Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision

Tommy C. Blanchard, [ID] **Caleb E. Strait, and Benjamin Y. Hayden**
Department of Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, New York

Blanchard TC, Strait CE, Hayden BY. Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *J Neurophysiol* 114: 2439–2449, 2015. First published September 2, 2015; doi:10.1152/jn.00711.2015.—We frequently need to commit to a choice to achieve our goals; however, the neural processes that keep us motivated in pursuit of delayed goals remain obscure. We examined ensemble responses of neurons in macaque dorsal anterior cingulate cortex (dACC), an area previously implicated in self-control and persistence, in a task that requires commitment to a choice to obtain a reward. After reward receipt, dACC neurons signaled reward amount with characteristic ensemble firing rate patterns; during the delay in anticipation of the reward, ensemble activity smoothly and gradually came to resemble the postreward pattern. On the subset of risky trials, in which a reward was anticipated with 50% certainty, ramping ensemble activity evolved to the pattern associated with the anticipated reward (and not with the anticipated loss) and then, on loss trials, took on an inverted form anticorrelated with the form associated with a win. These findings enrich our knowledge of reward processing in dACC and may have broader implications for our understanding of persistence and self-control.

persistence; dorsal anterior cingulate cortex; anticipatory activity; reward signaling

MANY REWARDS require persistent commitment to a choice before we obtain them. For example, marathon runners need to continue running to finish a race, predatory animals need to chase down their prey, and scientists need to spend many hours performing research to get published. Persistence is a central component of self-control, the ability to pursue a goal despite impulses to do otherwise (de Ridder et al. 2012).

Despite its importance, we know very little about the neural processes that occur during persistent commitment to a decision. Information about these neural processes, especially in brain regions linked to self-control, could help us to understand the neural mechanisms of persistence and, ultimately, self-control. We are particularly interested in the dorsal anterior cingulate cortex (dACC). The dACC has been associated with successful self-control in a number of contexts, including an intertemporal choice task (Peters and Büchel 2010), delay tasks (Narayanan et al. 2006; Narayanan and Laubach 2006), response inhibition tasks (Floden and Stuss 2006), and forced swim tasks (Warden et al. 2012). Activation of human dACC produces intense feelings of the will to persevere against challenges (Parvizi et al. 2013). One possibility is that the association between dACC and self-control is due to the role of dACC in facilitating persistence (Chudasama et al. 2013;

Picton et al. 2007). Although this cortical region has been closely linked to persistence, its precise role in this process remains unclear (Gusnard et al. 2003; Parvizi et al. 2013).

Previous research has shown that activation of the brain's reward circuitry is critical for successful persistence (Gusnard et al. 2003; McGuire and Kable 2015). One possibility is that brain activation elicited by an expected reward may serve to maintain motivation and ensure successful persistence (Howe et al. 2013). One way it could do this is to actively maintain a value signal that allows for decision-makers to overcome a tendency to succumb to temptation (Hillman and Bilkey 2010). Thus the evaluation and representation of expected rewards may be a critical part of persistence. Consistent with this idea, the ACC has been found to be critical for reward processing, particularly with respect to expected rewards (Alexander and Brown 2010; Amiez et al. 2006; Bush et al. 2002; Sallet et al. 2007).

Although the responses of dACC neurons during persistence remain unknown, studies have shed light on their function in similar contexts. In tasks requiring multiple steps before a reward is given, ensembles of putative dACC neurons in rats show a ramplike change in activity as the animals progress to a reward (Ma et al. 2014; Shidara and Richmond 2002). Similar ramping patterns have been found in macaque dACC (Hayden et al. 2011b; Toda et al. 2012). Numerous functions have been suggested for this ramping activity, such as keeping track of progress toward a reward (Ma et al. 2014; Shidara and Richmond 2002), timing an interval (Durstewitz 2003), and deliberate inhibition of an inappropriate action during a delay (i.e., persistence; Narayanan and Laubach 2006, 2009). Despite the functional importance of this ramping activity, little is known about the information it carries.

We analyzed the activity of dACC neurons in rhesus macaques as they performed a foraging task—specifically, a diet selection task—that required them to maintain fixation through a delay to receive a reward (Blanchard and Hayden 2014). On each trial, animals decided whether to accept or reject an option based on the reward size and delay being offered. To accept an option, subjects were required to fixate on the option throughout the delay, an action that requires a persistent commitment but probably very little effort. To reject, subjects had to avert their gaze.

Throughout prefrontal cortex and dACC, neurons have characteristic responses to rewards (see, e.g., Azzi et al. 2012; Blanchard et al. 2015; Blanchard and Hayden 2014; Strait et al. 2014; Toda et al. 2012; Wallis and Kennerley 2011). We found that throughout the delay neurons in our population gradually changed their activity to these characteristic responses, even on trials with risky options (which offered a 50% chance of

Address for reprint requests and other correspondence: T. C. Blanchard, Dept. of Psychology, Harvard Univ., 52 Oxford St., Cambridge MA 02138 (e-mail: tblanchard@g.harvard.edu).

reward). After a risky loss (no reward), neural responses took an inverted form—those that were excited by reward receipt tended to be inhibited and vice versa. We also found that the formats neurons used to encode reward size (that is, the ensemble pattern of responses) were different during the choice epoch of the task and when the reward was delivered. The coding format of anticipated reward size quickly changed shortly after the choice epoch but came to resemble the coding format used during the reward epoch prior to the delivery of reward. Thus, despite the receipt of reward being a discrete event occurring on a very short timescale, neurons were responding to reward size just prior to the actual receipt of the reward similarly as they did immediately after reward receipt.

These findings suggest that similar neural processes are active during persistence and reward receipt. We speculate that the reward representation we observe gradually arising when a reward is anticipated may facilitate persistence (Hillman and Bilkey 2010).

MATERIALS AND METHODS

Ethics statement. All procedures were approved by the University of Rochester Institutional Animal Care and Use Committee and were designed and conducted in compliance with the Public Health Service's *Guide for the Care and Use of Laboratory Animals.*

Task. On each trial of our task, an option appeared at the top of the screen and moved smoothly at a constant rate downward. All options were horizontally oriented rectangles 80 pixels high and of variable width (60–300 pixels; Fig. 1). Option color indicated reward value: orange (0.075 ml), gray (0.135 ml), green (0.212 ml), or cyan (0.293 ml). In addition to these four colors, one-fifth of the options were divided horizontally into half-cyan and half-red portions; these offered a 50% chance of receiving a 0.293-ml reward and a 50% chance of receiving no reward. Option width indicated the delay associated with that option. Option widths were 60, 120, 180, 240, and 300 pixels and corresponded to delays of 2, 4, 6, 8, and 10 s, respectively. Each

possible option (25 options, 5 widths × 5 colors) appeared with equal frequency; width and color were selected randomly and independently on each trial.
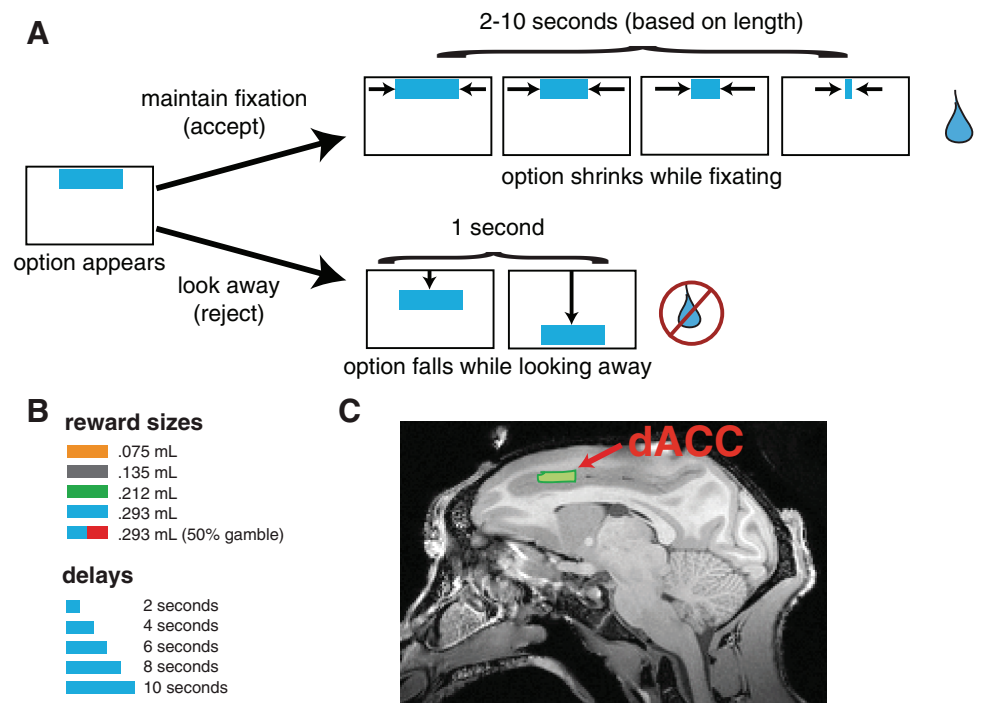
Two male rhesus macaques (*Macaca mulatta*; *monkey B* and *monkey J*) performed the task. On each trial, a subject could select an option by fixating it or reject the option by avoiding direct gaze on it. In the absence of any action, each option took 1 s to move vertically downward from the top of the display area of the computer monitor to the bottom, after which time it disappeared and could no longer be chosen. In this case, the trial would end and a 1.5-s intertrial interval (ITI) would begin. If the monkey selected an option by fixating it, the option would stop moving wherever it was and then would begin to shrink horizontally. Shrinking rate was constant (30 pixels/s), and thus option width served to identify the total remaining delay associated with each option.

If the monkey averted its gaze from the option during the shrinking phase, the option would (after a 0.25-s grace period) continue its movement toward the bottom of the screen. As it moved its width would remain at what it had been when gaze was averted, and if it was fixated upon again it would again pause and begin shrinking from its new, smaller width. If at any point the monkey held an option until it shrunk entirely, the appropriate reward would be delivered, the trial would end, and a 1.5-s ITI would follow.

Behavioral techniques. Horizontal and vertical eye positions were sampled at 1,000 Hz by an infrared eye-monitoring camera system (EyeLink 1000, SR Research, Mississauga, ON, Canada). We wrote our experiments in MATLAB (MathWorks), using the Psychophysics and Eyelink Toolbox extensions. A standard solenoid valve (Parker, Cleveland, OH) controlled the duration of water delivery. Immediately before recording, we performed a careful calibration of our solenoid system to establish a precise relationship between solenoid open time and water volume in our rigs.

Surgical procedures. Two male rhesus monkeys (*M. mulatta*) served as subjects. Initially, a head-holding mount was implanted with standard techniques. Four weeks later, animals were habituated to laboratory conditions and trained to perform oculomotor tasks for liquid reward. A second surgical procedure was then performed to

Fig. 1. Task and recording location. *A*: behavioral task. Delayed rewards, represented by horizontally oriented rectangles, appeared on the screen. Options differed in their reward size (indicated by color) and delay (indicated by width). Monkeys accepted an option by fixating or rejected by averting gaze. If accepted, the rectangle would shrink at a constant rate, and reward would be delivered when it had shrunk completely. If rejected, no reward would be delivered. See MATERIALS AND METHODS for more detail. *B*: reward sizes and delay lengths used in the task. *C*: recording location. Our recordings focused on the anterior portion of 6/32 and most of 8/32, specifically, the dorsal bank of the cingulate sulcus, from 26 mm interaural to 34 mm interaural. dACC, dorsal anterior cingulate cortex.

place a 19-mm plastic recording chamber (Crist Instruments) over dACC (32 mm anterior, 7 mm medial interaural). Animals received analgesics and antibiotics after all surgeries. The chamber was kept sterile with regular antibiotic washes and sealed daily with sterile plastic caps.

Microelectrode recording techniques. Single electrodes (Frederick Haer; impedance range 0.8–4 MΩ) were lowered with a microdrive (NAN Instruments) until the waveform of one or more (1–3) neuron(s) was isolated. Individual action potentials were identified by standard criteria and isolated on a Plexon system (Plexon). Neurons were selected for study solely on the basis of the quality of isolation and never on the basis of task-related response properties.

We approached dACC through a standard plastic recording grid (Crist Instruments). dACC was identified by structural magnetic resonance images taken before the experiment and concatenated with Brainsight (Rogue Research). Neuroimaging was performed at the Rochester Center for Brain Imaging, on a Siemens 3-T MAGNETOM Trio TIM using 0.5-mm voxels. Chamber placement was reconciled with Brainsight. We also used Brainsight to guide placement of grids and to determine the location of our electrodes. We confirmed recording locations by listening for characteristic sounds of white and gray matter during recording, which in all cases matched the loci indicated by the Brainsight system with an error of <1 mm. Our recordings came from areas 6/32 and 9/32 according to the Paxinos atlas.

Vector analyses. To analyze our population of neurons, we made use of vector analyses. For these analyses, we took some measurement from each neuron—either the normalized activity of the neuron or a regression coefficient of reward size onto normalized firing rate—for a specific time window. The measurement values for each neuron were then placed into a vector, giving us a vector of length n, where n is the number of neurons used in the analysis. We then compared this vector to vectors coming from another time window and measured the difference (or similarity) of the vectors to each other. This allowed us to measure how similarly our entire population of neurons was acting in two different time windows.

Euclidean distance. We used Euclidean distance as one of our measurements of the difference between two vectors. The Euclidean distance is a standard distance measurement that describes the length of a straight line connecting the two end points of two vectors. The equation for Euclidean distance between two vectors p and q is

$$d(p, q) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}$$

where n is the number of dimensions (here, the number of neurons) and q_i and p_i refer to the value of element i in vectors q and p (the value of the ith neuron), respectively.

Cross-validation for vector analyses. Including all trials in our vector analyses (i.e., Figs. 3, 4, and 5*C*) would lead to inflated correlations, especially when comparing temporally close windows, as shared neural noise would drive both vectors in the same direction. Therefore, for each neuron we selected half of the trials (even-numbered trials) to form what would be our template vector and the other half for our comparison vector. For our coding format vectors, we regressed firing rate onto reward size separately for each half, and the regression coefficients formed the vectors we would compare. For our activity pattern vectors, we simply took the mean firing rate from each half. Thus there was no overlap in the data used to generate the two vectors.

RESULTS

The data presented here are a new analysis of a data set previously summarized in Blanchard and Hayden (2014).

Behavioral results. Our diet selection task is based on a well-known problem from foraging theory (Krebs et al. 1977; Stephens and Krebs 1987). On each trial, monkeys could accept or reject the offer of a delayed reward that appeared on

a computer monitor. Options differed on two dimensions: offered reward size and delay to that reward if chosen (Fig. 1, *A* and *B*; see MATERIALS AND METHODS). No other offers appeared during the delay (which is called handling time in foraging theory), so delays imposed an opportunity cost. Thus reward-maximizing behavior in this task requires sensitivity to both the reward size and the delay length of options (Blanchard and Hayden 2014; Stephens and Krebs 1987). As described in detail in another report, both animals quickly learned the task and performed near-optimally (Blanchard and Hayden 2014). Consistent with our previous studies, monkeys were more likely to accept options with large reward sizes and less likely to accept options with long delays [logistic regression of acceptance onto reward size and delay length, reward size $\beta = 1.78$, $t(38830) = 84.82$, $P < 0.0001$; delay length $\beta = -2.19$, $t(38830) = -93.03$, $P < 0.0001$; Blanchard et al. 2013; Blanchard and Hayden 2015; Pearson et al. 2010]. See Blanchard and Hayden (2014) for a more detailed analysis of behavior in this task.

dACC activity correlates with anticipated reward size throughout the entire trial. We recorded responses of 124 neurons in the dACC of two monkeys performing the diet selection task ($n = 74$ in *monkey B*, $n = 50$ in *monkey J*; Fig. 1*C* shows the recording position). Of these neurons, 3 (all from *monkey B*) were from a version of the task that involved no risky trials; we excluded these 3 neurons from all analyses in the present study and focus here on the remaining 121. Aside from this, we did no preselection or screening for neurons during data collection or prior to analysis. In our previous study, we focused on the differences between accept and reject trials (Blanchard and Hayden 2014). Given our interest in persistence and in reward anticipation here, we focused exclusively on accept trials (61.4% of all trials), as no persistence was required during reject trials.
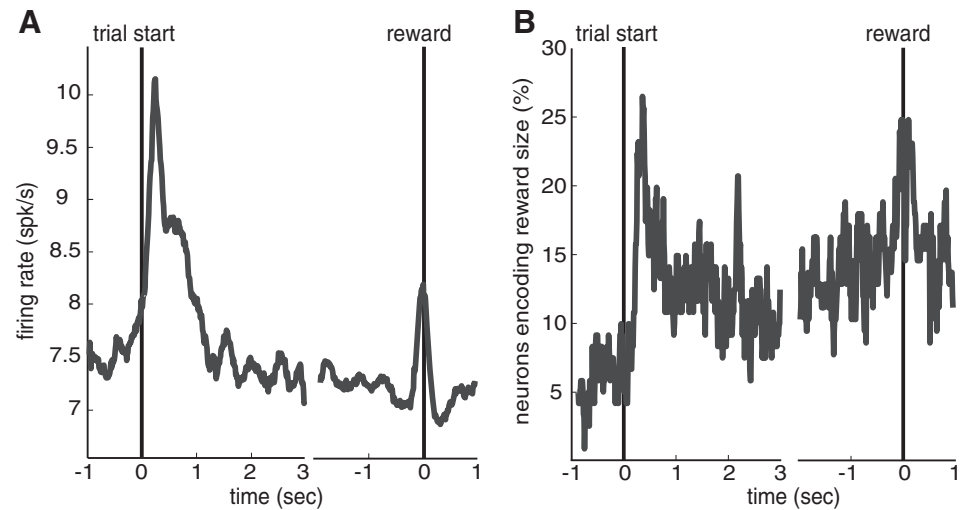
We focused first on the safe reward trials (80% of all offers). A basic event-aligned firing rate histogram revealed a burst of activity around the time of choice and the time of reward (Fig. 2*A*). Given the long-lasting nature of the bursts, we were concerned that the neural responses during the 2-s-delay trials would look significantly different from the longer-delay trials, possibly obscuring anticipation signals immediately before the reward. We therefore omitted these trials in our analyses of reward anticipation, unless otherwise noted.

We next looked for reward size correlations around the time of choice and reward delivery, using standard linear regression techniques. We wished to study the neural responses at the time of choice. In a previous study of this data set, we used behavioral data to estimate that the time of the decision is between 300 and 400 ms after option appearance in this task (Blanchard and Hayden 2014). This period corresponds to the peak of the initial rise in reward encoding we observe (Fig. 2*B*). We chose to define the choice epoch to be the same, but with some additional room to account for potential variability in neuronal latencies. Specifically, we used a 200-ms window from 250 to 450 ms after option appearance. Note that we use the term "encoding" of some variable to refer to correlations between firing rate and that variable.

We found that activity in 17.4% (21/121) of neurons was correlated with expected reward size during the choice epoch (linear regression, $\alpha = 0.05$). This proportion is significantly greater than the proportion expected by chance ($P < 0.001$,

Fig. 2. Basic neural responses. *A*: average firing rate of the population on safe trials where the option was accepted. *Left* side is aligned to trial start; *right* side is aligned to reward delivery. Vertical lines indicate the beginning of the trial and the reward delivery. *B*: % of neurons that significantly encode reward size as a function of time; sliding boxcar using 200-ms windows.

binomial test). Using a sliding boxcar analysis (boxcar width: 200 ms; step size: 10 ms), we found that the proportion of neurons whose activity was correlated with reward size remained relatively constant throughout the delay (Fig. 2*B*). We found that the population of neurons quickly responded to reward (Fig. 2*A*). Visual inspection revealed that the firing rate peaked around the time of delivery (Fig. 2*B*). We thus chose the 200-ms window beginning at reward delivery as our reward epoch. In the reward epoch, 20.7% (25/121) neurons significantly encoded reward size. This proportion is significantly greater than the proportion expected by chance ($P < 0.001$, binomial test). Thus we found quite modest but nonetheless significant correlations between firing rates and reward size, which appeared by conventional analyses to remain relatively constant in their preponderance throughout the various phases of the task.

Ensemble activity pattern gradually shifts toward reward state through delay. Average population firing rate provides little information about the pattern of activity changes. In particular, when averaging, neurons that increase their firing rate over the course of the delay could be counterbalanced with neurons that decrease their firing rate. This mixture of positive and negative responses to task variables is a ubiquitous feature of dACC and the prefrontal cortex more generally (Blanchard et al. 2015; Blanchard and Hayden 2014; Rigotti et al. 2013; Strait et al. 2014). Therefore, we next examined how the pattern of activity changed throughout the delay period.

We have opted to investigate ensemble responses, as opposed to traditional within-neuron analyses, for two reasons. First, because of the plurality of response profiles often found in dACC neurons, the use of conventional within-cell measures becomes difficult (cf. Ma et al. 2014). A simple linear trend may be found to be significant either when a neuron shows a smooth ramping in activity or when it shows no change in activity until a brief phasic burst. Second, we are interested here in describing how a signal in a population of neurons evolves over time. By using population-level measures to describe the changes in neural activity over the course of the trial, we are able to ask how the state of the ensemble of neurons is changing, instead of how each individual neuron is.

Our methods allowed us to get around the issues with looking for gradual changes in activity at the single-neuron level. These methods also provide an intuitive population-level

statistic instead of taking the additional step to summarize the statistics done on each neuron individually. However, we should stress that these methods are actually conceptually very simple—relying only on simple measures like Euclidean distance and correlation.

To study ensemble activity patterns, we formed normalized activity vectors with 121 elements each; each element corresponded to the normalized firing rate of a particular neuron during a 200-ms window. We could then use measures of similarity or distance between these vectors to judge how similar the pattern of activity in the population is within any two time windows.

We first used Euclidean distance to measure how similar two activity patterns were (MacEvoy and Epstein 2009). This technique measures distance between end points of vectors in an *N*-dimensional space, where a dimension is the firing rate of a specific neuron. The closer two vectors are, the more similar their patterns of activity. We used a simple cross-validation procedure to eliminate spurious similarities in the vectors caused by shared noise (see MATERIALS AND METHODS). We separated trials by delay length so we could investigate how activity patterns changed over the course of different delays. Because the subjects were unlikely to accept options with delays of 10 s, and we were therefore severely data-limited for this condition, we did not include options of *delay 10* for this analysis. A minority of neurons ($n = 15$) were from sessions in which the animal accepted fewer than five options of *delay 8*, and thus we removed these neurons and used the remaining 106 neurons for the *delay 8* analysis. We then ran two sliding boxcar analyses, one comparing activity in a sliding window to activity in a fixed window, the choice epoch normalized activity vector (Fig. 3*A*), and another comparing activity in the sliding window to the reward epoch normalized activity vector (Fig. 3*B*) for each delay length. We normalized the distance measure to be 1 at the greatest distance the vectors reached and 0 at the lowest distance.

We found that, regardless of delay length, there was only a brief period during which the sliding boxcar normalized activity vector was similar to the choice epoch normalized activity vector (Fig. 3*A*). Not surprisingly, the distance between the two vectors was lowest when the boxcar partially overlapped with the choice epoch. However, after this period, the boxcar activity became dissimilar quite rapidly. Specifically, the nor-
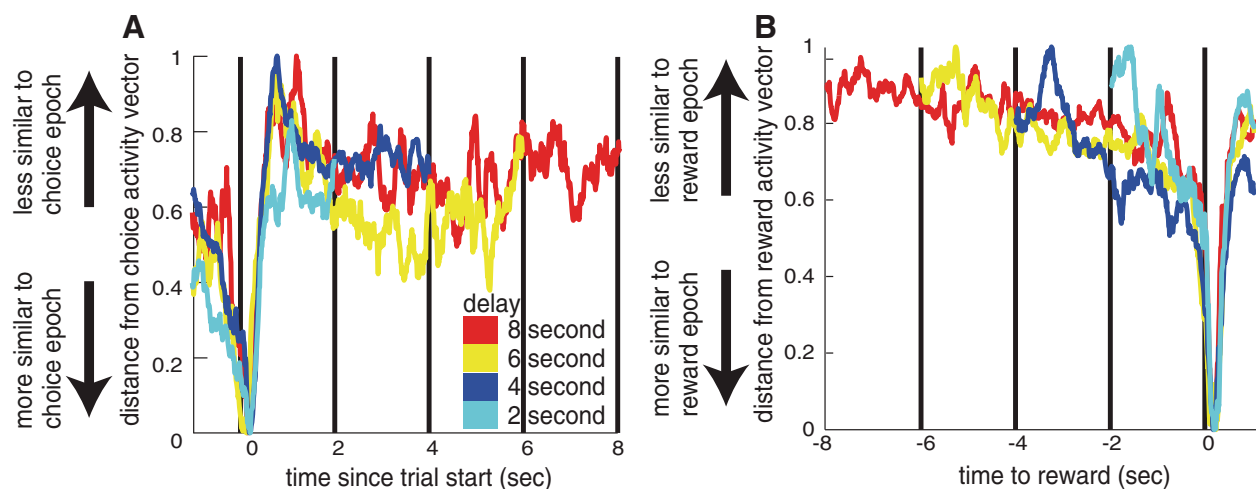
Fig. 3. Euclidean distance between activity pattern vectors shows gradual ramping toward reward activity pattern. *A* and *B*: ensemble activity pattern analysis using Euclidean distance: distance between the normalized activity vector from a running boxcar and the normalized activity vector from the choice epoch (*A*) and the reward epoch (*B*). Trials are separated out by delay length; color indicates the specific delay length. The ensemble activity pattern changes rapidly after the choice epoch and gradually comes to resemble the reward epoch pattern. Vertical lines indicate the start or end of a trial.

malized distance increased by 0.842 [bootstrapped 95% confidence interval (CI): 0.732–0.958] just 500 ms later (averaging across all duration conditions). In contrast to this rapid change with the choice epoch, we observed a slow and gradual increase in similarity between the sliding boxcar normalized activity vector and the reward epoch normalized activity vector throughout the duration of the trial. Thus from 500 ms after the choice epoch to 500 ms before the reward epoch, the normalized distance decreased significantly but by a small amount: 0.247 (95% CI: 0.196–0.312). The distance then decreased more rapidly in the final 500 ms before the reward, decreasing by 0.720 during this period (95% CI: 0.665–0.723).

This gradual ramping indicates that the pattern of activity shifts gradually toward the pattern of activity it will be in after receiving the reward. This gradual shift in activity is similar to that observed previously in dACC neurons while animals performed a task with explicit discrete steps that need to be performed in order to receive the reward (Ma et al. 2014).

Gradual shift in activity pattern vectors can be seen by using correlation as a similarity measure. Next, we wanted to ensure that the gradual ramping that we saw toward the reward activity pattern during the delay was not due to the similarity measure that we used. Thus we performed the same analysis, but using correlation instead of Euclidean distance. Positive correlation between two normalized activity vectors indicates that neurons are being modulated in similar ways in the two cases, independent of the overall population activity level (because correlation subtracts the means of the vectors). A correlation close to 0 means there is little relationship between the pattern of activity in the two vectors. To reduce noise, we did not separate trials by delay for this analysis.

Our analysis using correlation showed the same qualitative pattern of results as our analysis using Euclidean distance. The sliding boxcar normalized activity vector was only briefly highly correlated with the choice epoch normalized activity vector (Fig. 4*A*). They were significantly correlated only from the window beginning 530 ms before the choice epoch to the one beginning 370 ms after it, a total of 900 ms. It is notable that this significant correlation occurs prior to the beginning of the trial. We suspect that this early ramp-up is due to antici-

pation of the beginning of the trial. The ITI was constant (always 1.5 s), and thus the beginning of the trial could be predicted by the animal. Thus this early rise in correlation could be due to a rise in arousal or attention prior to the trial beginning or activity in preparation of choice or motor control. There was an increase of 0.701 in the correlation coefficient from when it first reached significance to the choice epoch ($r = 0.187$ 530 ms before choice epoch to $r = 0.887$ during choice epoch). Thus the correlation increased by 0.701 over 530 ms, an increase of 0.132 per 100 ms.

In contrast, the correlation between the boxcar and the reward epoch normalized activity vector increased more gradually. The correlation reached significance 1,380 ms prior to the reward epoch. At 1,380 ms prior to the reward epoch, the correlation coefficient was 0.180, and at the time of the reward epoch it was 0.938. Thus there was a change of 0.758 over a period of 1,380 ms, a change of 0.055 per 100 ms. The correlation remained significant until 730 ms after the reward epoch, for a total of 2,110 ms (Fig. 4*B*).

Interestingly, the boxcar and the reward epoch vector were briefly anticorrelated during a period immediately after the choice. Specifically, the boxcars from 580 ms to 1,910 ms after the choice epoch showed significant negative correlation with the reward epoch's normalized activity vector (Fig. 4*B*). The significance of this negative correlation is unclear. This negative correlation reached its most extreme 1,150 ms after option appearance ($r = -0.325$). We thus compared when the boxcar's correlation significantly differed from this low point. We found that the boxcar's correlation with the reward epoch was significantly different from this low point beginning 2,290 ms prior to the reward epoch. If this point is used as the beginning of the significant rise in correlation, there is a change of 0.45 per 100 ms ($r = -0.085$ 2,290 ms before the reward epoch to $r = 0.938$ during the reward epoch). In other words, the positive increase in correlation began early in the trial and had to first overcome this early negative correlation before becoming significantly positive.

Reward size coding format changes between choice and reward epochs. Next, we investigated how the coding of reward size changed throughout the delay. Given that the
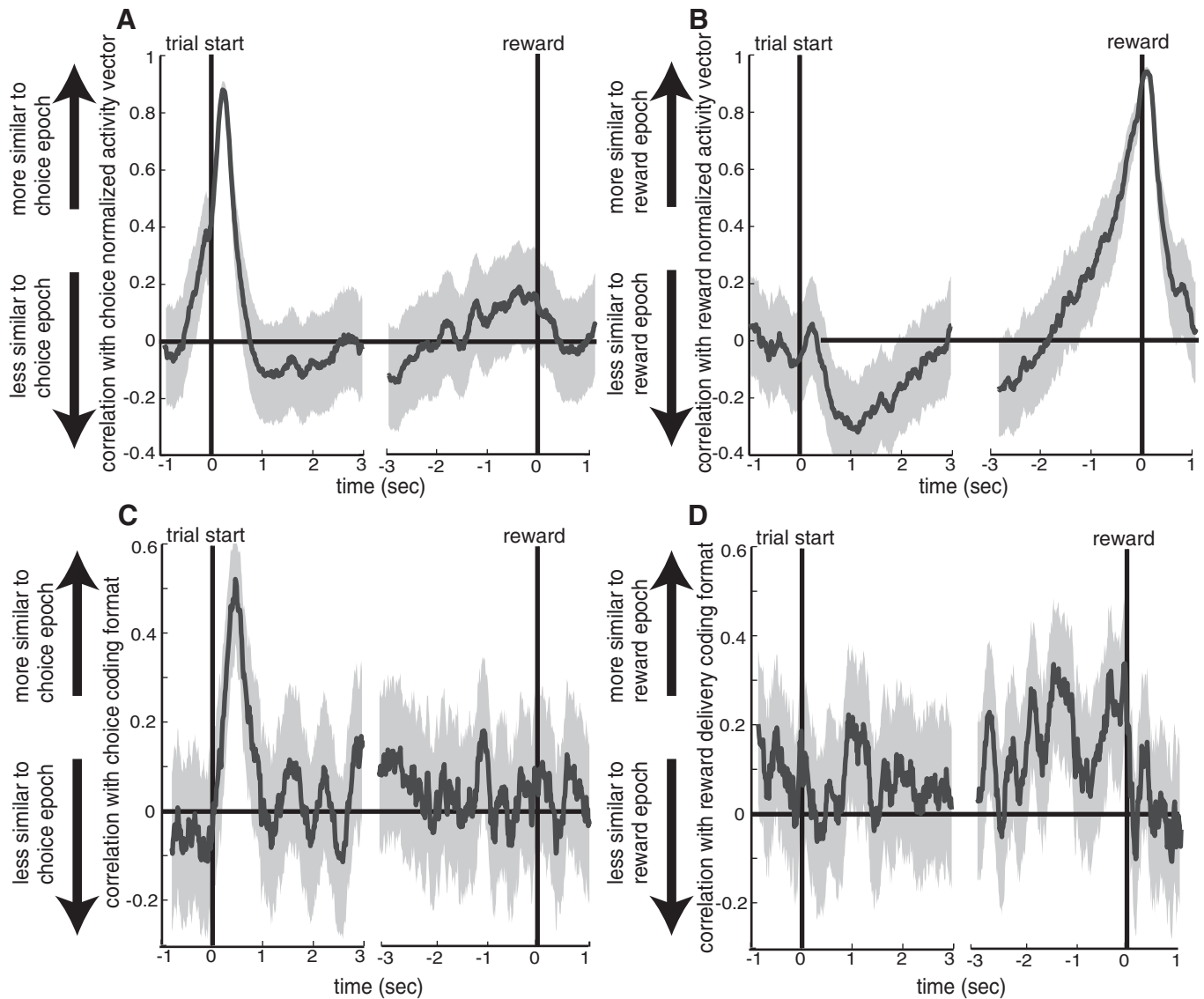
Fig. 4. Correlation between activity pattern vectors and coding format vectors shows gradual ramping toward reward activity pattern and coding format. *A* and *B*: ensemble activity pattern analysis: the correlation between the normalized activity vector from a running boxcar with the normalized activity vector from the choice epoch (*A*) and the reward epoch (*B*). The ensemble activity pattern changes rapidly after the choice epoch and gradually comes to resemble the reward epoch pattern. *C* and *D*: coding format analysis: the correlation between the coding format vector from a running boxcar with the coding format vector from the choice epoch (*C*) and the reward epoch (*D*). The coding format changes rapidly after the choice epoch and gradually comes to resemble the reward epoch format. Shaded regions surrounding plotted lines indicate 95% confidence intervals. Vertical lines indicate the start of the trial and the reward delivery.

proportion of neurons encoding reward size was similar between the choice and reward epochs, we wondered whether neurons used similar formats to encode this variable during these two epochs. To investigate this, we calculated the correlation between the coding formats for reward size during the choice and reward epochs. We used the regression coefficient from a linear regression of reward size onto firing rate as our measure of the coding format of an individual neuron (our term for the linear tuning properties of an individual neuron). Similar to our analysis using the normalized activity vectors to investigate ensemble activity patterns, we formed coding format vectors. Our coding format vectors were of length 121, with each element of the vector being a regression coefficient from a linear regression of firing rate onto reward size. We then calculated the correlation between these vectors.

It is worth emphasizing that our analysis approach differs in some ways from conventional analysis approaches, including our own past studies. Typically, we have focused on characterizing single neurons and then assessing the frequency of neurons similar to that in the population. Instead, here we ask what the population as a whole was doing, and what information it encoded. This approach has several advantages. First of all, it does not penalize neurons for diversity of response properties. Second of all, it is a more sensitive approach statistically. There are many cases where neurons show some effect but do not achieve significance individually, but show a strong effect when grouped together. The regression coefficient, even if not significant, remains the best estimate of a neuron's linear tuning to reward size. Third, the approach allows us to consider all neurons, rather than just a subset of significant ones, reducing the likelihood of double-dipping or voodoo correlations.

We found that the coding format vectors from the choice epoch and the reward epoch were not significantly correlated with each other (although there was a nonsignificant trend toward positive correlation, $r = 0.059$, $P = 0.052$, 95% CI: -0.120 to 0.235). Thus neurons in our population had no general tendency to respond similarly in the choice and reward epochs.

To further examine how coding format changed over time, we compared the choice epoch coding format vector to all other 200-ms windows in the trial (again using a sliding boxcar with 10-ms steps). To eliminate spurious correlations caused by shared neural noise, we continued to use the same cross-validation procedure as we did for the normalized activity vector analyses. Although much noisier, we found that the change in coding formats throughout the delay shared some of the same properties as the change in normalized activity vectors. The correlation between the sliding boxcar coding format vector and the choice epoch coding format vector only reached significance around the choice epoch (the boxcar beginning 210 ms before the choice epoch to the boxcar 280 ms after were significantly correlated with the choice epoch coding format; Fig. 4C). In other words, even though dACC encoded reward size throughout the delay (Fig. 2B), the tuning of individual neurons to reward changed after the choice epoch.

We then repeated this analysis but used the reward epoch as our fixed period. We found a significant correlation between sliding boxcar coding format vector and the reward epoch coding format vector prior to reward delivery (Fig. 4D). Specifically, we found a significant correlation beginning with the boxcar 570 ms prior to reward delivery. Further examination of the data revealed periods of significance prior to this, such as a period from 1,970 ms to 1,140 ms prior to the reward epoch. Thus we suspect that the lack of a significant correlation from 1,140 ms to 560 ms is likely attributable to noise (Fig. 4D). The high correlation prior to reward delivery indicates that dACC used similar coding formats to encode reward value before and after anticipated rewards. This lack of a shift in coding schemes suggests that dACC does not strongly differentiate anticipated from received rewards. One possible explanation is that dACC is not simply signaling reward value but may instead be generating signals involved in different types of control that correlate with reward value, and that these control processes are not particularly sensitive to the delivery of the reward itself (cf. Hayden et al. 2009, 2011a).

Timing of reward response. It is possible that our chosen reward epoch (0–200 ms after reward onset) may be too early to have included the neurons' responses to reward delivery and instead reflects anticipatory activity. To investigate this possibility, we plotted the populations' mean firing rate around the time of reward (Fig. 5A). This plot shows that the neural response to reward occurs rapidly and actually begins to drop off quickly after the time bin we used. Next, we investigated the proportion of neurons that significantly encoded reward (Fig. 5B). We found an odd dip in reward size encoding shortly after the 0–200 ms bin, and then a second peak slightly later. One possibility is that this first peak represents the peak of anticipatory activity, while the second is the peak of the actual reward response. The 100–300 ms bin occurs right in the middle of this second peak (whereas the 200–400 ms bin occurs after the peak). We therefore reran our key analyses,

using the 100–300 ms bin as our reward epoch. We find that this does not qualitatively change our results (Fig. 5, C and D).

Strategy adjustment following risky loss trials. We and others have previously argued that dACC responses instantiate explicit control signals that lead to a behavioral adjustment on the next trial (Botvinick et al. 2001; Hayden et al. 2011a, 2011b; Quilodran et al. 2008; Shenhav et al. 2013; Shima and Tanji 1998). In our task, monkeys exhibited greater reluctance to choose risky options after risky loss than after wins, suggesting that the brain generates control signals that adjust risk preferences after risky outcomes (0.223 rejection rate after risky win, 0.271 after risky loss; $P < 0.0001$, 95% CI: 6.9-2.8% difference in proportion, binomial 2-sample test; Fig. 6A). We hypothesized that behavioral adjustment signals may be present in ensemble activity in dACC.

Activity pattern changes in response to risky options. We next asked whether behavioral adjustment signals could be observed after outcome on risky trials (cf. Hayden et al. 2008, 2011a). After risky losses, neurons exhibited a sustained increase in firing rate compared with the firing rate after a risky win or safe reward (Fig. 6B; 6.84 spikes/s on risky wins, 7.30 spikes/s on safe trials, 8.86 spikes/s on risky losses). The normalized activity vector from risky loss trials during this time was negatively correlated with the normalized activity vector from the reward epoch of safe reward trials. Specifically, comparing the normalized activity vector for 800-1,000 ms after an unsuccessful risky outcome to the first 200 ms after a certain reward receipt, $r = -0.331$, $P < 0.0001$. We chose this epoch because it is when the response is strongest, but adjacent periods also showed significant negative correlations ($r = -0.233$, $P = 0.010$ and $r = -0.213$, $P = 0.019$ for 600–800 ms and 1,000-1,200 ms after reward receipt, respectively). There was also a significant negative correlation if the entire period from 500 to 1,500 ms after reward was used ($r = -0.258$, $P = 0.004$). This negative correlation means that neurons that tend to be excited compared with baseline during safe reward delivery will be suppressed after a risky loss, and vice versa, thus acting as a "no-reward" signal. We suspect that this negative correlation occurs some period after the outcome resolution because there was no explicit cue to the animals that they would not receive a reward. Instead, the lack of a reward had to be inferred from the combination of the visual stimulus disappearing and the lack of a liquid reward delivery.

These vectors were negatively correlated despite the mean population firing rates being very similar during these two periods (Fig. 6C). Specifically, the mean firing rate was 8.86 spikes/s for the risky loss trials vs. 8.82 spikes/s in the reward epoch following safe reward delivery. The similarity in the mean firing rates means that this signal would be invisible with aggregate measures of neural activity such as fMRI. The negative correlation we observe is consistent with the intuitive notion that risky losses and safe outcomes elicit opposing control signals, creating a possible link between the early ensemble ramping activity and control (Hayden et al. 2011b).

DISCUSSION

We measured responses of dACC neurons as monkeys performed a foraging task that required persistent commitment to a decision. We found that, throughout the delay prior to receiving a reward, the ensemble activity pattern in our popu-
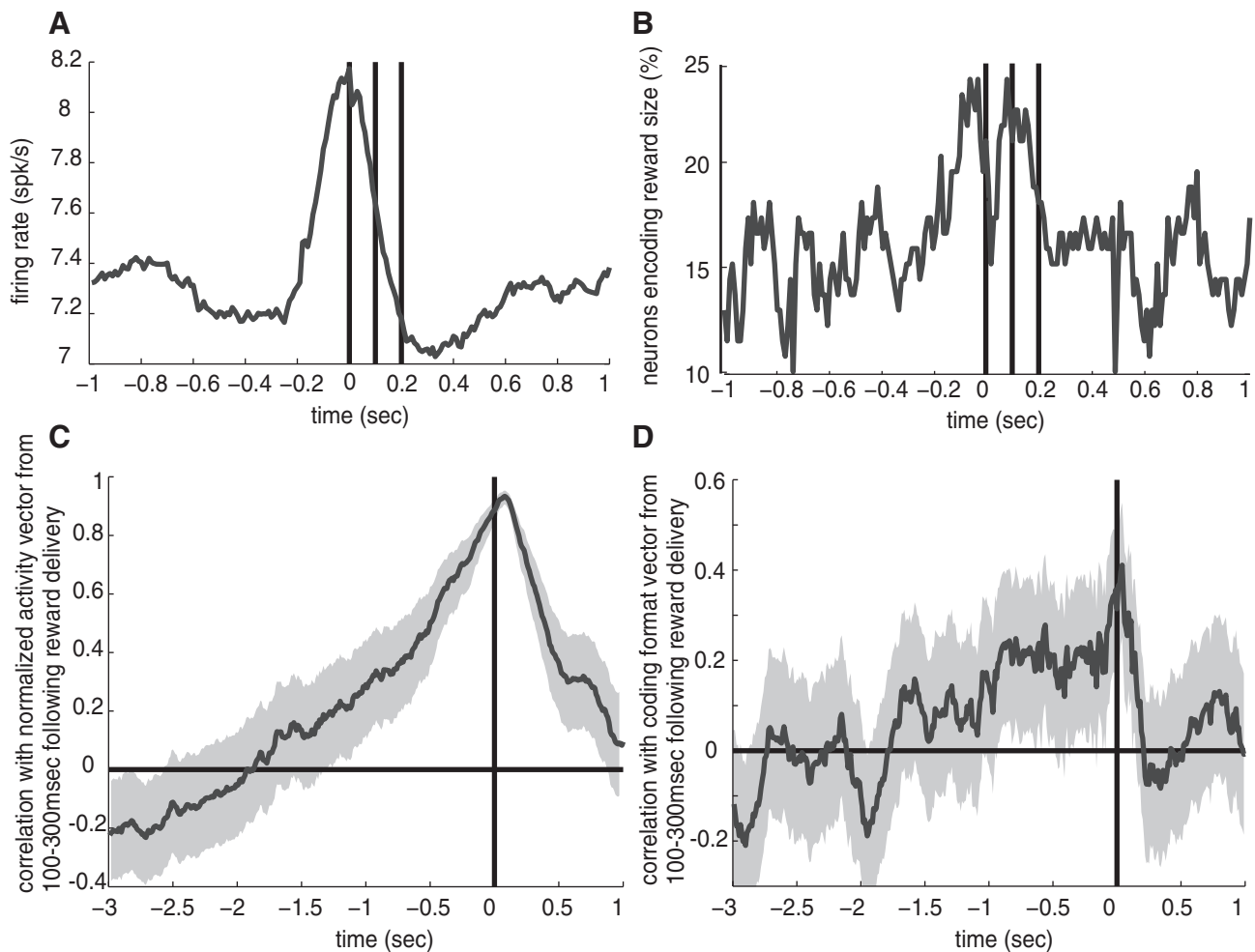
Fig. 5. Reward response timing and later reward epochs. *A* and *B*: timing of average population firing rate (*A*) and % of neurons that significantly encode reward size (*B*) in 200-ms sliding windows. *Time 0* indicates the time of reward. The timing is aligned to the beginning of each 200-ms bin. Thus the value of the graph at *time 0* represents the average firing rate in the bin 0–200 ms after reward delivery. Vertical lines indicate the 0–200 ms, 100–300 ms, and 200–400 ms bins. *C*: ensemble activity pattern analysis remains qualitatively the same when a later reward epoch is used: correlation between the normalized activity vector from a running boxcar with the normalized activity vector from the bin capturing the 100–300 ms period after reward delivery. Compare with Fig. 4*B*. *D*: coding format analysis remains qualitatively the same when a later reward epoch is used: correlation between the coding format vector from a running boxcar with the coding format vector from the bin capturing the 100–300 ms period after reward delivery. Compare with Fig. 4*D*. Shaded regions surrounding plotted lines indicate 95% confidence intervals. Vertical line indicates the time of reward delivery.

lation gradually came to resemble the ensemble activity pattern associated with receiving a reward. Although reward size was encoded during the choice and reward epochs, the coding format (meaning the ensemble tuning function) used to encode this variable was different during these two epochs. Much like the changes in the ensemble activity pattern, we found that the coding format of anticipated reward size quickly changed shortly after the choice epoch but came to resemble the coding format used during the reward epoch prior to the delivery of reward. Thus, despite the receipt of reward being a discrete event occurring on a very short timescale, neurons were responding to reward size just prior to the actual receipt of the reward similarly as they did immediately after reward receipt. In other words, there was no qualitative difference between anticipatory reward encoding and the encoding of a received reward. One possibility is that this emerging reward representation through the trial may serve to promote persistence (Hillman and Bilkey 2010).

Prior to the reward epoch, on the subset of trials with risky rewards, the ensemble of neurons showed a smooth ramping toward the pattern of activity evoked in response to a large and riskless reward. This ensemble ramping response was similar to the ensemble ramping found during safe trials. After a risky loss, ensemble activity pattern became not just uncorrelated but negatively correlated with the reward delivery state, meaning that neurons that were typically excited during reward delivery were typically inhibited after a risky loss, and vice versa. This rebound effect suggests that dACC encodes failure to receive a hoped-for reward by encoding its antithesis, one of the defining properties of reward prediction error (RPE) signals. Notably, this signal was not observed in average firing rates (and thus may potentially be invisible to aggregate methods like fMRI). The ensemble methods used in the present study allowed us to achieve a high-level view of how the population of neurons changed their activity over time. Instead of just investigating how and when individual neurons were changing, we instead were able to describe how the population shifted from one state to another. Thus it should be noted that the ensemble ramping activity we describe here is not necessarily the same as ramping activity observed in individual neurons. Instead, the en-
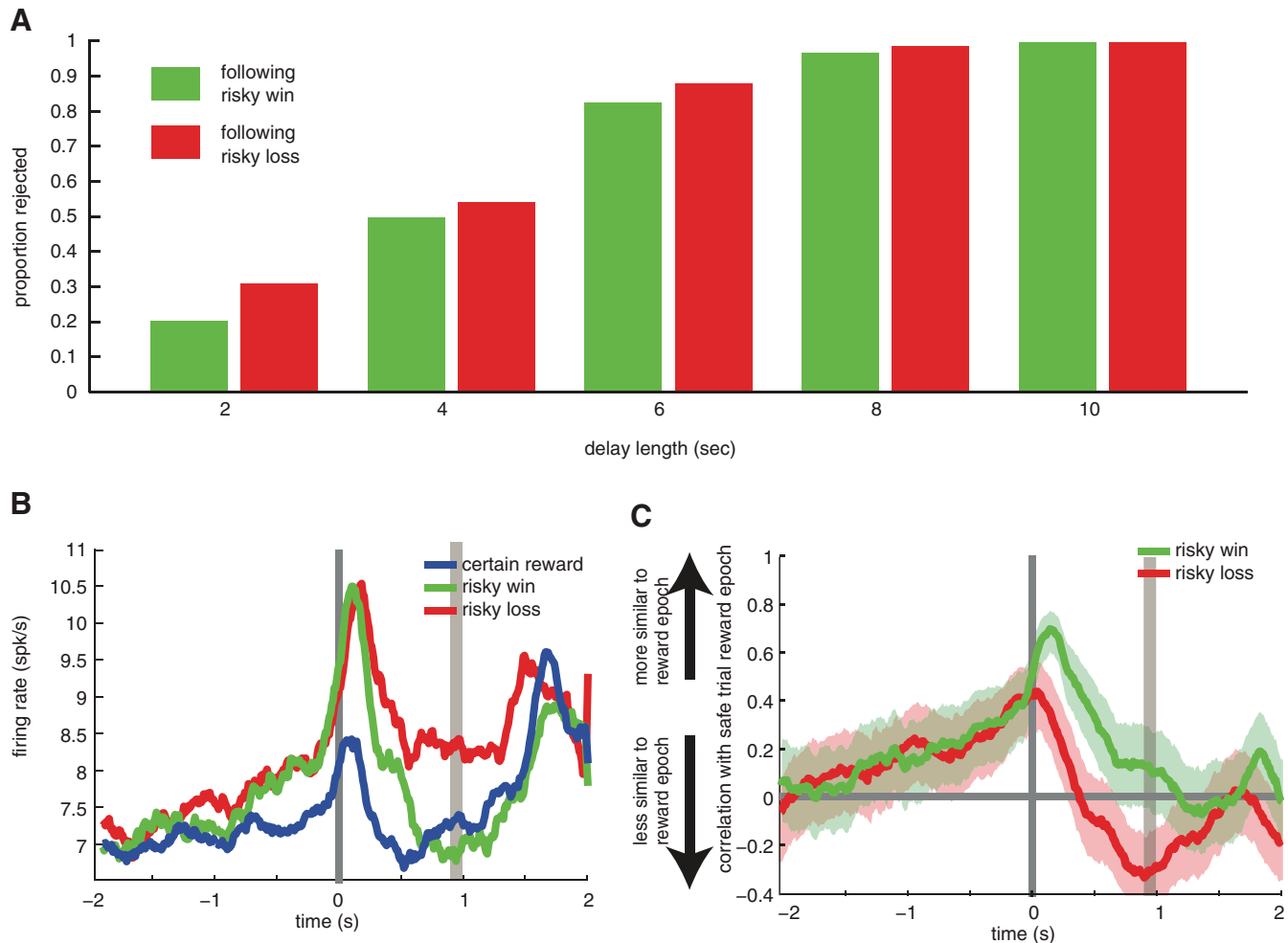
Fig. 6. Behavioral shift and no-reward signal after risky loss. *A*: acceptance rate of risky options after a risky win and a risky loss. Animals become more likely to reject a risky option after a risky loss. *B*: average firing rate of the population during trials with risky and safe options, aligned to reward delivery/risk resolution. Vertical lines indicate the start of the trial and the reward delivery. *C*: correlation between normalized activity vector from a running boxcar from risky trials with normalized activity vector from the reward epoch from large, safe reward trials, aligned to reward delivery/risk resolution. The correlation shows gradual rise leading up to risk resolution and a divergence following resolution. Unsuccessful risky outcome becomes negatively correlated 500-1,500 ms after reward delivery. Shaded regions surrounding plotted lines indicate 95% confidence intervals. Beige shaded period is the 200 ms analyzed further in the test.

semble ramping we describe here is a gradual change in ensemble activity pattern.

Throughout the trial, we saw a change in the ensemble activity pattern that ramped toward the ensemble activity pattern present during the reward epoch. The negative correlation we observed after a risky loss was negatively correlated with the ensemble activity pattern in the reward epoch. Thus we speculate that there may be a connection between these signals. Specifically, the positive correlation may be a signal that helps maintain the current status quo—a signal meant to maintain a behavior. The negative correlation may be the flip side of this signal—a signal meant to change the behavioral strategy (Hayden et al. 2011a, 2011b; Holroyd and Coles 2002; Quilodran et al. 2008).

When we strive to meet some goal that requires some time to achieve, it is necessary to maintain motivation to obtain the eventual reward. One method of maintaining motivation would be to maintain some representation of the reward as progress is made toward it, such that the progress itself is rewarding (Hillman and Bilkey 2010; Howe et al. 2013). Such a signal could potentially act as a control signal, keeping the animal

motivated to perform a particular action. A ramping ensemble reward function would ensure that when the reward is closest—when the least amount of investment is required to obtain the reward—the signal is strongest, and thus one is least likely to give up (Shidara and Richmond 2002). Thus a ramping ensemble anticipatory signal may be a general mechanism used by the brain to facilitate perseveration toward a goal.

Consistent with this idea, maintaining motivation toward a goal is the function that has been ascribed to recently observed ramps in dopamine (Howe et al. 2013). We speculate that the ensemble ramps in dACC may play a complementary role—given the noted function of dACC in behavioral control (Hayden et al. 2011a, 2011b; Kennerley et al. 2006; Quilodran et al. 2008; Rushworth et al. 2011), it is possible that ensemble ramps in dACC may implement an abstract control signal to have the animal continue to perform the actions required to achieve its goal, the strength of which may depend on the value of the goal being achieved (Holroyd and Coles 2002). Given dACC's dense projections to motor areas, this signal may then be translated into more concrete motor actions downstream (Paus 2001).

Alternatively, the anticipatory ramping ensemble signals observed in dACC may be a component of a reward anticipation circuit used to monitor the expected reward (Shidara and Richmond 2002). Dopamine neurons are known to encode the degree of reward expectancy, including when that expectation is spread out over time (Fiorillo 2003; Fiorillo et al. 2008; Schultz and Dickinson 2000). Recently, it has been suggested that dopamine ramps observed as animals approach a reward can be understood in the same framework (Gershman 2014; cf. Niv 2013). Thus our findings here of ramping ensemble signals in dACC as a reward approaches may be related to such a dopamine signal and play a role in monitoring expected rewards.

While in this study we focused on the responses of dACC neurons during the delay and at the time of reward, in a previous study using the same data set we analyzed the responses during the initial choice period (Blanchard and Hayden 2014). In that article, we reported that during the initial choice period dACC neurons responded qualitatively differently to our task variables depending on the behavior of the animal. Specifically, neurons encoded delay more strongly on accept trials than on reject trials and reward size more strongly on reject trials than on accept trials (although they did still significantly encode reward size on accept trials). One potential reason for the qualitative differences between accept and reject trials is that the same variables are being encoded on the two trial types to serve different functions. On a trial where the animal accepts the option, persistent fixation must be maintained, requiring a persistent control signal. We have suggested here that this control signal may be directly related to the representation of reward. Different types of control or learning processes may be required when the animal rejects an offer. For example, on reject trials dACC may encode an abstract "switch" signal to stop the animal from continuing to fixate and have it look away. Alternatively, dACC neurons may multiplex the action taken and the resulting outcome to drive learning (Camille et al. 2011; Hayden and Platt 2010).

Previous work has shown a link between dACC, persistence, self-control, and effort (Gusnard et al. 2003; Parvizi et al. 2013). However, little is known about exactly what the functional role of dACC in these processes is. On the basis of our findings here, we speculate that dACC may implement a control signal to allow for persistent commitment to a decision. This control signal may act to keep the animal progressing toward a goal. This potential function of dACC may suggest a framework for thinking about the neuroscience of persistence: to persist toward a goal, motivation and cognitive control must be maintained. These components may be implemented through a gradually arising reward representation.

DISCLOSURES

AUTHOR CONTRIBUTIONS

Author contributions: T.C.B. and B.Y.H. conception and design of research; T.C.B. performed experiments; T.C.B. analyzed data; T.C.B., C.E.S., and B.Y.H. interpreted results of experiments; T.C.B. prepared figures; T.C.B. drafted manuscript; T.C.B., C.E.S., and B.Y.H. edited and revised manuscript; T.C.B., C.E.S., and B.Y.H. approved final version of manuscript.

REFERENCES

Alexander WH, Brown JW. Competition between learned reward and error outcome predictions in anterior cingulate cortex. *Neuroimage* 49: 3210–3218, 2010.

Amiez C, Joseph JP, Procyk E. Reward encoding in the monkey anterior cingulate cortex. *Cereb Cortex* 16: 1040–1055, 2006.

Azzi JC, Sirigu A, Duhamel JR. Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc Natl Acad Sci USA* 109: 2126–2131, 2012.

Blanchard TC, Hayden BY. Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *J Neurosci* 34: 646–655, 2014.

Blanchard TC, Hayden BY. Monkeys are more patient in a foraging task than in a standard intertemporal choice task. *PLoS One* 10: e0117057, 2015.

Blanchard TC, Hayden BY, Bromberg-Martin ES. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* 85: 602–614, 2015.

Blanchard TC, Pearson JM, Hayden BY. Postreward delays and systematic biases in measures of animal temporal discounting. *Proc Natl Acad Sci USA* 110: 15491–15496, 2013.

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. *Psychol Rev* 108: 624–652, 2001.

Bush G, Vogt BA, Holmes J, Dale AM, Greve D, Jenike MA, Rosen BR. Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proc Natl Acad Sci USA* 99: 523–528, 2002.

Camille N, Tsuchida A, Fellows LK. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J Neurosci* 31: 15048–15052, 2011.

Chudasama Y, Daniels TE, Gorrin DP, Rhodes SEV, Rudebeck PH, Murray EA. The role of the anterior cingulate cortex in choices based on reward value and reward contingency. *Cereb Cortex* 23: 2884–2898, 2013.

Durstewitz D. Self-organizing neural integrator predicts interval times through climbing activity. *J Neurosci* 23: 5342–5353, 2003.

Fiorillo CD. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299: 1898–1902, 2003.

Fiorillo CD, Newsome WT, Schultz W. The temporal precision of reward prediction in dopamine neurons. *Nat Neurosci* 11: 966–973, 2008.

Floden D, Stuss D. Inhibitory control is slowed in patients with right superior medial frontal damage. *J Cogn Neurosci* 18: 1843–1849, 2006.

Gershman SJ. Dopamine ramps are a consequence of reward prediction errors. *Neural Comput* 26: 467–471, 2014.

Gusnard DA, Ollinger JM, Shulman GL, Cloninger CR, Price JL, Van Essen DC, Raichle ME. Persistence and brain circuitry. *Proc Natl Acad Sci USA* 100: 3479–3484, 2003.

Hayden BY, Heilbronner SR, Pearson JM, Platt ML. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J Neurosci* 31: 4178–4187, 2011a.

Hayden BY, Nair AC, McCoy AN, Platt ML. Posterior cingulate cortex mediates outcome-contingent allocation of behavior. *Neuron* 60: 19–25, 2008.

Hayden BY, Pearson JM, Platt ML. Fictive reward signals in the anterior cingulate cortex. *Science* 324: 948–950, 2009.

Hayden BY, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* 14: 933–939, 2011b.

Hayden BY, Platt ML. Neurons in anterior cingulate cortex multiplex information about reward and action. *J Neurosci* 30: 3339–3346, 2010.

Hillman KL, Bilkey DK. Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. *J Neurosci* 30: 7705–7713, 2010.

Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109: 679–709, 2002.

Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 500: 575–579, 2013.

Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9: 940–947, 2006.

Krebs JR, Erichsen JT, Webber MI, Charnov EL. Optimal prey selection in the great tit (*Parus major*). *Anim Behav* 25: 30–38, 1977.

Ma L, Hyman JM, Phillips AG, Seamans JK. Tracking progress toward a goal in corticostriatal ensembles. *J Neurosci* 34: 2244–2253, 2014.

MacEvoy SP, Epstein RA. Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Curr Biol* 19: 943–947, 2009.

McGuire JT, Kable JW. Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nat Neurosci* 18: 760–766, 2015.

Narayanan NS, Horst NK, Laubach M. Reversible inactivations of rat medial prefrontal cortex impair the ability to wait for a stimulus. *Neuroscience* 139: 865–876, 2006.

Narayanan NS, Laubach M. Top-down control of motor cortex ensembles by dorsomedial prefrontal cortex. *Neuron* 52: 921–931, 2006.

Narayanan NS, Laubach M. Delay activity in rodent frontal cortex during a simple reaction time task. *J Neurophysiol* 101: 2859–2871, 2009.

Niv Y. Neuroscience: dopamine ramps up. *Nature* 500: 533–535, 2013.

Parvizi J, Rangarajan V, Shirer WR, Desai N, Greicius MD. The will to persevere induced by electrical stimulation of the human cingulate gyrus. *Neuron* 80: 1359–1367, 2013.

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

Pearson JM, Hayden BY, Platt ML. Explicit information reduces discounting behavior in monkeys. *Front Hum Neurosci* 1: 237, 2010.

Peters J, Büchel C. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron* 66: 138–148, 2010.

Picton TW, Stuss DT, Alexander MP, Shallice T, Binns MA, Gillingham S. Effects of focal frontal lesions on response inhibition. *Cereb Cortex* 17: 826–838, 2007.

Quilodran R, Rothé M, Procyk E. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57: 314–325, 2008.

de Ridder DT, Lensvelt-Mulders G, Finkenauer C, Stok FM, Baumeister RF. Taking stock of self-control: a meta-analysis of how trait self-control relates to a wide range of behaviors. *Pers Soc Psychol Rev* 16: 76–99, 2012.

Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S. The importance of mixed selectivity in complex cognitive tasks. *Nature* 497: 585–590, 2013.

Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE. Frontal cortex and reward-guided learning and decision-making. *Neuron* 70: 1054–1069, 2011.

Sallet J, Quilodran R, Rothé M, Vezoli J, Joseph JP, Procyk E. Expectations, gains, and losses in the anterior cingulate cortex. *Cogn Affect Behav Neurosci* 7: 327–336, 2007.

Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci* 23: 473–500, 2000.

Shenhav A, Botvinick MM, Cohen JD. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79: 217–240, 2013.

Shidara M, Richmond BJ. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* 296: 1709–1711, 2002.

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

Stephens DW, Krebs JR. *Foraging Theory* (1st ed.). Princeton, NJ: Princeton Univ. Press, 1987.

Strait CE, Blanchard TC, Hayden BY. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* 82: 1357–1366, 2014.

Toda K, Sugase-Miyamoto Y, Mizuhiki T, Inaba K, Richmond BJ, Shidara M. Differential encoding of factors influencing predicted reward value in monkey rostral anterior cingulate cortex. *PLoS One* 7: e30190, 2012.

Wallis JD, Kennerley SW. Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. *Ann NY Acad Sci* 1239: 33–42, 2011.

Warden MR, Selimbeyoglu A, Mirzabekov JJ, Lo M, Thompson KR, Kim SY, Adhikari A, Tye KM, Frank LM, Deisseroth K. A prefrontal cortex-brainstem neuronal projection that controls response to behavioural challenge. *Nature* 492: 428–432, 2012.