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, Inuyama, Aichi 484-8506, Japan*

[2] *Department of Psychological Science, Kwansei Gakuin University, Nishinomiya, Hyogo 662-8501, Japan*

[3] *Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan*

[4] *Division of Biomedical Science, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan*

[5] *Research Fellow of the Japan Society for the Promotion of Science*

Correspondence should be addressed to M.M. (mmatsumoto@md.tsukuba.ac.jp)

SUMMARY

Animals monitor the outcome of their choice and adjust subsequent choice behavior using the outcome information. Together with the anterior cingulate cortex (ACC), the lateral habenula (LHb) recently attracts attention for its crucial role in monitoring negative outcome. To investigate their contributions to subsequent behavioral adjustment, we recorded single-unit activity from the LHb and ACC in monkeys performing a reversal learning task. The monkey was required to shift a previous choice to the alternative if the choice had been repeatedly unrewarded in past trials. We found that ACC neurons stored outcome information in several past trials whereas LHb neurons detected the ongoing negative outcome with shorter latencies. ACC neurons, but not LHb neurons, signaled a behavioral shift in the next trial. Our findings suggest that, although both the LHb and the ACC represent signals associated with negative outcome, these structures contribute to subsequent behavioral adjustment in different ways.

INTRODUCTION

To choose an appropriate action, animals monitor the outcome of their choice and adjust subsequent choice behavior using the outcome information. Previous studies have identified a number of brain structures that participate in outcome monitoring (Amiez et al., 2005; Hayden et al., 2008; Hollerman and Schultz, 1998; Holroyd et al., 2004; Lau and Glimcher, 2007; O'Doherty et al., 2001; Quilodran et al., 2008; Schoenbaum et al., 1998; Schoenbaum and Setlow, 2003; Stuphorn et al., 2000; Ullsperger and von Cramon, 2003; Wirth et al., 2009). These structures would be responsible for subsequent behavioral adjustment by signaling outcome information, although whether and how they cooperate to achieve adjustment remains to be determined.

Above all, it is generally accepted that the anterior cingulate cortex (ACC) plays important roles in monitoring and adjustment. Human event-related potential (ERP) and functional magnetic resonance imaging (fMRI) studies have shown that the ACC is activated when subjects receive negative feedback following inappropriate behavioral response (Carter et al., 1998; Gehring and Willoughby, 2002; Kiehl et al., 2000). Electrophysiological studies in awake animals have also reported that neurons in the ACC are activated by negative outcomes such as reward omission (Ito et al., 2003; Matsumoto et al., 2007; Niki and Watanabe, 1979; Quilodran et al., 2008) . Since its afferents and efferents are widely distributed in limbic and motor territories of the brain (Dum and Strick, 1991; Kunishio and Haber, 1994; Morecraft and Van Hoesen, 1998; Pandya et al., 1981; Vogt and Pandya, 1987), the ACC can function as an interface between the reward and the motor systems (Paus, 2001). Thus, the ACC is in a good position to integrate outcome information and motor commands, and thereby to

influence choice behavior using the information about past outcomes. Consistent with this view, lesioning or inactivation of the ACC impairs the ability of animals and humans to adjust choice behavior after negative outcome experiences (Shima and Tanji, 1998; Williams et al., 2004).

In a separate line of research, on the other hand, another brain structure called the lateral habenula (LHb) has recently attracted much attention for its crucial role in monitoring negative outcome. Most neurons in the LHb are activated by negative outcomes including reward omission, aversive stimulation, and cues predicting them (Gao et al., 1996; Matsumoto and Hikosaka, 2007, 2009a) . The LHb signal is used to inhibit the activity of midbrain dopamine neurons (Christoph et al., 1986; Ji and Shepard, 2007; Matsumoto and Hikosaka, 2007), a center of the brain's reward system (Schultz, 1998), which in turn project to the ACC (Miller et al., 2009; Williams and Goldman-Rakic, 1998). Since the LHb receives direct projections from the ACC (Chiba et al., 2001), these two structures can communicate with each other by their reciprocal relationship. It is therefore possible that the LHb and ACC cooperate to monitor negative outcome and to adjust subsequent choice behavior. Indeed, using optogenetic and pharmacological techniques, Lammel et al. (2012) reported that manipulation of signals transmitted from the LHb to the medial prefrontal cortex via dopamine neurons alters an avoidance behavior in rodents. However, despite recent advances in understanding the function of the LHb-ACC network, little is known about what signals are shared and what roles are divided between the two structures.

In order to investigate the roles of the LHb and ACC in behavioral adjustment, we recorded their single-unit activities in monkeys performing a reversal learning task. The monkey was required to store information about negative outcomes in several past

trials and needed to use that information to adjust choice behavior. We found that the role of the LHb is more oriented to quickly detect the negative outcome in the ongoing trial, while the ACC plays more crucial roles in storing the information about past negative experiences and in signaling behavioral adjustment in the next choice trial.

RESULTS

Reversal Learning Task and Behavioral Performance

We trained two monkeys (monkeys A and E) to perform a reversal learning task (**Figure 1A**). Each trial started with the presentation of a fixation point. While the monkey was fixating the point, two saccadic targets were presented on both the left and right sides of the point. The monkey was required to choose one of the targets with a saccade. Choosing one target was followed by a liquid reward with 50% probability, while choosing the other was followed by no reward. The reward-position contingency was fixed for a block of 20 to 40 trials, and then reversed for the next block without any instruction. The monkey had to adjust choice behavior to maximize reward gain by trial-and-error. Since the occurrence of the reward was probabilistic only for one target (not for the other), we referred to this task as a unilateral-probabilistic reversal learning task.

The monkey's choice behavior was influenced by outcome experiences in several past trials. **Figure 1B** shows choices and consequent outcomes during an example session in monkey A. Since the reward was delivered in the probabilistic manner, the monkey often failed to obtain it even by choosing a target that was associated with the reward. The monkey tended to continue choosing the same target (i.e., "stay" in **Figure 1B**) once the choice had been rewarded in a past trial. If choosing the same target had

been repeatedly unrewarded in past trials, the monkey tended to change the choice to the alternative (i.e., "shift" in **Figure 1B**). Thus, the monkey decided to stay (choosing the same target) or to shift (choosing the alternative) based on outcome experiences in several past trials. On average, the two monkeys quickly adjusted choice behavior to choose a target that was associated with 50% reward gain (**Figure 1C**).

To statistically analyze the effect of the outcomes on choice behavior, we first calculated the probability that the monkeys shifted a previous choice to the alternative after receiving the reward or no reward in the last trial (**Figure 1D**). The monkeys rarely shifted a previous choice if the choice was followed by the reward (0.1% for monkey A and 2.0% for monkey E), whereas they shifted more often if the choice was followed by no reward. However, the shift probability was only 18.0% for monkey A and 22.3% for monkey E even after receiving no reward in the last trial. Instead, the probability was gradually increased as choosing the same target was repeatedly unrewarded in past trials (correlation between the shift probability and the number of no-reward repetition; monkey A, $r = 0.33$, $p < 0.01$; monkey E, $r = 0.42$, $p < 0.01$) (**Figure 1E**). This suggests that the monkeys decided to shift a previous choice to the alternative by accumulating negative outcome experiences in several past trials.

Responses of LHb and ACC Neurons to the Positive and Negative Outcomes

While the monkeys were performing the reversal learning task, we recorded single-unit activity from 62 LHb neurons (41 in monkey A and 21 in monkey E) and 359 ACC neurons (256 in monkey A and 103 in monkey E) (see **Figure S1** online for histology). As reported in previous studies (Ito et al., 2003; Matsumoto and Hikosaka, 2007, 2009a; Matsumoto et al., 2007; Niki and Watanabe, 1979; Quilodran et al., 2008;

Williams et al., 2004), we found that many of the recorded neurons in both structures were strongly activated by the no-reward outcome (see **Figure 2A** for example neurons recorded in the LHb and ACC). We classified neurons showing a significantly stronger activation by the negative outcome (i.e., no reward) than by the positive outcome (i.e., the reward) as "negative-outcome type", and neurons showing a significantly stronger activation by the positive outcome than by the negative outcome as "positive-outcome type" ($p < 0.05$, Wilcoxon rank-sum test). In the LHb, most of the 62 neurons (49/62, 79%) were classified as the negative-outcome type, while only 4 neurons (6%) were classified as the positive-outcome type (**Figure 2B**). In the ACC, 98 of the 359 neurons (27%) were classified as the negative-outcome type, and 75 neurons (21%) were classified as the positive-outcome type. The proportion of the negative- and positive-outcome type neurons was significantly different from the chance level of 0.5 in the LHb ($p < 0.01$, binominal test) but not in the ACC ($p = 0.09$, binominal test). These data indicate that negative-outcome type neurons were dominant in the LHb and both types of neurons were equally observed in the ACC.

As a population, the outcome-dependent modulation was phasic and time-locked to outcome onset in the LHb, whereas it was relatively tonic and not time-locked in the ACC (**Figure 2C**). The modulation occurred with a significantly shorter latency in the LHb than in the ACC (LHb, negative-outcome type, mean ± SD = 195.2 ± 104.6 ms; ACC, negative-outcome type, mean ± SD = 460.0 ± 256.2 ms, positive-outcome type, mean ± SD = 411.5 ± 217.7 ms; $p < 0.01$, Wilcoxon rank-sum test), suggesting that LHb neurons more quickly detect the outcome in the ongoing trial compared with ACC neurons. The duration of the modulation, which was determined as the width at the half-height of the modulation (see EXPERIMENTAL PROCEDURES for details), was significantly

shorter in the LHb than in the ACC (LHb, negative-outcome type, mean ± SD = 423.0 ± 271.0 ms; ACC, negative-outcome type, mean ± SD = 868.4 ± 606.2 ms, positive-outcome type, mean ± SD = 827.9 ± 565.7 ms; $p < 0.01$, Wilcoxon rank-sum test). These data indicate that the outcome-dependent modulations of LHb and ACC neurons have different temporal features.

Effect of Past Outcome on No-Reward Evoked Response

We observed that the monkeys shifted their choice after receiving no reward (**Figure 1D**). Since we aimed at identifying neuronal signals that regulate the behavioral adjustment induced by the negative outcome, we focused on the responses of LHb and ACC neurons evoked by the negative outcome in the following analyses.

In our task, the monkeys decided to shift the current choice to the alternative in the next trial based not only on the negative outcome in the ongoing trial but also on outcome experiences in past trials (**Figure 1E**). To investigate the neural basis of the behavioral adjustment, we next examined the effect of past outcome experiences on neuronal activity. We first analyzed the effect of a single past outcome that had been obtained in the previous trial (**Figure 3**).

Figure 3A indicates the no-reward evoked responses of three example neurons that were classified as the negative-outcome type in the LHb, and the negative- and positive-outcome types in the ACC (we excluded the positive-outcome type in the LHb from the following analyses because this type constituted a very small population). The negative-outcome LHb neuron did not change its response to the current negative outcome dependent on whether the monkey had obtained the reward (past-rewarded condition) or no reward (past-unrewarded condition) in the previous trial. On the other

hand, the negative-outcome ACC neuron showed a stronger no-reward evoked response in the past-unrewarded condition, whereas the positive-outcome ACC neuron exhibited a stronger response in the past-rewarded condition. It should be noted here that we analyzed the data in which the monkey continued choosing the same target from the previous to current trials. Thus, the negative- and positive-outcome ACC neurons increased or decreased their no-reward evoked responses, respectively, if choosing the same target had been unrewarded in the previous trial.

To quantify the effect of a single past outcome on the no-reward evoked response for each neuron, we calculated receiver operating characteristic (ROC) value for discriminating the no-reward evoked response between the past-rewarded and the past-unrewarded conditions (**Figure 3B**). In the LHb, only a small proportion of the negative-outcome type neurons (5/49, 10%) exhibited a significant difference in the no-reward evoked response between the two conditions ($p < 0.05$, Wilcoxon rank-sum test). On the other hand, a larger proportion of the ACC neurons (negative-outcome type, 25/98, 26%; positive-outcome type, 15/75, 20%) exhibited a significant difference in the no-reward evoked response between the two conditions ($p < 0.05$, Wilcoxon rank-sum test). The proportion of neurons with a differential no-reward evoked response between the conditions was significantly larger than expected by chance, which was calculated by Monte Carlo analysis, in the ACC (negative-outcome type, $p < 0.01$; positive-outcome type, $p < 0.01$) but not in the LHb (negative-outcome type, $p = 0.18$) (**Figure 3C**). Moreover, the proportion was significantly larger in the negative-outcome ACC neurons than in the negative-outcome LHb neurons ($p < 0.05$, Fisher's exact test), though the difference failed to achieve a significant level between the positive-outcome ACC neurons and the negative-outcome LHb neurons ($p = 0.21$, Fisher's exact test) (**Figure**

3C). On average, the ROC value was not significantly different from 0.5 in either the LHb (negative-outcome type, $p = 0.24$, Wilcoxon signed-rank test) or the ACC (negative-outcome type, $p = 0.14$; positive-outcome type, $p = 0.20$, Wilcoxon signed-rank test) (**Figure 3B**), indicating that these neurons did not show a significant bias in the modulation direction induced by the previous outcome.

Effect of Negative Outcome Repetition in Past Trials on No-Reward Evoked Response

The above data suggest that, compared with the LHb, the ACC (the negative-outcome type, especially) more consistently retains the information about a single past outcome in the previous trial. We next examined whether neuronal activity was influenced by outcome experiences in several past trials. At the behavioral level, the monkeys more frequently shifted their choice to the alternative as choosing the same target was repeatedly unrewarded in past trials (**Figure 1E**). Here we analyzed how neuronal activity was modulated by the repetition of the negative outcome (**Figure 4**).

Figure 4A indicates the no-reward evoked responses of three example neurons that were classified as the negative-outcome type in the LHb, and the negative- and positive-outcome types in the ACC. Their activities are shown for trials before which choosing the same target was unrewarded in one, two, three or four consecutive trial(s). The negative-outcome LHb neuron did not change its response to the current negative outcome dependent on the repetition of the negative outcome in past trials. On the other hand, the negative-outcome ACC neuron increased the response as choosing the same target was repeatedly unrewarded in past trials, whereas the positive-outcome ACC neuron decreased the response. Thus, these ACC neurons accumulated negative

outcome experiences in the opposite directions.

To statistically analyze the relationship between neuronal activity and no-reward repetition, we calculated the correlation coefficient between the magnitude of the no-reward evoked response and the number of no-reward repetition (**Figure 4B**). In the LHb, only a small proportion of the negative-outcome type neurons (5/49, 10%) exhibited a significant correlation ($p < 0.05$). On the other hand, a larger proportion of the ACC neurons (negative-outcome type, 27/98, 28%; positive-outcome type, 12/75, 16%) exhibited a significant correlation ($p < 0.05$). On average, the no-reward evoked responses of the neurons with a significant positive or negative correlation gradually increased or decreased, respectively, as choosing the same target was repeatedly unrewarded in both the LHb and ACC (**Figure 4C**). The proportion of neurons with a significant correlation was significantly larger than expected by chance, which was calculated by Monte Carlo analysis, in the ACC (negative-outcome type, $p < 0.01$; positive-outcome type, $p < 0.01$) but not in the LHb (negative-outcome type, $p = 0.24$) (**Figure 4D**). Moreover, the proportion was significantly larger in the negative-outcome ACC neurons than in the negative-outcome LHb neurons ($p < 0.05$, Fisher's exact test), though the difference failed to achieve a significant level between the positive-outcome ACC neurons and the negative-outcome LHb neurons ($p = 0.43$, Fisher's exact test) (**Figure 4D**). These data suggest that, compared with the LHb, the negative-outcome type neurons in the ACC more largely contribute to the accumulation of negative outcome experiences in past trials. On average, the correlation coefficient was not significantly different from 0 in either the LHb (negative-outcome type, $p = 0.93$, Wilcoxon signed-rank test) or the ACC (negative-outcome type, $p = 0.10$; positive-outcome type, $p = 0.10$, Wilcoxon signed-rank test) (**Figure 4B**), indicating that

these neurons did not show a significant bias in the modulation direction induced by the repetition of the negative outcome in past trials.

No-Reward Evoked Response Predicts Subsequent Choice Behavior

We found that ACC neurons, especially the negative-outcome type, more consistently accumulated the effect of negative outcome experiences on their no-reward evoked responses (**Figure 4**), and that this accumulation effect seemed parallel with the behavioral shift from the previous choice to the alternative (see **Figures 1E** and **4C**). We next tested the hypothesis that the neuronal activity accumulating negative outcome experiences triggered the behavioral shift. We examined whether the no-reward related activity predicted that the monkey would shift the current choice to the alternative (next shift) or would stay with the current choice in the next trial (next stay) (**Figure 5**).

Figure 5A indicates the no-reward evoked responses of three example neurons that were classified as the negative-outcome type in the LHb, and the negative- and positive-outcome types in the ACC. The negative-outcome LHb neuron did not show a clear modulation reflecting the choice behavior in the next trial. On the other hand, the negative-outcome ACC neuron increased its activity in next shift trials compared with next stay trials, whereas the positive-outcome ACC neuron decreased its activity in next shift trials. Thus, these ACC neurons signaled the occurrence of the behavioral shift in the next trial in the opposite directions.

To quantify how neuronal activity predicted the choice behavior in the next trial, we calculated ROC value for discriminating the no-reward evoked response between next shift and next stay trials for each neuron (**Figure 5B**). In the LHb, only a small proportion of the negative-outcome type neurons (3/49, 6%) exhibited a significant

difference in the no-reward evoked response between next shift and next stay trials ($p < 0.05$, Wilcoxon rank-sum test). On the other hand, a larger proportion of the ACC neurons (negative-outcome type, 20/98, 20%; positive-outcome type, 9/75, 12%) exhibited a significant difference in the no-reward evoked response between these trials ($p < 0.05$, Wilcoxon rank-sum test). The proportion of neurons with a differential no-reward evoked response between the trials was significantly larger than expected by change, which was calculated by Monte Carlo analysis, in the ACC (negative-outcome type, $p < 0.01$; positive-outcome type, $p < 0.05$) but not in the LHb (negative-outcome type, $p = 0.83$) (**Figure 5C**). Moreover, the proportion was significantly larger in the negative-outcome ACC neurons than in the negative-outcome LHb neurons ($p < 0.05$, Fisher's exact test), though the difference failed to achieve a significant level between the positive-outcome ACC neurons and the negative-outcome LHb neurons ($p = 0.36$, Fisher's exact test) (**Figure 5C**). These data suggest that the no-reward evoked response was more strongly modulated in the ACC, especially in the negative-outcome type, depending on whether the monkey would take a shift or stay strategy in the next trial. On average, the ROC value was not significantly different from 0.5 in either the LHb (negative-outcome type, $p = 0.59$, Wilcoxon signed-rank test) or the ACC (negative-outcome type, $p = 0.62$; positive-outcome type, $p = 1.00$, Wilcoxon signed-rank test) (**Figure 5B**), indicating that these neurons did not show a significant bias in the modulation direction predicting the choice behavior in the next trial.

Past Outcome Experience and Subsequent Choice Behavior Individually Influence Neuronal Activity

We so far separately analyzed the effects of past outcome experiences and

subsequent choice behavior on neuronal activity. However, the monkeys were more likely to shift a previous choice to the alternative if the choice had been repeatedly unrewarded in past trials, indicating that past outcome experiences influenced subsequent choice behavior and thereby these two factors were not independent. Thus, the effect of one factor (e.g., subsequent choice behavior) on neuronal activity may be accounted for by the effect of the other factor (e.g., past outcome experiences). Contrary to this assumption, we found that past outcome experiences and subsequent choice behavior did not necessarily influence neuronal activity in the same fashion; some neurons showed only the effect of past outcome experiences whereas some other neurons showed only the effect of subsequent choice behavior (see **Figure 6A** and **B** for ACC neuron examples that were classified as the negative-outcome type).

In order to systematically evaluate the effects of each factor at the single-neuron level, we conducted multiple linear regression analysis that provides the regression coefficients of the effects of past outcome experiences and subsequent choice behavior (see EXPERIMENTAL PROCEDURES for details). We compared the coefficients of a single past outcome in the previous trial and choice behavior in the next trial (**Figure 6C**). In the LHb, 6 of the 49 negative-outcome type neurons showed a significant coefficient for either the past outcome or the choice behavior ($p < 0.05$), but none of them showed significance for both. In the ACC, 31 of the 98 negative-outcome type neurons and 15 of the 75 positive-outcome type neurons showed a significant coefficient for either one ($p < 0.05$), but only two of them (negative-outcome type, 2/31; positive-outcome type, 0/15) exhibited significance for both. Moreover, the absolute values of the coefficients of the past outcome and choice behavior were not significantly correlated with each other in either the LHb (negative-outcome type, $r = 0.05$, $p = 0.73$)

or the ACC (negative-outcome type, r = -0.03, p = 0.75; positive-outcome type, r = -0.11, p = 0.35), indicating that the magnitudes of the effects of the past outcome and choice behavior on neuronal activity were independent across neurons. The coefficients themselves (i.e., non-absolute coefficients) were also insignificantly correlated with each other (LHb, negative-outcome type, r = -0.28, p = 0.051; ACC, negative-outcome type, r = -0.07, p = 0.52, positive-outcome type, r = -0.07, p = 0.56). These data suggest that past outcome experiences and subsequent choice behavior were not necessarily represented by the same subgroup of neurons.

Comparison between the Dorsal and the Ventral Banks of the Anterior Cingulate Sulcus

The above observation may imply that different subgroups of ACC neurons are involved in accumulating past outcome experiences and adjusting subsequent choice behavior. Especially, since the ACC is subdivided into the dorsal and ventral banks of the anterior cingulate sulcus (dorsal and ventral ACC, respectively) that are structurally and functionally distinct (Bush et al., 2000; Paus, 2001), neurons in each subdivision might contribute to either accumulating past outcome experiences or adjusting subsequent choice behavior. To test this possibility, we next reanalyzed our ACC data set separately for dorsal (n = 216) and ventral ACC neurons (n = 143) (**Figure 7**).

We first compared the proportion of negative- and positive-outcome type neurons in the dorsal and ventral ACC (see "Responses of LHb and ACC Neurons to the Positive and Negative Outcomes" for details of the analysis) (**Figure 7A**). The proportion was not significantly different between the two regions (dorsal ACC, negative-outcome type,

66/216, 31%, positive-outcome type, 46/216, 21%; ventral ACC, negative-outcome type, 32/143, 22%, positive-outcome type, 29/143, 20%, p = 0.43, Fisher's exact test). We then compared the proportion of neurons in the dorsal and ventral ACC that showed a significant effect of a single past outcome in the previous trial (see "Effect of Past Outcome on No-Reward Evoked Response" for details of the analysis) (**Figure 7B**), and the proportion of neurons in the two regions that showed a significant effect of negative outcome repetition in past trials (see "Effect of Negative Outcome Repetition in Past Trials on No-Reward Evoked Response" for details of the analysis) (**Figure 7C**). Neither of these proportions was significantly different between the dorsal and the ventral ACC (single past outcome, dorsal ACC, 24/112, 21%, ventral ACC, 16/61, 26%, p = 0.57; negative outcome repetition, dorsal ACC, 27/112, 24%, ventral ACC, 12/61, 20%, p = 0.57, Fisher's exact test). Thus, we found no significant difference between the dorsal and the ventral ACC with respect to monitoring ongoing outcome (**Figure 7A**) or storing past outcome information (**Figure 7B** and **C**).

Notably, on the other hand, we found a significant difference in the roles of the dorsal and ventral ACC in subsequent choice behavior. We compared the proportion of neurons in the dorsal and ventral ACC that signaled whether the monkey would shift the current choice or stay with the current choice in the next trial (see "No-Reward Evoked Response Predicts Subsequent Choice Behavior" for details of the analysis) (**Figure 7D**). A significantly larger proportion of dorsal ACC neurons (24/112, 21%), than ventral ACC neurons (5/61, 8%), signaled the monkey's subsequent choice behavior ($p < 0.05$, Fisher's exact test). Moreover, the proportion was significantly larger than expected by chance, which was calculated by Monte Carlo analysis, in the dorsal ACC ($p < 0.01$) but not in the ventral ACC ($p = 0.29$). These results suggest that, although the

two subdivisions of the ACC are equally involved in accumulating past outcome experiences, the dorsal ACC contributes more largely to the adjustment of subsequent choice behavior than the ventral ACC.

DISCUSSION

The present study compared the roles of the LHb and ACC in behavioral adjustment after negative outcome experiences. We found that, although LHb and ACC neurons both responded to the negative outcome of a choice, these neurons transmitted different signals that would be instrumental in adjusting subsequent choice behavior in distinct ways.

One of the marked differences between LHb and ACC signals is the temporal feature of neuronal responses. The no-reward evoked activation was phasic and time-locked to the occurrence of the negative outcome in the LHb, whereas it was tonic and its onset timing varied across neurons in the ACC. On average, the no-reward modulation in the ACC was sustained until the beginning of the next trial. Such sustained neuronal activity is also observed in other brain regions (Asaad and Eskandar, 2011; Donahue et al., 2013; Hayden et al., 2008; Histed et al., 2009; Kennerley et al., 2011; Kim et al., 2009; Seo and Lee, 2009; Sul et al., 2010), and has been proposed to influence subsequent choice behavior by maintaining outcome information during an interval between trials (Hayden et al., 2008; Histed et al., 2009). Furthermore, the activity of ACC neurons not only maintained the information about the last outcome, but also stored outcome experiences in several past trials. Consistent with our finding in the ACC, Seo and Lee (2007) reported that the activity of neurons in the dorsal ACC, which at least partly overlaps our recording sites, was modulated by rewards that had

been obtained in past trials. As seen in our monkeys, animals decide what to choose by accumulating past outcome experiences. The ACC would be a major candidate of the neural mechanism underlying this accumulation process.

In contrast to the ACC, we found that the activity of LHb neurons was modulated mainly by the outcome in the ongoing trial. This observation seemingly conflicts with a previous study reporting that the activity of LHb neurons was influenced by outcomes in past trials (Bromberg-Martin et al., 2010). Their study investigated LHb neuron activity using a visually-guided saccade task, but not a choice task, in which monkeys were simply required to make an eye movement to a target. In their task, a reward or no reward was delivered in a pseudorandom order. For instance, the reward probability in a given trial was higher if the monkey had not obtained the reward in the previous trial, and was lower if the monkey had obtained the reward in the previous trial. Therefore, using the information about past outcomes, the monkey was able to estimate more accurately than expected by chance whether the reward would occur in a given trial. In our choice task, on the other hand, the reward was delivered independent of past outcomes so that the monkey was unable to estimate reward delivery using the past outcome information. Thus, the modulation of LHb neuron activity reported in the previous study might reflect the estimation of reward probability calculated from past outcome experiences, but might not represent the crude information as to whether the monkey had obtained the reward in past trials.

Another marked difference between LHb and ACC signals is seen in their relationship with monkey's choice behavior. A larger proportion of ACC neurons signaled whether the monkey would shift the current choice to the alternative in the next trial, whereas only a few LHb neurons signaled the subsequent behavioral shift.

Consistent with our finding in the ACC, Shima and Tanji (1998) reported that neurons in the rostral cingulate motor area, which includes at least a part of our recording sites, were particularly activated when monkeys voluntarily shifted their choice to an alternative that would cause a better consequence. They further reported that pharmacological inactivation thereof impaired the voluntary shift of choice behavior. In contrast to the ACC, we found that LHb neurons rarely encoded the signal associated with subsequent choice behavior. Thus, the LHb is unlikely to participate in the direct control of behavioral adjustment. However, manipulation of habenular activity has been shown to alter behaviors in fishes, rodents, and primates (Agetsuma et al., 2010; Lecourtier et al., 2004; Matsumoto and Hikosaka, 2011; Thornton and Bradbury, 1989), possibly through its strong effect on monoaminergic systems such as dopamine and serotonin (Amat et al., 2001; Amo et al., 2014; Lammel et al., 2012; Li et al., 2011). It is therefore conceivable that the LHb indirectly influences behavioral adjustment through the monoaminergic circuits. This issue is further discussed below.

Instead of storing past outcome experiences or signaling behavioral adjustment, LHb neurons exhibited a phasic and shorter-latency activation that was time-locked to the onset of the negative outcome. Such a phasic response would be suitable to quickly detect the negative outcome in the ongoing trial. Given the fact that the LHb sends direct and indirect projections to regions of midbrain dopamine neurons (Herkenham and Nauta, 1979; Jhou et al., 2009b; Omelchenko et al., 2009) that in turn project to the ACC (Miller et al., 2009; Williams and Goldman-Rakic, 1998), LHb neurons might transmit their signals to the ACC as a source of the outcome information accumulated across trials. In favor of this hypothesis, the dopamine neurons mediating the LHb-ACC circuit also encode signals associated with negative outcomes; these neurons are

inhibited by both of reward omission and aversive stimulation (Matsumoto and Hikosaka, 2009b). Such an inhibitory dopaminergic signal is at least partly caused by inputs from the LHb (Matsumoto and Hikosaka, 2007) via another relay nucleus called the rostromedial tegmental nucleus (Hong et al., 2011; Jhou et al., 2009a). In turn, the inhibitory dopaminergic signal has long been proposed to disinhibit ACC neurons via mesocortical dopaminergic projections, which thereby produces ACC activation associated with negative outcomes (Holroyd and Coles, 2002). Indeed, pharmacological blockade of the dopaminergic transmission reduces the event-related potential evoked by a negative outcome in the ACC (Vezoli and Procyk, 2009). Together with these literatures, our finding might suggest the possibility that the LHb provides the ACC with the information about the ongoing negative outcome via dopamine neurons, and that the ACC stores this outcome information through trials and adjusts subsequent choice behavior using the accumulated outcome information.

It should be mentioned here that, although we found substantial differences between LHb and ACC signals, these differences were significant only for the negative-outcome type neurons in the ACC. Compared with LHb neurons, a significantly larger proportion of the negative-outcome ACC neurons, but not the positive-outcome ACC neurons, signaled past outcome information and subsequent choice behavior. The reason why we observed the significant differences only for the negative-outcome ACC neurons remains unclear. However, the negative-outcome type neurons were defined as neurons that were more strongly activated by the negative outcome than the positive outcome, suggesting that these neurons preferentially signaled the information about the negative outcome. Thus, the negative-outcome type neurons might play more crucial roles in the behavioral adjustment that was associated

with the negative outcome.

As described above, LHb and ACC neurons transmitted different signals that would be instrumental in adjusting subsequent choice behavior in distinct ways. In addition, we also found a notable difference between the subdivisions of the ACC. Although neurons in both the dorsal and the ventral ACC were equally involved in accumulating past outcome experiences, a larger proportion of dorsal ACC neurons, than ventral ACC neurons, signaled whether the monkey would shift the current choice or would stay with the current choice in the next trial. Previous studies have reported that these subdivisions are structurally and functionally differentiated (Bush et al., 2000; Paus, 2001). For instance, Cai and Padoa-Schioppa (2012) recorded single-unit activity from the dorsal and ventral ACC while monkeys were performing an economic choice task. The monkey was required to choose one of two options (i.e., to make a saccade to one of two saccadic targets) that were associated with different values of reward. They found that, although neurons in both the dorsal and the ventral ACC encoded the subjective value of chosen reward, neuronal activity in the dorsal ACC alone was influenced by saccade direction. Thus, neurons in the dorsal ACC represented both the chosen outcome and the movement signals. They proposed that the dorsal ACC constitutes a gateway through which the choice system (or the outcome monitoring system) informs motor systems. This idea seems consistent with our findings that neurons in the dorsal ACC not only accumulated past outcome experiences but also signaled whether the monkey would take a shift or stay strategy in the next trial. Taken together, these studies suggest that the dorsal ACC, but not the ventral ACC, may play a crucial role in transforming past and current outcome information into future choice commands. Since we found that different subgroups of dorsal ACC neurons signaled past outcome

experiences and subsequent choice behavior, the transformation process might be implemented through the local circuit connecting these two subgroups.

Although we have so far focused on the LHb and ACC, another brain structures, the posterior cingulate cortex (PCC), also recently attracts attention for its significant contribution to outcome monitoring and behavioral adjustment (Hayden et al., 2008). As we found in the ACC, Hayden et al. (2008) reported that neurons in the PCC represented outcome experiences in several past trials and predicted subsequent behavioral adjustment. By electrically stimulating the PCC, they also found that monkeys often shifted their previous choice to an alternative in the next choice trial. Since the PCC has a strong reciprocal connection with the ACC (Kobayashi and Amaral, 2003), these cortical areas can communicate with each other. However, how they cooperate to achieve monitoring and adjustment remains unclear. Furthermore, many other brain structures have also been identified to participate in outcome monitoring and/or behavioral adjustment, such as the supplementary eye field (Stuphorn et al., 2000) and the orbitofrontal cortex (O'Doherty et al., 2001; Schoenbaum et al., 1998). In order to understand how the brain adjusts behavior through past negative experiences, future studies are called for to determine what signals are shared and what roles are divided between these brain structures.

In summary, we found that, although both the LHb and the ACC represent signals associated with negative outcome, these brain structures contribute to behavioral adjustment in different ways. Our results suggest that the LHb is suitable to quickly detect the negative outcome in the ongoing trial, whereas the ACC more largely contributes to accumulating past negative experiences and signaling subsequent behavioral adjustment.

EXPERIMENTAL PROCEDURES

Animals

We used two adult macaque monkeys, monkey A (*Macaca fuscata*, male, 10.0 kg) and monkey E (*Macaca mulatta*, male, 9.5 kg), for the present experiments. All procedures for animal care and experimentation were approved by the Institutional Animal Care and Use Committee of Primate Research Institute, Kyoto University (Permission Number: 2010-080) and by the University of Tsukuba Animal Experiment Committee (Permission Number: 12-415).

Behavioral Task

Behavioral task events were controlled by TEMPO system (Reflective Computing). The monkeys sat in a primate chair facing a frontoparallel computer monitor in a sound-attenuated and electrically shielded room. Eye movements were monitored using an infrared eye-tracking system (EYE-TRAC 6, Applied Science Laboratories) by sampling at 240 Hz.

The monkeys performed a reversal learning task (**Figure 1A**). Each trial began with the appearance of a central fixation point (0.5 degree diameter) and the animal was required to fixate the point. After 750 ms of fixation, the fixation point disappeared and two saccadic targets were presented on the left and right sides of the point (0.5 degree diameter, 8 degree eccentricity). The monkey was required to choose one of the targets with a saccade within 1000 ms. After the saccade, the targets were kept on for 750 ms during which the monkey had to keep fixating the chosen target. The completion of each choice (i.e., saccade and following fixation) was signaled by a tone (1 kHz frequency). Choosing one target was followed by a liquid reward with 50% probability, whereas

choosing the other target was not followed by the reward. The reward was delivered simultaneously with the tone. The reward-position contingency was fixed within a block of trials (20 to 30 trials for monkey A, 20 to 40 trials for monkey E), and was reversed in the next block without any external instruction. The probability of the contingency reversal was uniform in 20 to 30 trials in monkey A, and in 20 to 40 trials in monkey E.

Electrophysiology

A plastic head holder and recording chamber were fixed to the skull under general anesthesia and sterile surgical conditions. The recording chamber was placed over the midline of the frontoparietal lobes to be aimed at the LHb and ACC. The head holder and the recording chamber were embedded in dental acrylic resin that covered the top of the skull and were connected to the skull with plastic screws.

Single-unit recordings were performed using tungsten electrodes with impedance of 0.5-2.5 MΩ (Frederick Haer). The electrode was introduced into the brain through a stainless-steel guide tube using an oil-driven micromanipulator (MO-97-S, Narishige). The recording sites were determined using a grid system, which allowed recordings at every 1 mm between penetrations. For finer mapping of neurons, we also used a complementary grid which allowed electrode penetrations between the holes of the original grid.

Electrophysiological signals were amplified, band-pass filtered (100 Hz to 8 kHz; RZ5, Tucker-Davis Technologies) and stored in a computer at the sampling rate of 24.4 kHz. Single-unit potentials were isolated online using a window discrimination software (OpenEx, Tucker-Davis Technologies).

Data Analysis

To analyze neuronal activity, we combined data from the two monkeys because they were qualitatively identical for our major findings. We focused on the modulation of neuronal discharge rate evoked by the positive and negative outcomes (i.e., the reward and no reward). To analyze the modulation, we calculated the discharge rate of LHb neurons during 150 – 600 ms after the onset of the positive (reward and tone) and negative outcomes (tone only), and that of ACC neurons during 250 – 1250 ms after the onset. These time windows were determined on the basis of the averaged activities (see **Figure 2C** bottom). We compared the discharge rate of each neuron between rewarded and unrewarded trials. Neurons showing a significantly stronger activation in unrewarded trials were classified as the "negative-outcome type", whereas those showing a significantly stronger activation in rewarded trials were classified as the "positive-outcome type" ($p < 0.05$, Wilcoxon rank-sum test) (**Figure 2B**). To evaluate the effects of past outcome experiences (**Figures 3** and **4**) and subsequent choice behavior (**Figure 5**) on the no-reward evoked response of each neuron, we used the discharge rate during the same time window.

The peri-stimulus time histograms (PSTHs) (bin width, 10 ms) were smoothed by averaging across a 60 ms sliding window with a 10 ms step.

To visualize the time course of the neuronal modulation by the outcomes for each neuron, we calculated ROC value for discriminating the discharge rate between rewarded and unrewarded trials using a 200 ms test window sliding with a 10 ms step (**Figure 2C** top). Using the sliding ROC value, we further calculated the duration of the neuronal modulation for each neuron. We first determined the maximum peak of the sliding ROC value. The duration of the neuronal modulation was determined as the

width at the half-height of the peak.

The latency of the neuronal modulation by the outcomes was determined, for each neuron, by comparing the discharge rate between rewarded and unrewarded trials. We performed the comparison by Wilcoxon rank-sum test using a 100 ms test window sliding with a 1 ms step. The latency was taken as the midpoint of the first of 20 consecutive test windows, if the first and at least 19 of the 20 windows showed a significant difference ($p < 0.05$, Wilcoxon rank-sum test).

We performed a Monte Carlo analysis to test whether the proportion of neurons showing a significant effect of a single past outcome in the previous trial (**Figures 3C** and **7B**), the proportion of neurons showing a significant effect of no-reward repetition in past trials (**Figures 4D** and **7C**), and the proportion of neurons showing a significant effect of subsequent choice behavior (**Figures 5C** and **7D**) were larger than expected by chance. For each neuron, we shuffled the firing rate of each trial and assigned it to another trial at random to form a new data set. The new data sets of all neurons were combined, and the proportion of neurons showing a significant effect was calculated. We repeated such shuffling and calculation 1000 times, and obtained the chance-level distribution of the proportion. We compared the original proportion of significant neurons with this distribution.

In order to evaluate the effects of outcomes (i.e., the reward or no reward) in past trials and the choice behavior (i.e., stay or shift) in the next trial on the no-reward evoked response of each neuron, we conducted a multiple regression analysis (Bayer and Glimcher, 2005; Hayden et al., 2008). This analysis provided a set of weights (ß values) representing the magnitudes of the effects of the past outcomes and subsequent choice behavior, in the following form:

$$F_i = \beta_0 + \beta_{i+1}C_{i+1} + \beta_{i-1}R_{i-1}C_{i-1} + \beta_{i-2}R_{i-2}C_{i-2} + \cdots + \beta_{i-5}R_{i-5}C_{i-5}$$

where F_i indicates the z-scored discharge rate of each neuron in the response to the no-reward outcome in the ith trial calculated using the time window described above. R_i indicates the past outcome in the ith trial and is set to 1 or -1 for the reward and no reward, respectively. C_i indicates the choice behavior in the ith trial and set to 1 or -1 for choosing the same target (stay) and changing the choice to the alternative (shift), respectively.

Histology

At the end of the recording session in monkey A, we selected representative locations of electrode penetration into the LHb and ACC, and made electrolytic microlesions at each recording site (12 μA and 35 s). Then monkey A was deeply anaesthetized with pentobarbital sodium, and perfused with 10% formaldehyde. The brain was blocked and equilibrated with 30% sucrose. Frozen sections were cut every 60 μm in the coronal plane. The sections were stained with cresyl violet.

SUPPLEMENTAL INFORMATION

Supplemental information includes 1 figure and can be found with this article online.

AUTHOR CONTRIBUTIONS

T.K. performed the experiments and analyzed the data. T.K., H.Y., N.S., M.T., and M.M. discussed the results and wrote the manuscript. M.M organized this project.

ACKNOWLEDGEMENTS

We thank T. Kuroda and I. Nagumo for technical assistance, and K. Bunzui for animal care. This research was supported by Funding Program for Next Generation World-Leading Researchers (LS074) to M.M. from Cabinet Office, Government of Japan; Grants-in-Aid for Scientific Research (26120707) to M.M. from the Ministry of Education, Science, Sports, Culture, and Technology of Japan; the Takeda Science Foundation to M.M.; the Inamori Foundation to M.M.; and the Uehara Memorial Foundation to M.M. This research was conducted as part of the cooperative research program at the Primate Research Institute, Kyoto University.

REFERENCES

Agetsuma, M., Aizawa, H., Aoki, T., Nakayama, R., Takahoko, M., Goto, M., Sassa, T., Amo, R., Shiraki, T., Kawakami, K.*, et al.* (2010). The habenula is crucial for experience-dependent modification of fear responses in zebrafish. Nat Neurosci *13*, 1354-1356.

Amat, J., Sparks, P.D., Matus-Amat, P., Griggs, J., Watkins, L.R., and Maier, S.F. (2001). The role of the habenular complex in the elevation of dorsal raphe nucleus serotonin and the changes in the behavioral responses produced by uncontrollable stress. Brain Res *917*, 118-126.

Amiez, C., Joseph, J.P., and Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. Eur J Neurosci *21*, 3447-3452.

Amo, R., Fredes, F., Kinoshita, M., Aoki, R., Aizawa, H., Agetsuma, M., Aoki, T., Shiraki, T., Kakinuma, H., Matsuda, M.*, et al.* (2014). The habenulo-raphe serotonergic circuit encodes an aversive expectation value essential for adaptive active avoidance of danger. Neuron *84*, 1034-1048.

Asaad, W.F., and Eskandar, E.N. (2011). Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. J Neurosci *31*, 17772-17787.

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron *47*, 129-141.

Bromberg-Martin, E.S., Matsumoto, M., Nakahara, H., and Hikosaka, O. (2010). Multiple timescales of memory in lateral habenula and dopamine neurons. Neuron *67*, 499-510.

Bush, G., Luu, P., and Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci *4*, 215-222.

Cai, X., and Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. J Neurosci *32*, 3791-3808.

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., and Cohen, J.D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. Science *280*, 747-749.

Chiba, T., Kayahara, T., and Nakano, K. (2001). Efferent projections of infralimbic and prelimbic areas of the medial prefrontal cortex in the Japanese monkey, Macaca fuscata. Brain Res *888*, 83-101.

Christoph, G.R., Leonzio, R.J., and Wilcox, K.S. (1986). Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. J Neurosci *6*, 613-619.

Donahue, C.H., Seo, H., and Lee, D. (2013). Cortical signals for rewarded actions and strategic exploration. Neuron *80*, 223-234.

Dum, R.P., and Strick, P.L. (1991). The origin of corticospinal projections from the premotor areas in the frontal lobe. J Neurosci *11*, 667-689.

Gao, D.M., Hoffman, D., and Benabid, A.L. (1996). Simultaneous recording of spontaneous activities and nociceptive responses from neurons in the pars compacta of substantia nigra and in the lateral habenula. Eur J Neurosci *8*, 1474-1478.

Gehring, W.J., and Willoughby, A.R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. Science *295*, 2279-2282.

Hayden, B.Y., Nair, A.C., McCoy, A.N., and Platt, M.L. (2008). Posterior cingulate cortex

mediates outcome-contingent allocation of behavior. Neuron *60*, 19-25.

Herkenham, M., and Nauta, W.J. (1979). Efferent connections of the habenular nuclei in the rat. J Comp Neurol *187*, 19-47.

Histed, M.H., Pasupathy, A., and Miller, E.K. (2009). Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. Neuron *63*, 244-253.

Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci *1*, 304-309.

Holroyd, C.B., and Coles, M.G.H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychological Review *109*, 679-709.

Holroyd, C.B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R.B., Coles, M.G., and Cohen, J.D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. Nat Neurosci *7*, 497-498.

Hong, S., Jhou, T.C., Smith, M., Saleem, K.S., and Hikosaka, O. (2011). Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. J Neurosci *31*, 11457-11471.

Ito, S., Stuphorn, V., Brown, J.W., and Schall, J.D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science *302*, 120-122.

Jhou, T.C., Fields, H.L., Baxter, M.G., Saper, C.B., and Holland, P.C. (2009a). The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. Neuron *61*, 786-800.

Jhou, T.C., Geisler, S., Marinelli, M., Degarmo, B.A., and Zahm, D.S. (2009b). The

mesopontine rostromedial tegmental nucleus: A structure targeted by the lateral habenula that projects to the ventral tegmental area of Tsai and substantia nigra compacta. J Comp Neurol *513*, 566-596.

Ji, H., and Shepard, P.D. (2007). Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. J Neurosci *27*, 6923-6930.

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

Kiehl, K.A., Liddle, P.F., and Hopfinger, J.B. (2000). Error processing and the rostral anterior cingulate: an event-related fMRI study. Psychophysiology *37*, 216-223.

Kim, H., Sul, J.H., Huh, N., Lee, D., and Jung, M.W. (2009). Role of striatum in updating values of chosen actions. J Neurosci *29*, 14701-14712.

Kobayashi, Y., and Amaral, D.G. (2003). Macaque monkey retrosplenial cortex: II. Cortical afferents. J Comp Neurol *466*, 48-79.

Kunishio, K., and Haber, S.N. (1994). Primate cingulostriatal projection: limbic striatal versus sensorimotor striatal input. J Comp Neurol *350*, 337-356.

Lammel, S., Lim, B.K., Ran, C., Huang, K.W., Betley, M.J., Tye, K.M., Deisseroth, K., and Malenka, R.C. (2012). Input-specific control of reward and aversion in the ventral tegmental area. Nature *491*, 212-217.

Lau, B., and Glimcher, P.W. (2007). Action and outcome encoding in the primate caudate nucleus. J Neurosci *27*, 14502-14514.

Lecourtier, L., Neijt, H.C., and Kelly, P.H. (2004). Habenula lesions cause impaired cognitive performance in rats: implications for schizophrenia. Eur J Neurosci *19*,

2551-2560.

Li, B., Piriz, J., Mirrione, M., Chung, C., Proulx, C.D., Schulz, D., Henn, F., and Malinow, R. (2011). Synaptic potentiation onto habenula neurons in the learned helplessness model of depression. Nature *470*, 535-539.

Matsumoto, M., and Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. Nature *447*, 1111-1115.

Matsumoto, M., and Hikosaka, O. (2009a). Representation of negative motivational value in the primate lateral habenula. Nat Neurosci *12*, 77-84.

Matsumoto, M., and Hikosaka, O. (2009b). Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature *459*, 837-841.

Matsumoto, M., and Hikosaka, O. (2011). Electrical stimulation of the primate lateral habenula suppresses saccadic eye movement through a learning mechanism. PLoS One *6*, e26701.

Matsumoto, M., Matsumoto, K., Abe, H., and Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. Nat Neurosci *10*, 647-656.

Miller, M.W., Powrozek, T.A., and Vogt, B.A. (2009). Dopamine systems in the cingulate gyrus: organization, development, and neurotoxic vulnerability. In Cingulate Neurobiology and Disease, B.A. Vogt, ed. (New York: Oxford University Press), pp. 163-187.

Morecraft, R.J., and Van Hoesen, G.W. (1998). Convergence of limbic input to the cingulate motor cortex in the rhesus monkey. Brain Res Bull *45*, 209-232.

Niki, H., and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. Brain Res *171*, 213-224.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C. (2001).

Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci *4*, 95-102.

Omelchenko, N., Bell, R., and Sesack, S.R. (2009). Lateral habenula projections to dopamine and GABA neurons in the rat ventral tegmental area. Eur J Neurosci *30*, 1239-1250.

Pandya, D.N., Van Hoesen, G.W., and Mesulam, M.M. (1981). Efferent connections of the cingulate gyrus in the rhesus monkey. Exp Brain Res *42*, 319-330.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat Rev Neurosci *2*, 417-424.

Quilodran, R., Rothe, M., and Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. Neuron *57*, 314-325.

Schoenbaum, G., Chiba, A.A., and Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat Neurosci *1*, 155-159.

Schoenbaum, G., and Setlow, B. (2003). Lesions of nucleus accumbens disrupt learning about aversive outcomes. J Neurosci *23*, 9833-9841.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. J Neurophysiol *80*, 1-27.

Seo, H., and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. J Neurosci *27*, 8366-8377.

Seo, H., and Lee, D. (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. J Neurosci *29*, 3627-3641.

Shima, K., and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. Science *282*, 1335-1338.

Stuphorn, V., Taylor, T.L., and Schall, J.D. (2000). Performance monitoring by the supplementary eye field. Nature *408*, 857-860.

Sul, J.H., Kim, H., Huh, N., Lee, D., and Jung, M.W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron *66*, 449-460.

Thornton, E.W., and Bradbury, G.E. (1989). Effort and stress influence the effect of lesion of the habenula complex in one-way active avoidance learning. Physiol Behav *45*, 929-935.

Ullsperger, M., and von Cramon, D.Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. J Neurosci *23*, 4308-4314.

Vezoli, J., and Procyk, E. (2009). Frontal feedback-related potentials in nonhuman primates: modulation during learning and under haloperidol. J Neurosci *29*, 15675-15683.

Vogt, B.A., and Pandya, D.N. (1987). Cingulate cortex of the rhesus monkey: II. Cortical afferents. J Comp Neurol *262*, 271-289.

Williams, S.M., and Goldman-Rakic, P.S. (1998). Widespread origin of the primate mesofrontal dopamine system. Cereb Cortex *8*, 321-345.

Williams, Z.M., Bush, G., Rauch, S.L., Cosgrove, G.R., and Eskandar, E.N. (2004). Human anterior cingulate neurons and the integration of monetary reward with motor responses. Nat Neurosci *7*, 1370-1375.

Wirth, S., Avsar, E., Chiu, C.C., Sharma, V., Smith, A.C., Brown, E., and Suzuki, W.A. (2009). Trial outcome and associative learning signals in the monkey hippocampus. Neuron *61*, 930-940.

FIGURE LEGENDS

Figure 1. Reversal learning task and behavioral performance

(**A**) Unilateral-probabilistic reversal learning task. (**B**) Choices and consequent outcomes during an example session in monkey A. The monkey chose the left or right target (vertical row) in a given trial (horizontal row). Droplet marks indicate rewarded trials, whereas those with a red cross indicate unrewarded trials. Gray rectangles indicate positions associated with the reward with 50% probability. (**C**) Change in the probability of choosing the reward-associated target after the reversal of reward-position contingency. Magenta and cyan plots indicate the data from monkeys A and E, respectively. (**D**) Shift probability after receiving the reward or no reward in the last trial. Double asterisks indicate a significant difference ($p < 0.01$, Wilcoxon signed-rank test). (**E**) Shift probability after choosing the same target was repeatedly unrewarded in past trials. Error bars indicate SEM. See also **Figure S1**.

Figure 2. Responses of LHb and ACC neurons to the positive and negative outcomes

(**A**) Activity of example neurons recorded in the LHb (left) and ACC (right). Rasters and peristimulus time histograms (PSTHs) are aligned by outcome onset. Rewarded and unrewarded trials are indicated by blue and red, respectively. (**B**) Proportions of negative-outcome type neurons (red), positive-outcome type neurons (blue), and neurons with no significant modulation by the outcomes (gray) in the LHb (left) and ACC (right). (**C**) Top, outcome-dependent modulation of the activity of all recorded neurons in the LHb (left, $n = 62$) and ACC (right, $n = 359$). The modulation of each neuron is presented as a row of pixels. The color of each pixel indicates the magnitude of the modulation,

which is expressed as an ROC value discriminating the discharge rate between rewarded and unrewarded trials. The ROC value was calculated by a 200 ms test window sliding with a 10 ms step. Warm colors (ROC > 0.5) indicate higher discharge rates in unrewarded trials, while cool colors (ROC < 0.5) indicate higher discharge rates in rewarded trials. Bottom, mean ROC value of the negative- and positive-outcome type neurons in the LHb (left) and ACC (right). The ROC value of the positive-outcome type neurons was flipped at 0.5 before averaging. Gray areas indicate the periods that were used to analyze the outcome-evoked responses of LHb and ACC neurons.

Figure 3. Effect of a single past outcome in the previous trial on the no-reward evoked response

(**A**) Activity of three example neurons; negative-outcome type in the LHb (left), negative-outcome type in the ACC (middle), and positive-outcome type in the ACC (right). Rasters and PSTHs are aligned by the onset of the negative outcome and are shown for the past-rewarded condition (brown) and the past-unrewarded condition (purple). (**B**) Distributions of the ROC values of the negative-outcome LHb neurons (n = 49, left), the negative-outcome ACC neurons (n = 98, middle), and the positive-outcome ACC neurons (n = 75, right) for discriminating their no-reward evoked responses between the past-rewarded and the past-unrewarded conditions. ROC values more than 0.5 indicate higher discharge rates in the past-unrewarded condition. Black bars indicate neurons showing a significant difference in their no-reward evoked responses between the two conditions ($p < 0.05$, Wilcoxon rank-sum test). Arrowheads indicate mean ROC values. n.s. indicate no significant deviation from 0.5 ($p > 0.05$, Wilcoxon signed-rank test). (**C**) Percentage of neurons showing a significant difference in their

no-reward evoked responses between the past-rewarded and the past-unrewarded conditions ($p < 0.05$, Wilcoxon rank-sum test). Dotted lines indicate chance level calculated by Monte Carlo analysis. Double asterisks above bars indicate a significantly larger proportion than the chance level ($p < 0.01$). Single asterisk between bars indicates a significant difference between the proportions ($p < 0.05$, Fisher's exact test).

Figure 4. Effect of negative outcome repetition in past trials on the no-reward evoked response

(**A**) Activity of three example neurons; negative-outcome type in the LHb (left), negative-outcome type in the ACC (middle), and positive-outcome type in the ACC (right). PSTHs are aligned by the onset of the negative outcome and are shown for the activity after choosing the same target was unrewarded in one (gray), two (yellow), three (orange), and four (red) consecutive trial(s). (**B**) Distributions of correlation coefficients between the magnitude of the no-reward evoked response and the number of no-reward repetition for the negative-outcome LHb neurons ($n = 49$, left), the negative-outcome ACC neurons ($n = 98$, middle), and the positive-outcome ACC neurons ($n = 75$, right). Black bars indicate neurons showing a significant correlation ($p < 0.05$). Arrowheads indicate mean correlation coefficients. n.s. indicate no significant deviation from 0 ($p > 0.05$, Wilcoxon signed-rank test). (**C**) Averaged magnitudes of the normalized (z-scored) no-reward evoked responses of the neurons with a significant positive (top) and negative correlation (bottom) ($p < 0.05$) plotted against the number of no-reward repetition. Light gray, the negative-outcome LHb neurons (top, $n = 2$; bottom $n = 3$). Black, the negative-outcome ACC neurons (top, $n = 17$; bottom, $n = 10$). Dark gray, positive-outcome ACC neurons (top, $n = 4$; bottom, $n = 8$). Error bars indicate SEM. (**D**)

Percentage of neurons showing a significant correlation ($p < 0.05$). Dotted lines indicate chance level calculated by Monte Carlo analysis. Double asterisks above bars indicate a significantly larger proportion than the chance level ($p < 0.01$). Single asterisk between bars indicates a significant difference between the proportions ($p < 0.05$, Fisher's exact test).

Figure 5. Effect of subsequent choice behavior on the no-reward evoked response

(**A**) Activity of three example neurons; negative-outcome type in the LHb (left), negative-outcome type in the ACC (middle), and positive-outcome type in the ACC (right). Rasters and PSTHs are aligned by the onset of the negative outcome and are shown for next shift trials (orange) and next stay trials (green). (**B**) Distributions of the ROC values of the negative-outcome LHb neurons ($n = 49$, left), the negative-outcome ACC neurons ($n = 98$, middle), and the positive-outcome ACC neurons ($n = 75$, right) for discriminating their no-reward evoked responses between next shift and next stay trials. ROC values more than 0.5 indicate higher discharge rates in next shift trials. Black bars indicate neurons showing a significant difference in their no-reward evoked response between the trials ($p < 0.05$, Wilcoxon rank-sum test). Arrowheads indicate mean ROC values. n.s. indicate no significant deviation from 0.5 ($p > 0.05$, Wilcoxon signed-rank test). (**C**) Percentage of neurons showing a significant difference in their no-reward evoked responses between next shift and next stay trials. ($p < 0.05$, Wilcoxon rank-sum test). Dotted lines indicate chance level calculated by Monte Carlo analysis. Double and single asterisks above bars indicate a significantly larger proportion than the chance level ($p < 0.01$ and 0.05, respectively). Single asterisk between bars indicates a significant difference between the proportions ($p < 0.05$, Fisher's exact test).

Figure 6. Individual effects of past outcome experience and subsequent choice behavior

(**A**) No-reward evoked response of an ACC neuron example showing a clear modulation between next stay and next shift trials (right) but no clear modulation between the past-unrewarded and the past-rewarded conditions (left). (**B**) No-reward evoked response of another ACC neuron example showing a clear modulation between the past-unrewarded and the past-rewarded conditions (left) but no clear modulation between next stay and next shift trials (right). (**C**) Regression coefficients for the outcome in the previous trial (ordinate) and the choice behavior in the next trial (abscissa) for the negative-outcome LHb neurons (n = 49, left), the negative-outcome ACC neurons (n = 98, middle), and the positive-outcome ACC neurons (n = 75, right). Purple and orange dots indicate neurons showing a significant regression coefficient for the outcome in the previous trial and the choice behavior in the next trial, respectively ($p < 0.05$). Cyan dots indicate neurons showing a significance for both. White dots indicate neurons with no significance. Numbers in the scatter plot indicate the regression coefficients of the neurons shown in (A) and (B).

Figure 7. Comparison between the dorsal and the ventral ACC

(**A**) Proportions of negative-outcome type neurons (red), positive-outcome type neurons (blue), and neurons with no significant modulation by the outcomes (gray) in the dorsal (left) and ventral ACC (right). (**B**-**D**) Percentage of neurons showing a significant difference in their no-reward evoked responses between the past-rewarded and the past-unrewarded conditions ($p < 0.05$, Wilcoxon rank-sum test) (B), percentage of neurons showing a significant correlation between the normalized magnitude of the

no-reward evoked response and the number of no-reward repetition ($p < 0.05$) (C), and percentage of neurons showing a significant difference in their no-reward evoked responses between next shift and next stay trials ($p < 0.05$, Wilcoxon rank-sum test) (D). Dotted lines indicate chance level calculated by Monte Carlo analysis. Double asterisks above bars indicate a significantly larger proportion than the chance level ($p < 0.01$). Single asterisk between bars indicates a significant difference between the proportions ($p < 0.05$, Fisher's exact test). n.s. indicate no significance ($p > 0.05$, Fisher's exact test).

Figure 1
Click here to download Figure: Figure_1_revise_c.tif

Figure 1

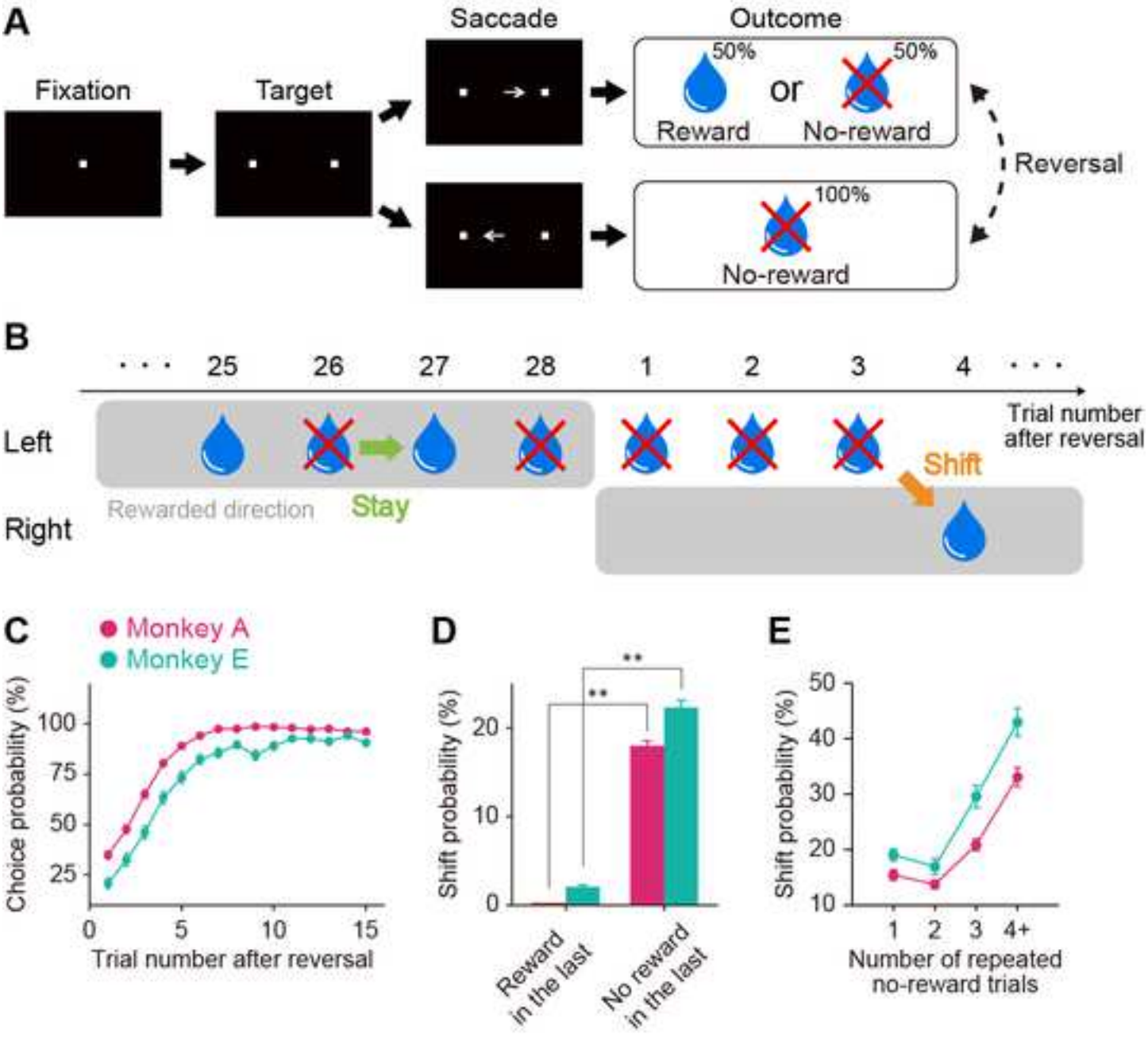


Figure 2

Click here to download Figure: Figure_2_revise_c_tk2.tif

Figure 2

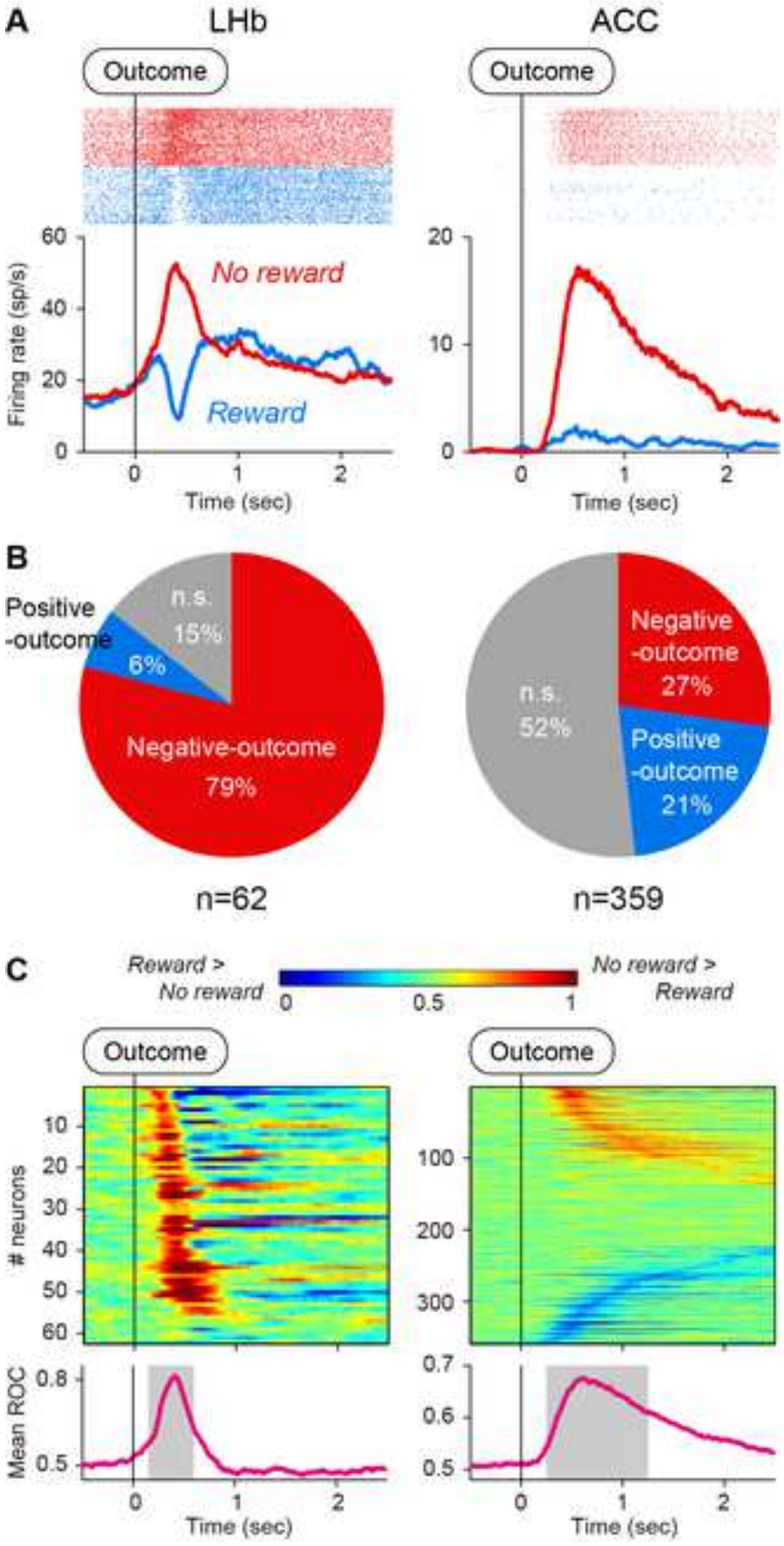


Figure 3
Click here to download Figure: Figure_3_revise_c2_tk.tif

Figure 3

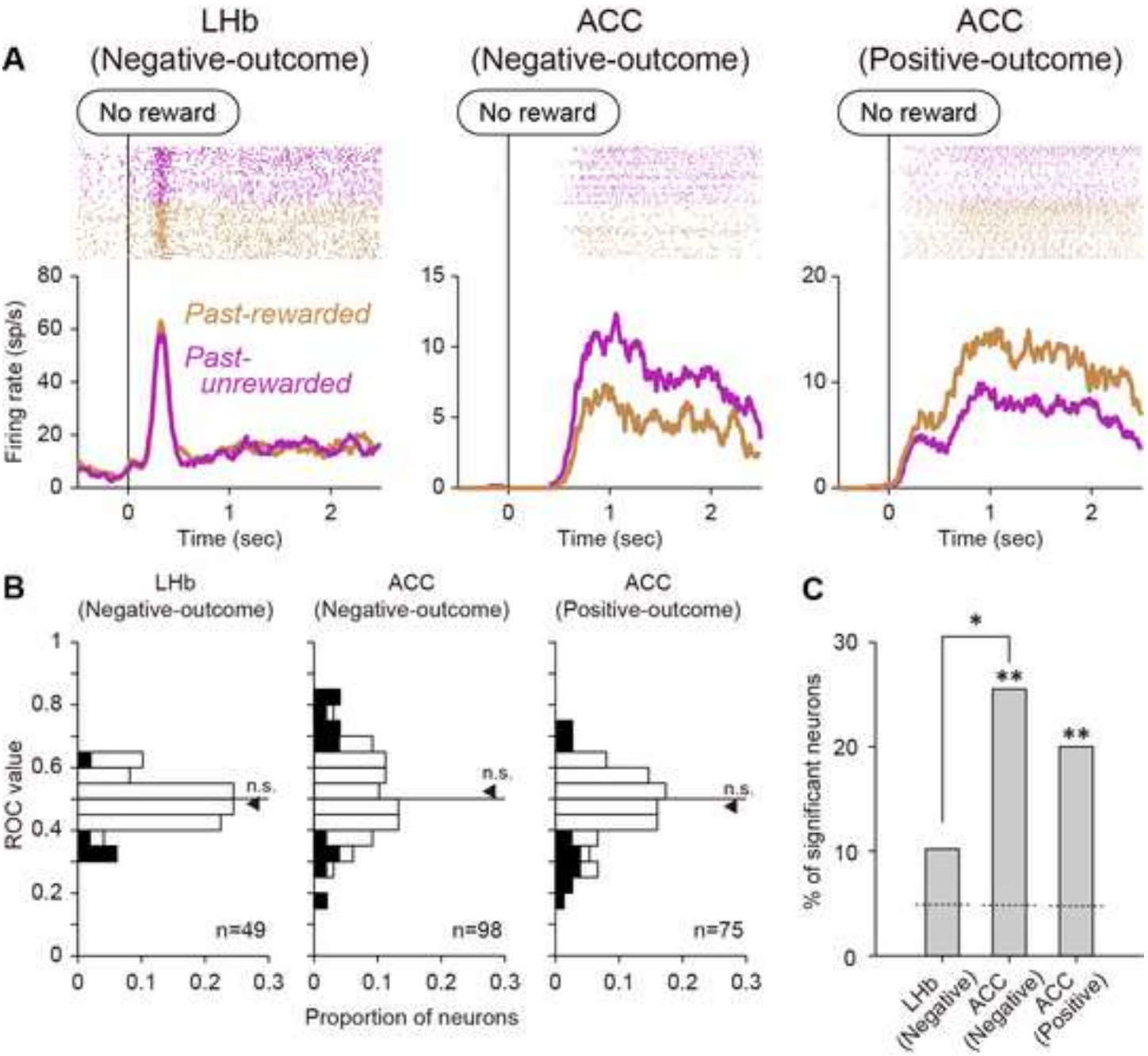


Figure 4

Click here to download Figure: Figure_4_revise_c2_tk.tif

Figure 4

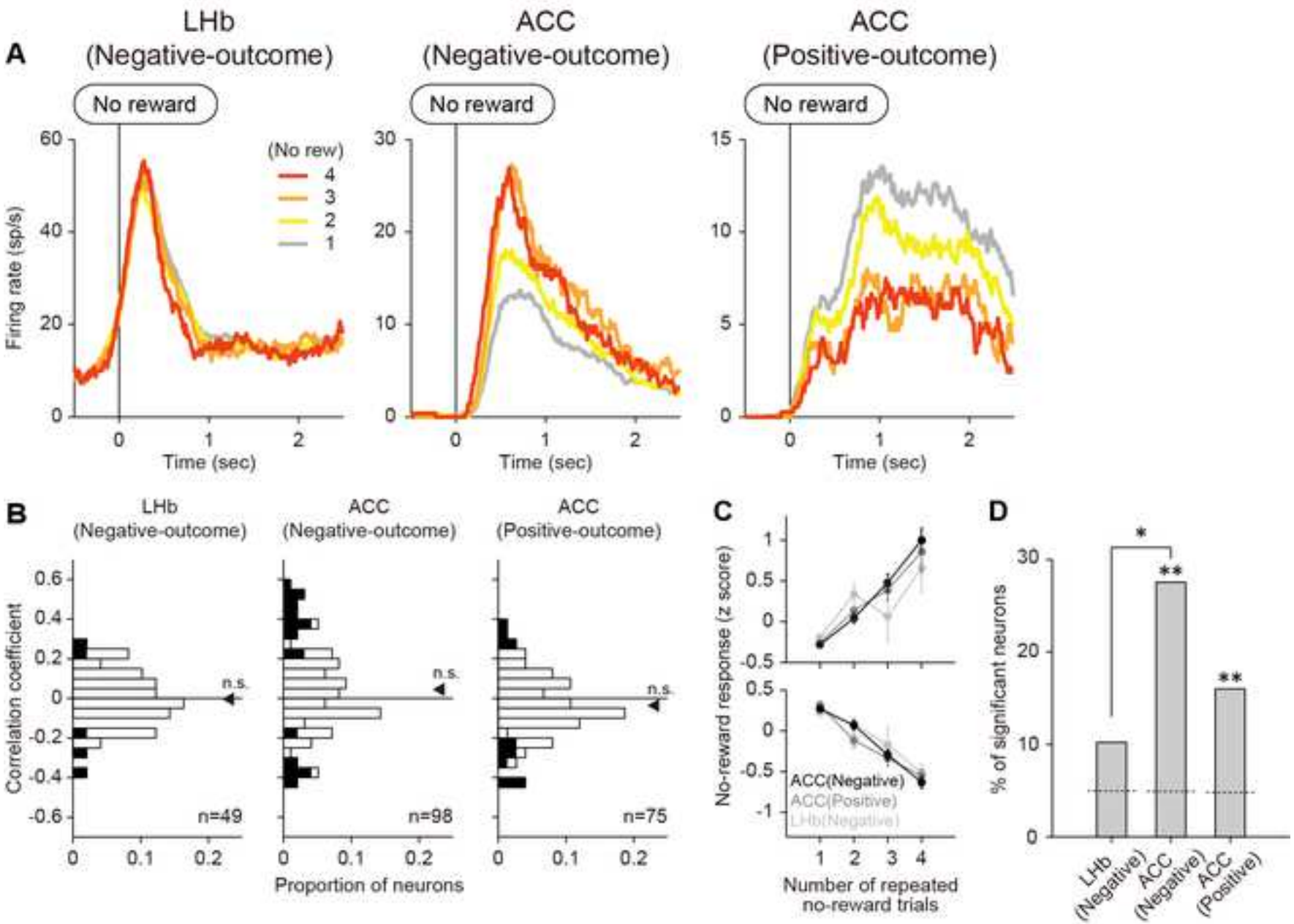


Figure 5
Click here to download Figure: Figure_5_revise_c2_tk.tif

Figure 5

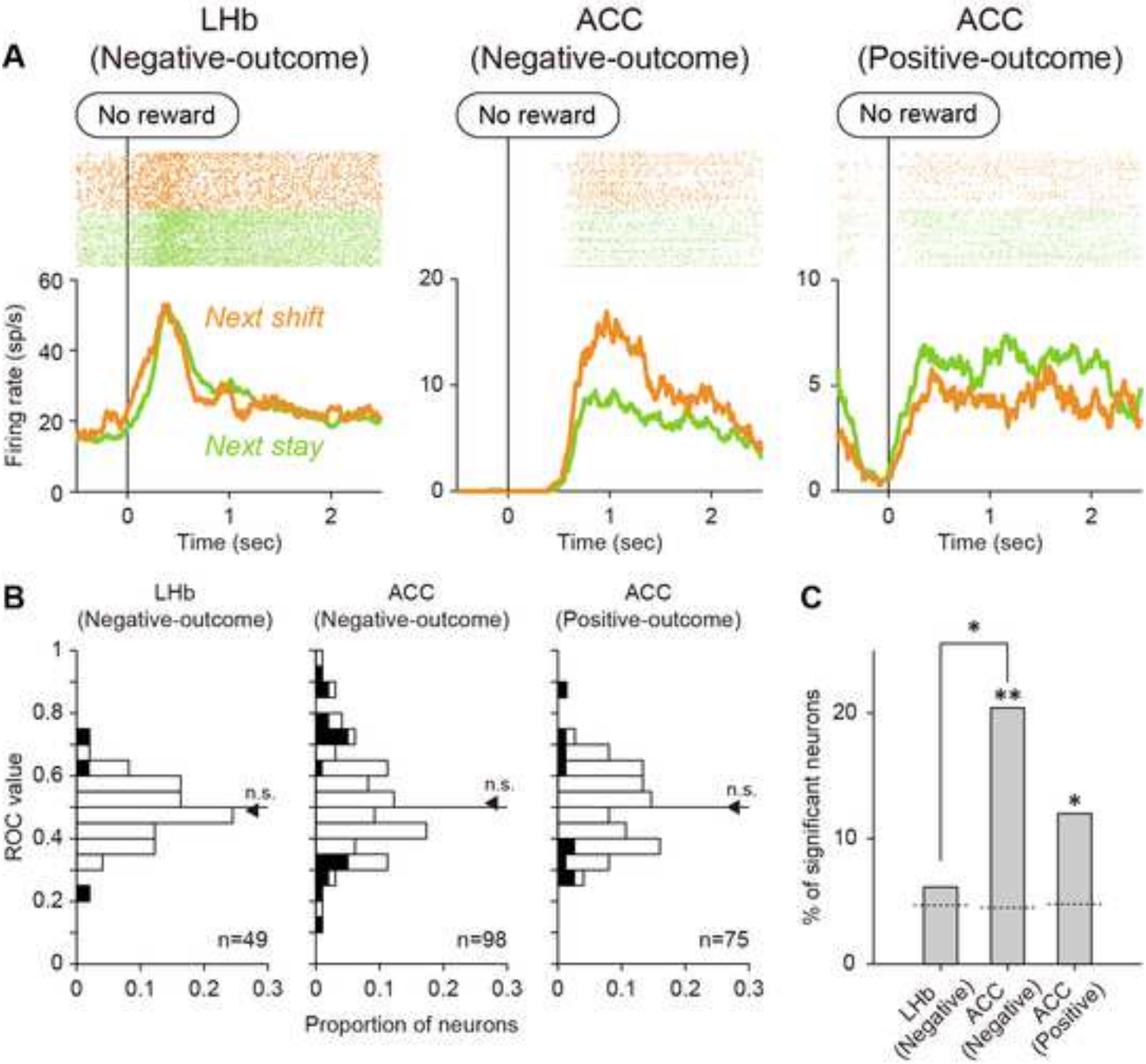


Figure 6

Click here to download Figure: Figure_6_revise_c.tif

Figure 6

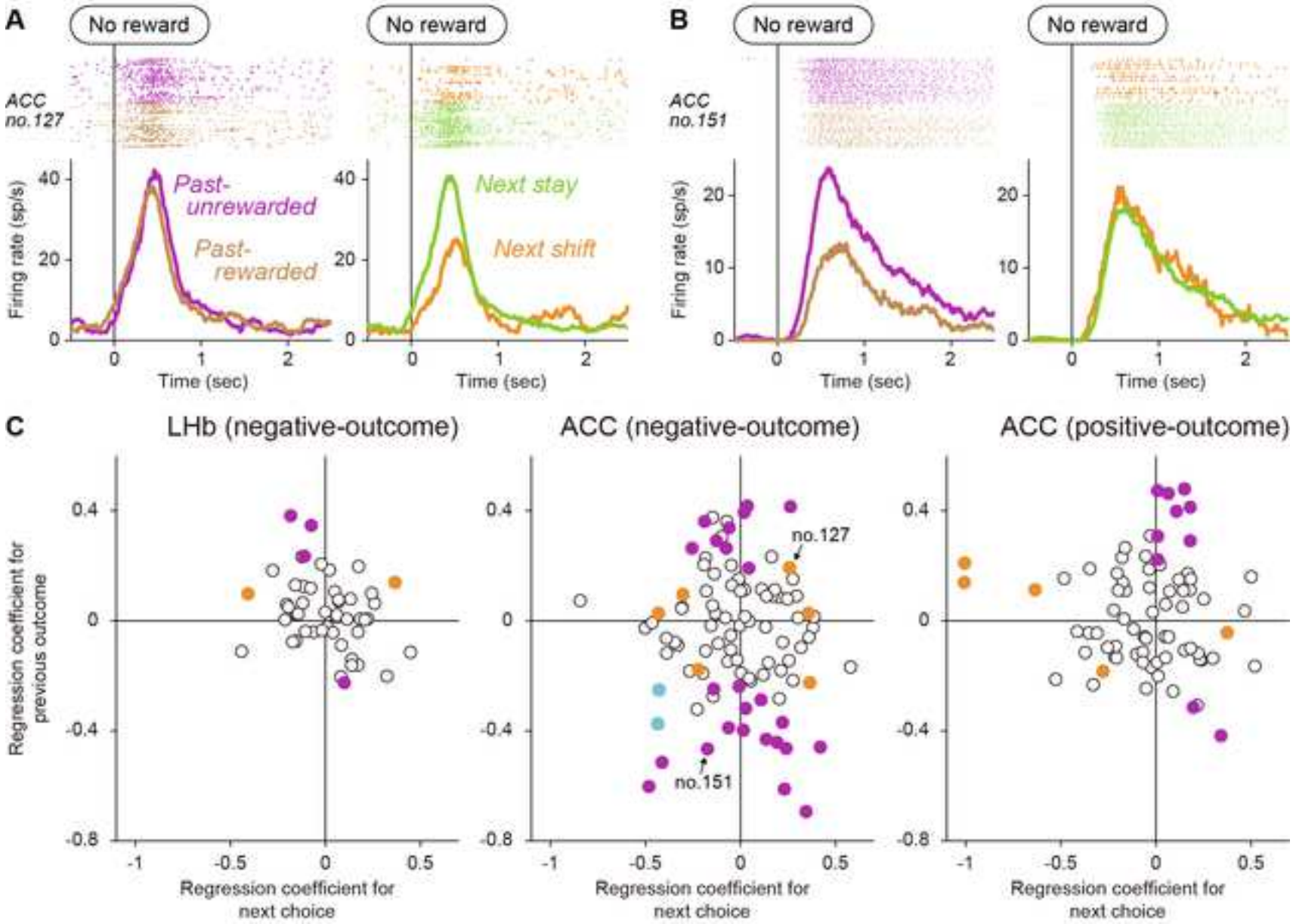
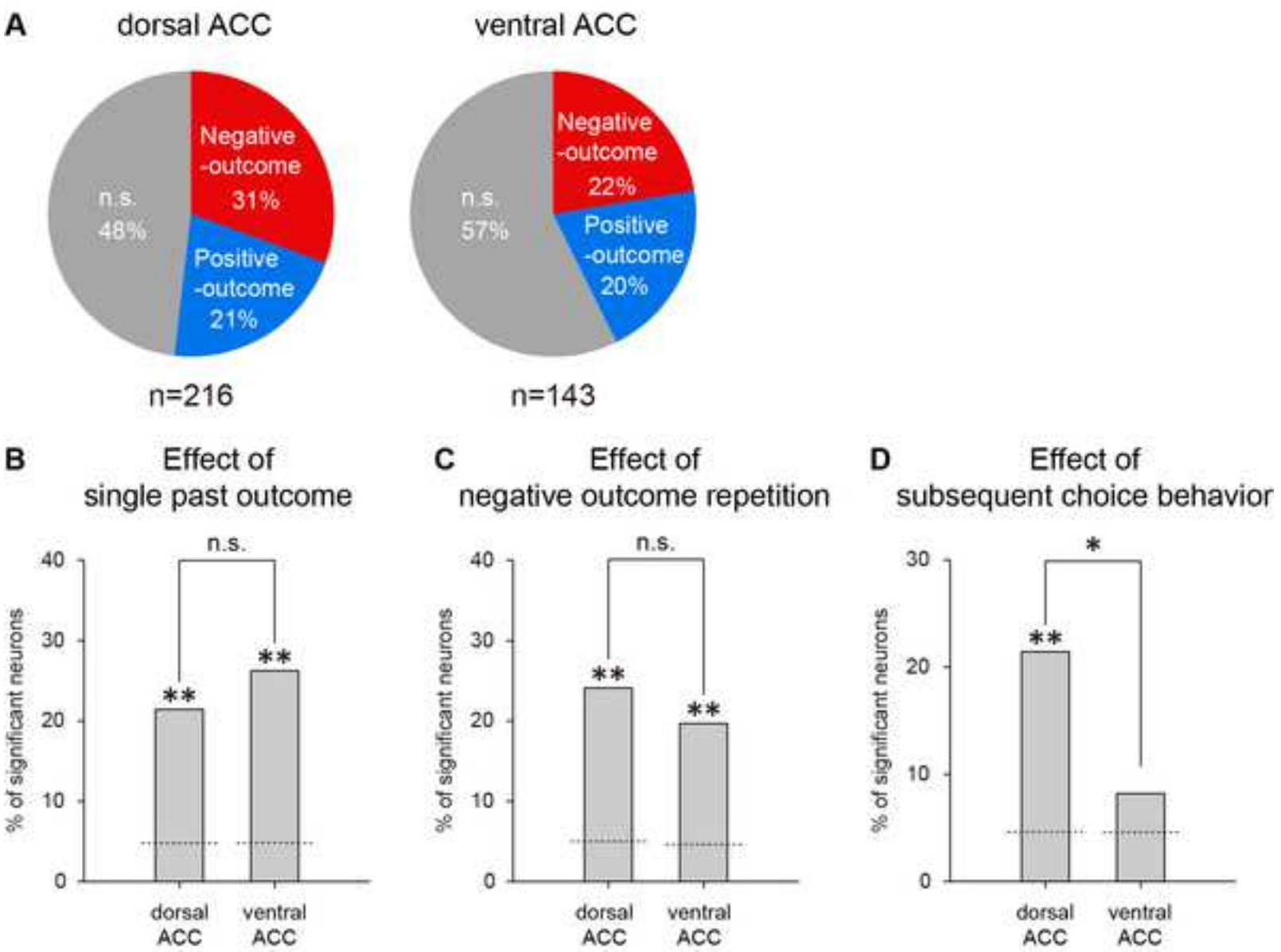


Figure 7

Click here to download Figure: Figure_7_revise_c2_tk.tif

Figure 7



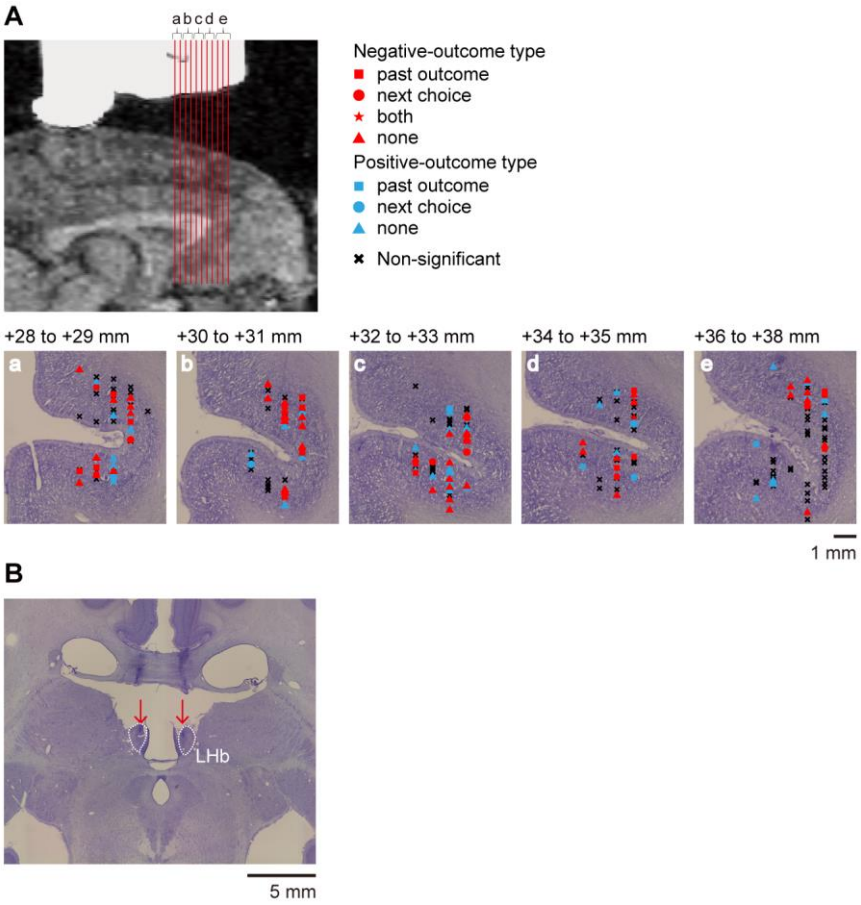


Figure S1 (related to Figure 1). Recording sites of LHb and ACC neurons

(**A**) Sagittal view of the brain by MRI and histological reconstruction of ACC recording sites in monkey A. Red lines in the MRI indicate the positions of the coronal sections depicted in a lower row. Red and blue symbols indicate negative- and positive-outcome type neurons, respectively. Colored squares, circles, stars, and triangles indicate neurons showing a significant regression coefficient for the outcome in the previous trial, for the choice behavior in the next trial, for both, and for neither of them, respectively ($p < 0.05$) (see EXPERIMENTAL PROCEDURES for the regression coefficient). Black crosses indicate neurons with no significant modulation by the outcomes. (**B**) Histological reconstruction of LHb recording sites. Arrows point to the electrolytic microlesions made at representative locations of electrode penetration.