Single caudate neurons encode temporally discounted value for formulating motivation for action

Yukiko Hori[1], Koki Mimura[1], Yuji Nagai[1], Atsushi Fujimoto[1], Kei Oyama[1], Erika Kikuchi[1], Ken-ichi Inoue[2], Masahiko Takada[2], Tetsuya Suhara[1], Barry J. Richmond[3] and Takafumi Minamimoto[1*]

1 Department of Functional Brain Imaging, National Institute of Radiological Sciences, National Institutes for Quantum and Radiological Science and Technology
4-9-1, Anagawa, Inage-ku, Chiba 263-8555, Japan

2 Systems Neuroscience Section, Primate Research Institute, Kyoto University, Inuyama, Aichi 484-8506, Japan

3 Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services
Bethesda, Maryland 2089-4415, USA

* Corresponding author
E-mail: minamimoto.takafumi@qst.go.jp (TM)

Abstract

Temporal discounting captures both choice preferences and motivation for delayed rewards. While temporally discounted value for choice is represented in brain areas including the dorsolateral prefrontal cortex (DLPFC) and the striatum, the neural process of motivation for delayed rewards remains unidentified. Here we show that neuronal activity of the dorsal part of the primate caudate head (dCDh) — a striatal region receiving projection from the DLPFC — signals temporally discounted value essential for computing motivation for delayed rewards. Macaque monkeys performed an instrumental task, in which a visual cue indicated the forthcoming size and delay duration before reward. Single dCDh neurons represented the temporally discounted value without reflecting changes in the animal's physiological state. Bilateral pharmacological or chemogenetic inactivation of dCDh specifically distorted a normal motivational performance based on the integration of reward size and delay. These results suggest a major contribution of dCDh to encoding a temporally discounted value, the integrated multidimensional information critical for formulating the motivation for action.

Introduction

Motivation for engaging in action depends on the expected value of its outcome, e.g., when and how much money or food will be available as a reward. Intuitively, the larger and earlier the reward, the greater the motivation will be. Animals and humans slower and less accurate when they expect the reward to be delayed. This decline in motivation is conceptualized as discounting of reward value as a function of time, namely temporal discounting [1, 2]. Temporal discounting was originally proposed to describe choice preferences for earlier smaller rewards rather than later larger rewards [3-5], implying

that motivation and decision-making may share the common brain process. Besides temporal discounting, motivational processes also consider internal drive for reward, such as hunger and thirst, integrating these two factors into motivational value [6-8].

Several lines of evidence suggest that the neural circuit involving the dorsolateral prefrontal cortex (DLPFC) and the striatum plays a major role in decision-making via temporal discounting of value. Human neuroimaging studies have shown that the structural and functional connectivity between these regions are associated with the individual difference in temporal discounting [9-11]. Neurons in DLPFC carry a representation of the temporally discounted value in monkeys [12] and in an avian analogue of PFC [13]. Additionally, disrupting DLPFC alters intertemporal choice behavior in humans [14] and non-choice behavior for delayed rewards in monkeys [15]. Neurons in the caudate nucleus are also known to respond in relation to temporally discounted values during intertemporal choice [16]. The dorsal part of the head of the caudate nucleus (dCDh), a striatal region receiving strong projections from DLPFC [17, 18], is well situated to be important for temporal discounting processes by providing a shared representation for both decision-making and motivation. Numerous studies have also reported that the caudate nucleus plays an important role in encoding an expected reward value, leading to motivation for an action with dopaminergic modulation [19-22]. However, it is not yet clear how dCDh contributes to the computation of motivational value with temporal discounting.

Here, we examined single unit activity in dCDh of macaque monkeys while they performed a delayed reward task. In the task a visual cue indicated the forthcoming

reward size and the delay duration to the reward after simple action. From each animal's behavior, we were able to infer the value for temporally discounted rewards including their interactions with satiation. We found that a subpopulation of single dCDh neurons increased their activity during the time period from the cue onset to the execution of action. The activity of many neurons was correlated with the temporally discounted value related to expected value of outcome. However, the activity was not influenced by the level of satiation. To determine whether the value-related activity might be causally related to the behavior, pharmacological inactivation (local muscimol injection) and chemogenetic inactivation (designer receptor exclusively activated by designer drugs, DREADD) [23, 24] of dCDh was carried out; both of these inactivations produced consistent impairments in motivational behaviors reflected as a distorted integration of reward size and delay, while behaviors based on the integration of reward size and physiological state remained intact.

Results

Temporal discounting accounts for monkeys' behavior

We studied computation of the motivational value using temporal discounting in macaque monkeys induced delaying reward delivery (Fig. 1A). In the basic task, the monkey must release a bar when a red spot turns green. A visual cue appears at the beginning of each trial and remains on throughout. Each of the 6 cues was linked to one combination of reward size (1 small drop; 3 or 4 large drops) and delay to reward (one of 0, 3.3, and 6.9 seconds; Fig. 1B). In this and similar tasks, the error rate, i.e., the proportion of trials with an incorrect response (either releasing the bar too early or too late), reflects the monkey's motivation for action, which can be interpreted as the motivational value or decision

utility for whether to act or not, according to its prediction about the forthcoming reward. In our previous studies, the error rate was inversely related to the subjective value [1]. In previous behavioral studies, the subjective value of delayed reward was formulated as a hyperbolic discounting model [3-5],

$$V = \frac{R}{1 + kD} \#(1),$$

where V is the value of delayed reward (i.e., temporally discounted value), R is the magnitude of reward, k is a discount parameter, and D is the delay to the reward. Accordingly, to describe error rates in this delayed reward task, we extended the inverse relation, incorporating it into a hyperbolic discounting model as shown in Eq. 2, with error rates (E), reward size (R), delay (D), and a monkey-specific free-fitting parameter (a) [1],

$$E = \frac{1 + kD}{aR} \#(2).$$

As shown in Figure 1C, the error rates were higher when a small reward size was expected, and for both reward sizes, the errors increased linearly as the expected delay duration increased. This pattern of the error rates was well described by the inverse relation with hyperbolic delay discounting (Eq. 2) (R^2 = 0.96, 0.88, and 0.91 for monkeys BI, FG and ST, respectively; Fig. 1C, solid lines). We verified that the hyperbolic model fitted the error rates significantly better than the exponential discounting model (Eq. 3; Fig. 1C, dashed curves) for all three monkeys as well as for seven additional monkeys ($p < 0.05$; cross-validation analysis), although the latter also explained well the majority of the cases (7/10 monkeys, $R^2 > 0.93$).

Fig 1. Task, behavioral performance and recording sites. (A) Sequence of events of behavioral tasks. (B) Example of relationship between cue and outcome in delayed reward task. (C) Ratio of error trials (mean ± sem) as a function of delay duration in monkeys BI, FG and ST. Data of small (1 drop) and large reward (3 or 4 drops) trials are indicated by black and red, respectively. Solid lines and dotted curves are best fit of Eqs. 1 and 3, respectively. (D) Series of coronal sections illustrating locations of recorded neurons plotted by dots. Anterior-posterior positions of sections (distance, in mm) are indicated by plus and minus numbers from anterior commissure (AC), respectively. Red, cue-responsive neurons with DV coding; Pink, cue-responsive neurons without DV coding; Gray, neurons without cue response. Coronal sections of CT-MR fusion image in top left visualize an electrode (*) in left caudate nucleus. CD, caudate nucleus; Put, putamen.

The reaction times also covaried with both reward size and reward delay; reaction times were shorter for larger rewards (two-way ANOVA; $p < 0.001$, 8/10 monkeys including monkeys BI, FG and ST) and shorter delays ($p < 0.001$, 9/10 monkeys including monkeys BI, FG and ST, S1A Fig). Although the monkeys were not required to fixate during the task, they usually gazed at the cue during the cue period. We did not find any significant effect of forthcoming reward size or delay duration on the duration of gazing at the cue (two-way ANOVA; main effect of reward size, effect size $\eta^2 = 0.004$; main effect of delay, $\eta^2 = 0.002$; reward size × delay, $\eta^2 = 0.003$; S1B Fig).

Together, these results suggest that the monkeys adjusted their motivation of action based on the temporally discounted value, which forms a hyperbolic relation between expected size and delay of forthcoming reward.

Neuronal activity of dCDh reflects temporally discounted value

We examined the role of the caudate nucleus in the motivational control of action based on the temporally discounted value. Specifically, we focused on the dorsal part of the caudate head (referred as 'dCDh'), the region receiving innervation from the ventral/dorsal part of the principal sulcus and anterior bank of the arcuate sulcus (S2 Fig). We recorded the activity of 150 presumed projection neurons (i.e., phasically active neurons; see methods) from dCDh (Fig. 1D) while the monkeys performed the delayed reward task. Most of the neurons (n=118) significantly increased their activity around more than one of three task phases: *cue* (immediately after cue appearance), *release* (at the time of bar release), and/or *reward* (at the time of reward delivery) (Figs. 2A-C; $p < 0.05$, χ^2 test). The cue response was the most prominent activity of dCDh neurons during the task (Figs. 2A and 2C); the proportion of cue-responsive neurons (100/150) was significantly larger than that of release-responsive neurons (49/150; $p < 0.01$, χ^2 test) and reward-responsive neurons (Fig. 2D; 49/150; $p < 0.0001$, χ^2 test).

Fig 2. Task-related responses of CD neurons. (A) Example of a neuron that responded exclusively to cue. Rasters and spike density histograms for all trials are aligned at the cue signal (left), bar release (middle) and reward delivery (right). Rasters are shown in order of occurrence of trials from bottom to top. Shaded areas are time windows when discharge probability is significantly higher than baseline ($p < 0.05$, χ^2 test). (B) Example of a neuron that responded exclusively to reward delivery. (C) Example of a neuron that responded to cue, bar release and reward delivery. (D) Distribution of neurons that responded in three task phases shown in Venn diagram. Numbers in parentheses represent

numbers of neurons showing significant response to each event. The proportions of responded neurons in each monkey are as follows: Cue, 88%, 88%, and 83%; Release, 37%, 42%, and 46%; Reward, 41%, 50%, 38%; for monkeys BI, FG, and ST, respectively.

Some of the cue responses signaled a temporally discounted value (DV) of the forthcoming reward (Eq. 1). An example of the activity shown in Figure 3A exhibited the strongest activation after the cue associated with a large and immediate reward. The cue response became smaller as the delay duration became longer, and with the smallest reward with long delay, the neuron did not respond at all. The neuron presented in Figure 3B had the opposite response pattern; the activation was stronger when the cue predicted smaller rewards with longer delays. We related spike discharge rates to DV estimated using the hyperbolic function obtained from individual behavior (Eqs. 1 and 2). The firing rate during the cue period of example neurons (Figs. 3A and 3B) correlated with DV positively (Fig. 3C, $R^2 = 0.86$, $p < 0.01$) and negatively (Fig. 3D, $R^2 = 0.77$, $p < 0.05$). A significant regression coefficient for DV ($p < 0.05$, t-test) was found in 27 of 100 cue-responsive neurons (11, 6, and 10 in monkey BI, FG, and BI, respectively); 18 and 9 exhibited positive and negative correlations, respectively. In contrast, significant DV relation was relatively minor in release-related (5/49) and reward responses (3/49).

Fig 3. Cue responses of temporally discounted value coding. (A-B) Activity of example neurons during cue period. Rasters and spike density histograms are aligned at cue onset. The color corresponds to each reward condition. Rasters are shown in order of occurrence of trials from bottom to top in each condition. Shaded areas on rasters are time windows for evaluating the magnitude of cue response. (C-D) Relationship between firing

rate (mean ± sem) and temporally discounted value (V, Eq. 1) for neuronal activities shown in (A) and (B), respectively.

Besides the DV relation, the cue response might solely reflect reward size or delay duration. We compared the effect of size or delay alone on cue response with that of DV using multiple linear regression analysis (Eq. 8). We found that only 3 and 4 neurons showed a significant exclusive effect of size or delay on their cue response, respectively (Figs. 4A and 4B, blue and green). In contrast, 19 and 5 neurons had a significant effect of DV or both delay and size, respectively (Figs. 4A and 4B, red and pink), the proportions of which were significantly larger than that of neurons by chance coding both delay and reward size ($p < 0.01$; χ^2 test). The strength of size or delay effect was relatively smaller than that of DV. Thus, DV-related neurons were not just a selected population from the neurons representing mixtures of these delay and size by chance; rather, the entire neuronal population seemed to represent reward size and delay in an integrated manner. Such population level DV-relation was also observed in the release response, but not in the reward response (S3 Fig).

Fig 4. Impact of DV and comparison with delay and size on cue response. (A-B) Scatterplots of standardized partial regression coefficients (SPRC) of DV (ordinate) against those for reward size or delay (abscissa) for discharge rates during cue period, respectively. Colored dots indicate neurons with significant ($p < 0.05$) coefficient, while gray dots correspond to neurons without any significant effect (NA). DV/DV & Other, neurons with significant coefficient of DV; Size & Delay, those with both size and delay; Size, those exclusively with size; Delay, those exclusively with delay. Numbers in

parentheses indicate number of neurons.

Together, our results suggest that the temporally discounted value of the forthcoming reward is represented in dCDh neurons, that is, mainly in a subpopulation of neurons. In the following section, we will focus on this subset of neurons, and refer to neurons with and without significant correlation with DV as *DV coding neurons* (n = 27) and *non-DV coding neurons* (n = 73), respectively. DV coding neurons were not confined to specific locations, but were found throughout the CD (Fig. 1D).

To quantify the time course of DV coding of the cue responses, the effect size of DV (R^2) in a linear regression analysis (Eq. 7) was calculated (200-ms window, 10-ms steps) for each DV coding neuron (Figs. 5A and 5B). On average, the effect size rises from 100 ms after cue onset, reaching a peak at 750 ms after the cue (red curve, Fig. 5C). Thereafter, it gradually decreased to the bar release (Fig. 5D). The effect size does not become 0, indicating that a few neurons (n = 5) also signaled DV around bar release. In non-DV coding neurons, the effect size was only slightly changed from 0 during the cue period, whereas it increased after the bar release (black curve, Figs. 5C and 5D). Thus, DV coding started just after the monkey was informed about the reward size and delay of the forthcoming reward, and it continued until the time point of execution of an action.

Fig 5. Time course of DV coding. (A-B) Time-dependent change of DV coding. Each row represents color-coded effect size (R^2) of DV for a single DV coding neuron. Responses were aligned by cue onset and bar release, respectively. (C-D) Time-dependent change of effect size of DV for DV coding (red) and non-DV coding neurons

(black) aligned by cue onset and bar release, respectively. Thick curve and shaded areas indicate mean ± sem, respectively. Arrows indicate time of go signal (first 3 of 5 with variable interval).

DV-coding is insensitive to satiation

The motivational value of reward should decrease as the physiological drive state changes from thirst to satiation. In every daily session, the monkeys were allowed to work until they stopped by themselves, meaning that the data were collected as the monkeys were approaching satiation. As the normalized cumulative reward (R_{cum}) increased, the overall error rate in each combination of reward size and delay also increased (Fig. 6A). When we looked at the data from one quarter (e.g., Fig. 6B, R_{cum} = 0.75 - 1), the error rate increased linearly as the delay duration increased with each reward size. These observations were well in accordance with the psychological concepts of incentive motivation assuming a multiplicative interaction between the value of outcome (i.e., discounted value) and the satiation effect [6-8] (Eq. 6；see methods). The error rates were well explained by Eq. 4 for each individual monkey (R^2 = 0.89 ± 0.06; mean ± sem) as well as for the average across 9 monkeys (R^2 = 0.98, Figs. 6A and 6B). The satiation effect, $F(R_{cum})$ (Fig. 6C), indicated that the motivational value of reward decreased at a rate of more than 15% (16%, 33% and 17% for BI, FG and ST, respectively) in a single recording session (i.e., 120 trials) according to the number of average success trials in a daily session.

Fig 6. Negligible effect of satiation on DV-coding. (A) Ratio of error trials (mean ± sem) as a function of normalized cumulative reward (R_{cum}) on average across 9 monkeys.

Dotted curves are the best fit of Eq. 4 to the data. (B) Error rates (mean ± sem) as a function of delay duration for each quarter of R_{cum}. (C) Satiation function, $F(R_{cum})$ along with R_{cum} in 3 individual monkeys. Since average total trials were 934, 512 and 493 in BI, FG and ST, motivational value became 84%, 67% and 83% through 120 trials (i.e., 16%, 33% and 17% devalued, respectively). (D) Example of comparison of cue responses in 1st and 2nd half of recording period for each reward condition in single dCDh neuron (monkey ST). Spike density histograms are aligned at cue onset; 1 and 3 drops in reward size, respectively. (E) Comparison of cue responses in 1st and 2nd half of recording period for each trial type in positive DV-coding neurons (n= 18). Responses were normalized by firing rate of cue response in immediate large reward trials during 1st half of the period.

Although satiation significantly influenced behavior, satiation did not influence dCDh activity, even when coding DV. When we compared the cue responses between the 1st and 2nd half of 120 successful trials, the activity patterns were indistinguishable between the 1st and 2nd half of the recording period in a single neuron (Fig. 6D). Similarly, the normalized mean discharge rate of cue responses for each reward condition did not significantly change between the 1st and 2nd half in 18 positive DV-coding neurons (repeated measures two-way ANOVA; main effect of trial type, $F_{(5, 119)} = 16.8$, $p < 10^{-8}$; main effect of satiation, $F_{(1, 119)} = 1.7$, $p = 0.29$; Fig. 6E). Additional neuron-by-neuron analysis using a multiple linear regression model (Eq. 9) demonstrated that a significant satiation effect was not found in any of the cue responsive dCDh neurons (97/100) except for 3 non-DV coding neurons (S4 Fig). Therefore, dCDh neurons encode the expected temporally discounted value in their cue response without reflecting internal

physiological drive.

Inactivation of dCDh specifically impairs behavioral pattern to delay discounting

In our results, the activity of a subset of dCDh neurons encoded DV after the cue, but not reward size or delay alone. This raises the question of whether the activity is needed to judge the values reflected by DV. To test this, we inactivated bilateral dCDh by local injection of muscimol ($GABA_A$ receptor agonist) or by a chemogenetic technology (DREADDs), two complementary methods to produce the comparable behavioral change when applied to the primate striatum [25]. Two monkeys had muscimol injected locally into the dCDh, which was confirmed by CT images of injection cannulae overlaying MR images, matching with the recording sites (Figs. 7A and 7B; see Figs. 1D and S2D Fig for comparison). Another monkey received injections of a viral vector expressing an inhibitory DREADD, hM_4Di, into the dCDh bilaterally. A positron emission tomography (PET) scan with a DREADD-selective radioligand, [^{11}C]deschloroclozapine (DCZ) [23], confirmed that hM_4Di expression covered the dCDh (Fig. 7C). Chemogenetic silencing was achieved by systemic administration of the selective DREADD agonist DCZ [23]. Both pharmacological and chemogenetic inactivation resulted in a significant shift in error rate patterns with respect to reward size and delay in all three monkeys (Fig. 7D, left); the behavioral patterns were idiosyncratic across the monkeys, but they were generally not in accordance with the temporal discounting model (i.e., Eq. 2; R^2 = 0.41, 0.19 and 0.76, for monkeys BI, RI and ST, respectively). In contrast, the error rate pattern following vehicle injection remained well explained by the model (Fig. 7D, R^2 > 0.86).

Fig 7. Bilateral inactivation of dCDh disrupted normal motivational performance based on size and delay. (A) CT-based localization of muscimol injection sites. CT image visualizing injection cannulae targeting caudate nucleus bilaterally (hot color) overlaid on MR image (gray scale) in monkey BI. Scale bar, 5 mm. (B) Injection sites located 3 mm anterior and 3 mm posterior to anterior commissure (AC) are mapped. Magenta and blue circles indicate injection sites of muscimol and saline, respectively. Session numbers are shown on the circles. (C) [^{11}C]DCZ-PET visualizing hM_4Di expression *in vivo* in monkey ST. Parametric image of specific binding (BP_{ND}) of [^{11}C]DCZ-PET overlaying MR image. Scale bars, 5 mm. (D) Error rates (mean ± sem) as function of delay duration under inactivation (left) and control condition (right). Black and red symbols are low and high reward trials, respectively. Dotted lines represent best-fit of function of hyperbolic temporal discounting (Eq. 2). (E) Distribution of sum of squared residuals (SSR) of best-fit function (Eq. 2) to averaged resample data obtained by bootstrap method (n=20,000). Blue and red lines indicate SSR of best-fit of Eq. 2 to mean error rates in control and inactivation sessions, respectively.

Despite changing error patterns, inactivation did not produce statistically significant changes in the overall error rates (inactivation vs. control; two-way ANOVA for treatment × reward condition; main effect of treatment, $F_{(1, 2)} = 13.6$, $p = 0.07$; interaction, $F_{(5, 164)} = 2.1$, $p = 0.07$). Apart from the error rates, the inactivation did not affect other behavioral parameters. The total reward earned during the task was unchanged (one-way ANOVA; treatment, $F_{(1, 26)} = 2.8$; $p = 0.11$, S5A Fig). The reaction times were not prolonged by inactivation in monkeys BI and RI (two-way ANOVA, effect size of treatment, $\eta^2 < 0.01$) but became shorter in monkey ST (effect size of treatment, $\eta^2 = 0.06$), while they

continued to reflect the reward condition (effect size of reward condition, $\eta^2 > 0.06$ in all monkeys, S5B Fig). The duration of gazing at the cue was slightly but not significantly longer during muscimol inactivation (t-test, $p = 0.063$, S5C Fig). Together, the bilateral inactivation of dCDh did not cause impairments in overall motivation or motor performance.

These results demonstrated that dCDh inactivation appeared to produce a consistent impairment, namely, alteration of error rate pattern without changing overall error rates. To quantify behavioral deviation from normal temporal discounting, we normalized the error rates in each session for baseline, inactivation and control condition (S6 Fig). Bootstrap analysis revealed that, compared to baseline data, inactivation, but not control, caused significant deviations in the error rate patterns away from the temporal discounting model in all monkeys ($p < 0.05$, Fig. 7E, red line), suggesting that dCDh silencing distorted normal motivational value processing based on the integration between reward size and delay.

What do the distorted error patterns indicate? Fig. 8 illustrates simulated error rate patterns due to alterations of parameter or variables comprising temporal discounting. Greater temporal discounting (i.e., increase of parameter *k*) will lead to an increase in the steepness of the slope of linear relation, causing increases in overall error rates (Fig. 8A). Alteration of reward size estimation will cause a change in the ratio of error rates between high and low reward trials (Fig. 8B). Alteration of delay estimation will induce a non-linear relationship across delay duration (Fig. 8C). Alterations of both reward size and delay estimation will produce an error pattern with a twisted, non-linear relation with

delay for two reward-size conditions in parallel (Fig. 8D). However, the error patterns observed during the inactivation of dCDh did not comply with any of these 4 simulated patterns. Instead, the patterns reflected an altered interactive estimation between reward size and delay (Fig. 8E). Thus, the behavioral change due to dCDh inactivation cannot be explained solely by the changing steepness of temporal discounting, or by altered evaluation of reward size, delay duration, or both. Instead, it is attributable to impaired interactive calculation between reward size and delay duration, namely, temporal discounting.

Fig 8. Simulation of disrupted error rate pattern due to alteration of parameters or variables comprising temporal discounting

Black circles and red triangles represent simulated data for R = 1 and 3, respectively. Dotted lines indicate a representative best-fit function (Eq.2) from original data (monkey ST). (A) An error rate pattern in a case of a change in steepness of temporal discounting, where k becomes 1.5 times of original ($k' = 1.5 \times k$). (B) An error rate pattern in a case of altered reward size sensitivity, where reward size [R = (1, 3)] is considered as if R' = (2, 2.5). (C) An error rate pattern in a case of altered delay sensitivity, where delay [D = (0, 3.3, 3.9)] is perceived as if D' = (0, 4, 4). (D) An error rate pattern in the case of both C and D simultaneously occurs. (E) An error rate pattern in the case of a conditional (interactive) change in sensitivity to reward size and delay, where the combinations of reward size and delay [(R, D) = (1, 0), (1, 3.3), (1, 6.9), (3, 0), (3, 3.3), (3, 6.9)] are recognized as if (R', D') = (1, 0), (1, 4), (1, 4), (2.5, 0), (2.5, 3.3), (2.5, 6.9).

Was the impairment specifically related to the temporally discounted value? Alternatively,

it may reflect the dysfunction of the motivational process in general. Since the temporally discounted value is often referred to as 'subjective value', dCDh inactivation could produce a general dysregulation of computation for motivational value — a subjective impact of the upcoming reward on performance. To examine the effects of dCDh inactivation on motivational value without delay, we tested two monkeys in a reward-size task in which the task requirement remained the same as the delayed reward task, but a successful bar release was immediately rewarded with one of four reward sizes (1, 2, 4, or 8 drops) associated with a unique cue (Fig. 9A). It has been repeatedly shown that the error rates of this task are well explained by the joint function of reward size and satiation (Eq. 6) [1, 26, 27]. Pharmacological or chemogenetic inactivation of bilateral dCDh did not alter the pattern of the error rate in this task; in both cases they remained to be well explained by the model (Fig. 9C, $R^2 > 0.7$) and were equally well compared with the baseline data ($p > 0.15$, bootstrap significance test; Fig. 9D). The inactivation did not change the overall error rates (three-way ANOVA, treatment, $F_{(1, 243)} = 1.35$, $p = 0.45$) or the interactive effect with reward size on the error rates (treatment × size, $F_{(3, 243)} = 1.69$, $p = 0.17$). These results suggest that, unlike for delay discounting, the motivational value computation for predicted reward size does not rely on the dCDh activity.

Fig 9. Reward-size task and behavioral performance. (A) Cue stimuli used in reward-size task uniquely associated with forthcoming reward size (1, 2, 4, or 8 drops). (B) *left:* Error rates (mean ± sem) as function of reward size in muscimol treatment (magenta) and non-treatment control session (black) for monkey RI, respectively. *right:* Error rates (mean ± sem) as function of reward size after DCZ treatment (red) and after vehicle treatment (black) for monkey ST, respectively. Dotted curves represent best-fit of inverse

function. (C) Error rates (mean ± sem) as function of normalized cumulative reward (R_{cum}) for monkeys RI (top) and ST (bottom), respectively. Each reward size condition was shown in a different color. (D) Distribution of sum of squared residuals (SSR) of best-fit function (Eq. 6) to averaged resample data obtained by bootstrap method (n=20,000). Blue and red lines indicate SSR of best-fit of Eq. 6 to the mean error rates in control and inactivation sessions, respectively.

Discussion

In the present study, we investigated the role of dCDh in formulating the motivational value of expected delayed rewards. The behavior showed that the likelihood of carrying out the trials for delayed rewards was well described by a model with hyperbolic discounting and satiation. There were two main findings. First, a substantial number of single dCDh neurons represented the temporally discounted values, combining the information about the reward size and delay in delivery. However, these same neurons did not reflect a decrease in internal physiological drive seen in the behavior as the monkeys became more satiated. Second, bilateral pharmacological or chemogenetic inactivation of dCDh distorted the motivational valuation derived from the integration of reward size and delay duration, whereas the relationship from the integration of reward size and physiological state remained intact. These results suggest a major contribution of dCDh in mediating the integrated external information that is critical for formulating the motivation for action.

Neurons in dCDh are sensitive to intertemporal choice, and specifically in relation to differences in the temporally discounted values for alternative targets [16]. This neuronal signal was similar to that observed in the dorsal striatum, reflecting the action values of a specific movement [16, 28-30]. This implies that dCDh should contribute to selecting an action to maximize future rewards. We found that some of the cue responses of dCDh neurons reflected temporally discounted value inferred from individual behavior in our task. This cannot be a simple reflection of the physical features of a visual cue, since the neuronal signal was observed irrespective of cue sets used for assigning delayed reward, and since the neuronal correlates disappeared when the cue was randomized with respect

to the outcome (data not shown). The DV signal emerged just after cue onset, gradually increased, and then disappeared before execution of the action (*cf.* Fig. 5). This time course suggests that the neuronal signal is not a simple representation of the Pavlovian value, but is related to the process mediating the outcome prediction underlying the decision whether to act or not. Similar neuronal modulation reflecting expected reward value (e.g., presence, amount, or kind of reward) has been reported in the dorsal striatum [19, 20, 22].

Compared with DV, the effect of the reward size or delay duration on cue responses was relatively weak (cf. Fig. 4), indicating that the signal integration does not occur in dCDh but rather in some upstream brain area(s). The first plausible source of temporal discounting is a prefronto-striatal projection. Our recordings were carried out from the head of the caudate nucleus, corresponding to the region receiving direct input from the DLPFC (cf. S2 Fig) [31-33]. DLPFC neurons encode DV as well as reward, delay duration, and target position during an intertemporal choice task [12, 13], exhibiting strong modulations in response to the delay combined with the amount of reward [34]. The second possible source is a nigrostriatal dopaminergic input. When a stimulus signaled the timing of reward delivery, the stimulus response of dopaminergic neurons declined hyperbolically with the delay duration [35]. The third possible source is a thalamostriatal input arising from the thalamic nuclei, including the centromedian-parafascicular (CM-Pf) complex [36]. Neuronal activity in CM reflects the predicted outcome value [37], but at this point there is no evidence that it is involved in delay discounting.

Independent of temporal discounting, the motivational value of the cue should also decrease according to a shift in the internal physiological drive state. However, the effect of drive has been investigated separately from temporal discounting, and it has generally not been taken into account during studies of choice behavior. In our task, the changing motivational value was well approximated as being exponentially decreased along with reward accumulation (*cf.* Fig. 6C), while the relative effect of reward size and delay on decision appeared to be constant. This was in good agreement with psychological concepts of motivation, in which motivational value arises from a multiplicative interaction between external stimulus and physiological state [6, 7]. This also suggests that temporal discounting and reward devaluation may be two independent processes, one exerting a hyperbolic effect of delay duration on the reward size changing in a trial-by-trial manner, and the other slowly decreasing the motivational value of the reward in response to reward accumulation. Our data support the notion that dCDh may be involved in the former process only; DV coding in dCDh was not sensitive to changes in internal drive (*cf.* Fig. 6 and S4 Fig). A similar insensitivity to satiation has been reported in terms of cue-related neuron activity in the ventral striatum that was correlated with reward value [38]. This leaves an intriguing possibility, namely, that the insensitivity of internal drive may result from the motor output used; different data could be obtained if we tested monkeys with saccadic eye movements, in which neurons in dCDh are known to be involved. Satiety-dependent changes in neuronal activity have been seen in the orbitofrontal cortex (OFC), ventromedial prefrontal cortex [39-41], rostromedial caudate nucleus (rmCD) and ventral pallidum (VP) [26]. Perhaps satiety-related signals would be represented in a network believed to be critical for guiding a choice of food based on internal drive [42, 43]. To formulate the motivational value for action, the physiological

state or drive signal from this network may be integrated with temporal discounting in the basal ganglia-thalamocortical circuit, brain structures downstream from dCDh.

The causal contribution of DV coding in dCDh to temporal discounting was examined by pharmacological and chemogenetic inactivations, which are complementary and applicable to silencing primate striatal activity [25]. Muscimol inactivation is a standard procedure that has repeatedly been used in monkey studies. It has, however, major drawbacks: (1) the extent of an effective area is difficult to be controlled or identified (although we monitored the location of injection sites by computed tomography (CT)), and (2) when the experiments are repeated, mechanical damage to tissue would accumulate. The chemogenetic tool DREADD, on the other hand, overcomes these problems; once a silencing DREADD, hM_4Di, is delivered, substantially the same neuronal population can be inactivated non-invasively and the effective region can be confirmed by PET imaging, as demonstrated here, and by traditional post-mortem histochemistry. We found that inactivation of dCDh by either method produced consistent behavioral impairments; inactivation abolished the normal pattern of error rates derived from the integration of reward size and delay duration (cf. Fig. 7). This impairment cannot be explained simply by changes in the temporal discounting rate or alterations in the evaluation of single incentive factors. Our results are consistent with previous findings that both lesioning and inactivation of the dorsomedial striatum in rats, a homologue of dCDh in primates, reduced the sensitivity of instrumental performance to shifts in the outcome value [44, 45]. In contrast, dCDh inactivation did not impair motivation based on reward size alone or according to the integration of reward size and physiological state (i.e., motivational value; cf. Fig. 9). Thus, the impairment can be attributed to the loss of

DV coding seen in the activity of single dCDh neurons. Similar specific impairments have also been found in monkeys with bilateral ablation of DLPFC [15]. Given intact motivational evaluation for the reward size alone in these cases, the motivational process appears to gain access to value signals bypassing the DLPFC-CD pathway. A plausible network for the reward size process is prefronto-basal ganglia projections from OFC to rmCD/ventral striatum and/or VP [18], since ablation or inactivation of these related areas abolished the normal relationship between reward size and error rate in the reward-size task [15, 25, 26]. Therefore, our findings, together with our previous results, support the concept that incentive motivation is processed through the prefronto-basal ganglia circuit in accordance with certain topographic organization [46, 47]. Our findings additionally provide evidence that defines a specific role of dCDh in incentive motivation, as dCDh signals the integrated multi-dimensional factors and contributes to computation of the motivational value.

Our findings may also have some clinical relevance. Dysregulation of normal temporal discounting is associated with increased impulsive behavior. Impulsive behavior and preference are often manifested in patients with psychiatric disorders, including depression, schizophrenia, bipolar disorders, obsessive-compulsive disorders, and substance use disorders [48, 49]. Human imaging studies have revealed the structural and functional connectivity between DLPFC and the striatum with the individual differences in temporal discounting [10, 11]. Since silencing dCDh did not induce impulsivity (steepened temporal discounting or facilitating reaction was not observed), it could be difficult in the present study to address the link between dCDh activity and mechanisms underlying impulsivity. Nevertheless, our findings may provide a framework to elucidate

dysregulation of motivational systems in impulsive individuals with psychiatric disorders.

In summary, our work indicates that dCDh neurons encode, at a single-neuron level, temporally discounted values of forthcoming rewards without reflecting any internal state alteration. These signals are likely to be used in the downstream brain structures for formulating motivation of action especially when multi-dimensional factors have to be jointly evaluated.

Materials and Methods

Subjects

Eleven macaque monkeys (5 -11 kg; 10 male rhesus monkeys and one female Japanese monkey) were used in this study. Ten monkeys were used for behavioral testing. Of these, three (BI, FG and ST) were also used for recording, and one (ST) and two (BI and RI) for chemogenetic and pharmacological inactivation experiments, respectively. Another one was used for histological analysis. All surgical and experimental procedures were approved by the National Institutes for Quantum and Radiological Science and Technology (11-1038-11) and by the Animal Care and Use Committee of the National Institute of Mental Health (Annual Report ZIMH002619), and were in accordance with the Institute of Laboratory Animal Research *Guide for the Care and Use of Laboratory Animals*.

Behavioral tasks

The monkeys squatted on a primate chair inside a dark, sound-attenuated, and electrically shielded room. A touch-sensitive bar was mounted on the chair. Visual stimuli were

displayed on a computer video monitor in front of the animal. Behavioral control and data acquisition were performed using a real-time experimentation system (REX) [50]. Neurobehavioral Systems Presentation software was used to display visual stimuli (Neurobehavioral Systems).

Ten monkeys were trained and tested with the delayed reward task (Fig. 1AB) [1]. In each of the trials, the monkey worked for one of six combinations of reward size and delay. Every trial had the same requirement for obtaining the reward: releasing the bar when a colored spot changed from red to green. Trials began when the monkey touched the bar at the front of the chair. A visual cue and a red spot (wait signal) sequentially appeared in the center of the monitor with a 0.1 s interval. After a variable interval, the red target turned green (go signal). If the monkey released the bar between 0.2 and 1 s after this go signal, the trial was considered correct and the spot turned blue (correct signal). A liquid, either small (1 drop, ca. 0.1 mL) or large reward (3 drops, except for monkey BI, 4 drops) was delivered immediately (0.3 ± 0.1 s) or with an additional delay of either 3.3 ± 0.6 s or 6.9 ± 1.2 s after correct release of the bar. Each combination of reward size and delay was chosen with equal probability and independently of the preceding reward condition. An inter-trial interval (ITI) of 1 s was enforced before allowing the next trial to begin. We used a fixed ITI instead of adjusted ITIs with post-reward delays (for example [51]), because monkeys are insensitive to post-reward delays in our tasks (please see Fig. 3 in [1]). Anticipatory bar releases (before or no later than 0.2 s after the appearance of the go signal) and failures to release the bar within 1 s after the appearance of the go signal were counted as errors. In error trials, the trial was terminated immediately, all visual stimuli disappeared and, following ITI, the trial was repeated, that is, the reward size/delay

combination remained the same as in the error trial.

In the behavioral experiment, the visual cue indicated a unique combination of reward size and delay. Two sets of cues were used: a stripe set (for 9 monkeys except for BI) and an image set (for monkey BI) (Fig. 1B). Prior to the behavioral experiment, all monkeys had been trained to perform color discrimination trials in a cued multi-trial reward schedule task for more than 3 months. We collected behavioral data with the delayed reward task for 5-25 daily testing sessions. Each session ended when the monkey would no longer initiate a new trial.

Two monkeys (RI and ST) was also tested with the reward-size task, in which the reward was always delivered immediately (0.3 ± 0.1 s), but the size of the reward (1, 2, 4, and 8 drops) varied and was assigned by unique cue (Fig. 9A) [1]. The sequence and timing of events were the same as those in the delayed reward task.

Surgery

After behavioral training, magnetic resonance (MR) images at 1.5T (monkey FG) and 7T (monkeys BI, RI and ST) were obtained under anesthesia (intravenous infusion of propofol 0.2-0.6 mg/kg/min, or pentobarbital sodium 15-30 mg/kg) to determine the position of the recording or local injection. After obtaining each MR image, a surgical procedure was carried out under general isoflurane anesthesia (1~2%) to implant chambers for unit recording and/or chemical inactivation. For monkeys BI and FG, we implanted a rectangle chamber (22 x 22 mm ID; KDS Ltd.) from vertical in the coronal plane aiming for the bilateral caudate nucleus. We implanted one or two cylinder

chambers (19 mm ID; Crist Instrument Co., Inc.) angled 10° or 20° from vertical in the coronal plane targeting the right or bilateral caudate for monkeys ST and RI, respectively. Based on measurements made from the MR images, the centers of the chambers were placed to target the caudate nucleus near the anterior commissure. A post for head fixation during data collection was also implanted.

Recording neuronal activity and mapping recording location

Single-unit activity was recorded (51, 31, and 68 from monkeys BI, FG and ST, respectively) while monkeys performed the delayed reward task in a block usually consisting of 120 trials. Action potentials of single neurons were recorded from the left CD using epoxy-coated 1.1-1.5 MΩ tungsten microelectrodes (Microprobes for Life Science; 1.1-1.5 MΩ at 1 kHz) or glass-coated 1.0 MΩ tungsten microelectrodes (Alpha Omega Engineering Ltd). A guide tube was inserted through the grid hole in the implanted recording chamber into the brain, and the electrodes were advanced through the guide tube by means of a micromanipulator (Narishige MO-97A or Alpha Omega EPS). Spike sorting to isolate single neuron discharges was performed with a time-window algorithm (TDT-RZ2, Tucker Davis Technologies) or custom-made software written in LabVIEW (National Instruments). Striatal neuronal activities were classified into two subtypes: presumed projection neurons and tonically active neurons (TANs, presumed cholinergic interneurons) based on their background discharge rates and action potential waveforms, as previously described [52]. We exclusively examined the activity of the presumed projection neurons, which are characterized as having a low spontaneous discharge rate (< 2 spikes/s) and exhibiting phasic discharges in relation to one or more behavioral task events. The activity of TANs recorded from the CD of monkeys performing a similar task

was reported in a previous study [53]. The timing of action potentials was recorded together with all task events at millisecond precision. Eye movements were also monitored for corneal reflection of an infrared light beam through a video camera at a sampling rate of 120Hz (i_rec, http://staff.aist.go.jp/k.matsuda/eye/).

To confirm the recording location, MR or CT (3D Accuitomo 170, J.MORITA CO.) images were acquired with a tungsten microelectrode (Fig. 1D). Recording sites extended from 2 mm anterior to the anterior commissure (AC) to 3 mm posterior to the AC for monkey BI, from 4 mm anterior to the AC to 3 mm posterior to the AC for monkey FG, and from 3 mm anterior to the AC to 2 mm posterior to AC for monkey ST.

Localization of DLPFC recipient striatum

A monkey (#214) was injected with AAV2-CMV-mKO (1.7 x $10e^{13}$ particles/mL) into the left DLPFC (S2A and S2B Figs). Details of the surgical procedures for the viral vector injection were described elsewhere [23]. The vector was injected at 1 to 2 µl per site. 33 days after injection, the monkey was deeply anesthetized with an overdose of sodium pentobarbital (80 mg/kg, i.v.) and transcardially perfused with saline at 4°C, followed by 4% paraformaldehyde in 0.1 M phosphate buffered saline, pH 7.4 (PBS). The brain was removed from the skull, postfixed in the same fresh fixative overnight, saturated with 30% sucrose in PB at 4°C, and then cut serially into 50-µm-thick sections on a freezing microtome. For visualization of immunoreactive signals of mKO, a series of every 6th section was immersed in 1% skim milk for 1 h at room temperature and incubated overnight at 4°C with rabbit anti-monomeric Kusabira-Orange (1:500; PM051M, MEDICAL & BIOLOGICAL LABORATORIES CO., LTD) in PBS containing 0.1%

Triton X-100 and 1% normal donkey serum for 2 days at 4°C. The sections were then incubated in the same fresh medium containing biotinylated donkey anti-rabbit IgG antibody (1:1,000; Jackson ImmunoResearch) for 2 h at room temperature, followed by avidin-biotin-peroxidase complex (ABC Elite, Vector Laboratories) for 1.5 h at room temperature. For visualization of the antigen, the sections were reacted in 0.05 M Tris-HCl buffer (pH 7.6) containing 0.04% diaminobenzidine (DAB), 0.04% $NiCl_2$, and 0.003% H_2O_2. The sections were mounted on gelatin-coated glass slides, air-dried, and cover-slipped. A part of other sections was Nissl-stained with 1% Cresyl violet.

Chemogenetic inactivation

One monkey (ST) received bilateral injections of an adeno-associated virus vector (AAV1-hSyn-hM$_4$Di-IRES-AcGFP; 3 μL/site; 4.7 x 10e^{13} particles/mL) at two locations into each side of the CD. The injection procedure was as described previously [25]. 49 days post vector injection, the monkey underwent a PET scan with [^{11}C]deschloroclozapine (DCZ) to visualize *in vivo* hM$_4$Di expression. Chemogenetic silencing was achieved by intramuscular injection (i.m.) with a DREADD selective agonist, DCZ (HY-42110, MedChemExpress; 0.1 mg/kg). DCZ was dissolved in 2% dimethyl sulfoxide (DMSO) in saline to a final volume of 0.65 ml. DCZ solution or vehicle (as control) was administered intramuscularly. Five to ten min following administration, the animal was allowed to start performing the tasks, which continued for 100 min. Based on a previous study, chemogenetic silencing would be effective 15-120 min after DCZ administration [23]. We performed at most one inactivation study per week. Note that we verified that the DCZ administration (0.1 mg/kg, i.m.) does not cause any significant motivational/motor impairments or alteration of the incentive effect of the

performance of reward-size task in monkeys without expressing DREADDs (n = 3) [23]. Detailed protocols for PET imaging were described elsewhere [23].

Pharmacological inactivation

To inactivate neuronal activity, we injected a $GABA_A$ receptor agonist, muscimol (M1523, Sigma-Aldrich), locally into the bilateral CD of monkeys BI and RI. We used two stainless steel injection cannulae inserted into the caudate (O.D. 350 μm; BRC Inc., Japan), one in each hemisphere. Each cannula was connected to a 5-μl microsyringe (Hamilton, #7105KH) via polyethylene tubing. These cannulae were advanced through the guide tube by means of an oil-drive micromanipulator. Muscimol (4 μg/1 μL saline) was injected at a rate of 0.2 μL/min by auto-injector (Legato210, KD Scientific Inc.) for a total volume of 3 μL in each side. Soon after the injection was completed, the animal was allowed to start performing the tasks, which continued for 100 min. We performed at most one inactivation study per week. For control, we injected saline at other times using the same parameters as those used for muscimol. At the end of each session, a CT scan was conducted to visualize the injection cannulae in relation to the chambers and skull. The CT images were overlaid on MR images by using PMOD® and VirtualPlace (AZE Inc.) image analysis software to assist in identifying the injection sites (Fig. 1D and 7A).

Data analysis

The R statistical computing environment (R Development Core Team 2004) was used for all data analyses.

Behavioral data analysis. Error rates in task performance were calculated by dividing

the total number of errors by the total number of trials for each reward condition and then averaged across all sessions. The average error rates in the delayed reward task were fitted with the inverse function of reward size with hyperbolic (Eq. 2) or that with exponential temporal discounting [1] as follows:

$$E = \frac{e^{-kD}}{aR} \#(3).$$

We fitted these 2 models to the data with least-squares minimization using 'optim' function in R, and compared the models by leave-one-out cross-validation as described previously [1].

To examine the effects of satiation, we divided each session into quartiles based on normalized cumulative reward, R_{cum}, which was 0.125, 0.375, 0.625, and 0.875 for the first through fourth quartiles, respectively. We fitted the error rates in delayed reward task obtained from each monkey and the average data across monkeys to the following model:

$$E = \frac{1 + kD}{aR \times F(R_{cum})} \#(4),$$

where the satiation effect, $F(R_{cum})$, as the reward value was exponentially decaying in R_{cum} at a constant λ [54].

$$F(R_{cum}) = e^{-\lambda R_{cum}} \#(5).$$

For modeling satiation effects of the error rates in reward-size task, we used an inverse model integrating satiation effect (Eq. 5), as follows:

$$E = \frac{1}{aR \times F(R_{cum})} \#(6).$$

We also applied conventional ANOVA modeling to the behavioral data. The proportional behavioral data were transformed using the variance stabilizing arcsine transformation before hypothesis testing [55].

Significance of deviation from baseline data was examined by means of the parametric bootstrapping method (n = 20,000). We first constructed distribution of the sum of squared residuals (SSR) of the best fit of the model to the averaged resampled error rates (n = 5 or 4 sessions for delayed reward and reward-size task, respectively) from the pooled sample in baseline conditions in each subject. For this analysis, we used normalized error rates by the maximum error rates among reward conditions in each session to remove variance across sessions. P-values for deviation from the distribution were obtained for SSR of the best fit of the model-to-test data (control or inactivation).

Neuronal data analysis. Only neuronal data from correct trials were used for the analyses. For each neuron, we collected data from 20 - 30 correct trials for each combination of reward-size-and-delay duration, a total of 120 - 180 successful trials. For each neuron, we first determined the start and end of event-related responses by using a series of χ^2 tests [56]. The background window was defined as the activity between 500 and 0 ms before cue onset. The test window spanned 100 ms for cue responses, and it moved in 10-ms increments, from 0 to 1,500 ms, after cue appearance. For bar-release responses, the 100-ms test window moved from 300 ms before to 300 ms after bar release. For reward-related responses, the 100-ms test window moved from 0 to 500 ms after reward appearance. For each 100-ms test window, an χ^2 test was used to determine whether the proportions of filled to empty 1-ms bins in the 100-ms test interval were significantly

different from the proportion in the 500-ms background window. Start of the response was taken to be the middle of the first of four consecutive 100-ms test intervals showing a significant difference ($p < 0.05$) in spike count between the test and background window. End of the response was defined as the middle of the last window showing a significant difference. Duration of the response was defined as the difference between the start and end of the response. The procedure worked well for all tested neurons, yielding latencies that matched those we would have chosen by visual inspection. A neuron was classified as responsive to the three events when a significant response could be detected in at least five consecutive windows.

To quantify the influence of temporal discounting of reward value on the response, we applied linear regression analysis. For each significant response, firing rates (Y) were fitted by the following linear regression model:

$$Y = \beta_0 + \beta_V V \#(7),$$

where β_V is the regression coefficient and β_0 is the intercept, and V is the temporally discounted value formulated by a hyperbolic function (Eq. 1) [3-5]. The effect of V was compared with that of delay and reward size information on the response by the following multiple linear regression model:

$$Y = \beta_0 + \beta_{delay}\, D + \beta_{size}\, R + \beta_V\, V \#(8),$$

where D and R are delay duration and reward size, respectively, β_{delay}, β_{size}, and β_V are the regression coefficients, and β_0 is the intercept. Another linear regression analysis was performed to quantify the influence of temporal discounting of reward value and satiation on the response, as follows:

$$Y = \beta_0 + \beta_V\, V + \beta_{Rcum}\, R_{cum} \#(9).$$

where V and R_{cum} are the temporally discounted value and cumulative reward, respectively, β_{V} and β_{Rcum} are the regression coefficients, and β_0 is the intercept.

Acknowledgements

We thank George Dold, Keiji Matsuda, Risa Suma, Tomoni Kokufuta, Anzu Maruyama, Jun Kamei, Yuichi Matsuda, Ryuji Yamaguchi, Yoshio Sugii, Maki Fujiwara, and Mayuko Nakano for their technical assistance. One Japanese monkey used in this research was provided by NBRP "Japanese Monkeys" through the National BioResource Project of the MEXT.

References

1. Minamimoto T, La Camera G, Richmond BJ. Measuring and modeling the interaction among reward size, delay to reward, and satiation level on motivation in monkeys. J Neurophysiol. 2009;101(1):437-47.
2. Shadmehr R, de Xivry JJO, Xu-Wilson M, Shih TY. Temporal Discounting of Reward and the Cost of Time in Motor Control. J Neurosci. 2010;30(31):10507-16.
3. Mazur JE. Tests of an equivalence rule for fixed and variable reinforcer delays. J. Exp. Psychol.: Anim. Behav. Processes. 1984;10(4):426-36.
4. Mazur JE. Hyperbolic value addition and general models of animal choice. Psychol Rev. 2001;108(1):96-112.
5. Green L, Myerson J. A discounting framework for choice with delayed and probabilistic rewards. Psychol Bull. 2004;130(5):769-92.
6. Berridge KC. Motivation concepts in behavioral neuroscience. Physiol Behav. 2004;81(2):179-209.
7. Toates F. Motivational Systems. Cambridge: Cambridge University Press. 1986.
8. Zhang J, Berridge KC, Tindell AJ, Smith KS, Aldridge JW. A Neural Computational Model of Incentive Salience. PLoS Comput Biol. 2009;5(7). doi: 10.1371/journal.pcbi.1000437.
9. Peper JS, Mandl RCW, Braams BR, de Water E, Heijboer AC, Koolschijn P, et al. Delay Discounting and Frontostriatal Fiber Tracts: A Combined DTI and MTR Study on Impulsive Choices in Healthy Young Adults. Cereb Cortex. 2013;23(7):1695-702.

10. van den Bos W, Rodriguez CA, Schweitzer JB, McClure SM. Connectivity Strength of Dissociable Striatal Tracts Predict Individual Differences in Temporal Discounting. J Neurosci. 2014;34(31):10298-310.
11. van den Bos W, Rodriguez CA, Schweitzer JB, McClure SM. Adolescent impatience decreases with increased frontostriatal connectivity. Proc Natl Acad Sci U S A. 2015;112(29):E3765-E74
12. Kim S, Hwang J, Lee D. Prefrontal coding of temporally discounted values during intertemporal choice. Neuron. 2008;59(1):161-72.
13. Kalenscher T, Windmann S, Diekamp B, Rose J, Gunturkun O, Colombo M. Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. Curr Biol. 2005;15(7):594-602.
14. Figner B, Knoch D, Johnson EJ, Krosch AR, Lisanby SH, Fehr E, et al. Lateral prefrontal cortex and self-control in intertemporal choice. Nat Neurosci. 2010;13(5):538-9.
15. Simmons JM, Minamimoto T, Murray EA, Richmond BJ. Selective ablations reveal that orbital and lateral prefrontal cortex play different roles in estimating predicted reward value. J Neurosci. 2010;30(47):15878-87.
16. Cai X, Kim S, Lee D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. Neuron. 2011;69(1):170-82.
17. Haber SN, Kunishio K, Mizobuchi M, Lyndbalta E. The Orbital and Medial Prefrontal Circuit through the Primate Basal Ganglia. J Neurosci. 1995;15(7):4851-67.

18. Haber SN, Kim KS, Mailly P, Calzavara R. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. J Neurosci. 2006;26(32):8368-76.
19. Cromwell HC, Schultz W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J Neurophysiol. 2003;89(5):2823-38.
20. Hassani OK, Cromwell HC, Schultz W. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. J Neurophysiol. 2001;85(6):2477-89.
21. Hori Y, Minamimoto T, Kimura M. Neuronal encoding of reward value and direction of actions in the primate putamen. J Neurophysiol. 2009;102(6):3530-43.
22. Kawagoe R, Takikawa Y, Hikosaka O. Expectation of reward modulates cognitive signals in the basal ganglia. Nat Neurosci. 1998;1(5):411-6.
23. Nagai Y, Miyakawa N, Takuwa H, Hori Y, Oyama K, Ji B, et al. Deschloroclozapine, a potent and selective chemogenetic actuator enables rapid neuronal and behavioral modulations in mice and monkeys. Nat Neurosci. in press. Priprint available from bioRxiv; https://doi.org/10.1101/854513.
24. Roth BL. DREADDs for Neuroscientists. Neuron. 2016;89(4):683-94.
25. Nagai Y, Kikuchi E, Lerchner W, Inoue K, Ji B, Eldridge MAG, et al. PET imaging-guided chemogenetic silencing reveals a critical role of primate rostromedial caudate in reward evaluation. Nature Commun. 2016;7. doi: 10.1038/ncomms13605.
26. Fujimoto A, Hori Y, Nagai Y, Kikuchi E, Oyama K, Suhara T, et al. Signaling Incentive and Drive in the Primate Ventral Pallidum for Motivational Control of Goal-Directed Action. J Neurosci. 2019;39(10):1793-804.

27. Minamimoto T, Yamada H, Hori Y, Suhara T. Hydration level is an internal variable for computing motivation to obtain water rewards in monkeys. Exp Brain Res. 2012;218(4):609-18.

28. Ito M, Doya K. Distinct neural representation in the dorsolateral, dorsomedial, and ventral parts of the striatum during fixed- and free-choice tasks. J Neurosci. 2015;35(8):3499-514.

29. Lau B, Glimcher PW. Value representations in the primate striatum during matching behavior. Neuron. 2008;58(3):451-63.

30. Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. Science. 2005;310(5752):1337-40.

31. Averbeck BB, Lehman J, Jacobson M, Haber SN. Estimates of projection overlap and zones of convergence within frontal-striatal circuits. J Neurosci. 2014;34(29):9497-505.

32. Calzavara R, Mailly P, Haber SN. Relationship between the corticostriatal terminals from areas 9 and 46, and those from area 8A, dorsal and rostral premotor cortex and area 24c: an anatomical substrate for cognition to action. Eur J Neurosci. 2007;26(7):2005-24.

33. Selemon LD, Goldman-Rakic PS. Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. J Neurosci. 1985;5(3):776-94.

34. Hosokawa T, Kennerley SW, Sloan J, Wallis JD. Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. J Neurosci. 2013;33(44):17385-97.

35. Kobayashi S, Schultz W. Influence of reward delays on responses of dopamine neurons. J Neurosci. 2008;28(31):7837-46.

36. Smith Y, Raju DV, Pare JF, Sidibe M. The thalamostriatal system: a highly specific network of the basal ganglia circuitry. Trends Neurosci. 2004;27(9):520-7.

37. Minamimoto T, Hori Y, Kimura M. Complementary process to response bias in the centromedian nucleus of the thalamus. Science. 2005;308(5729):1798-801.

38. Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. J Neurosci. 2009;29(42):13365-76.

39. Bouret S, Richmond BJ. Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. J Neurosci. 2010;30(25):8591-601.

40. Critchley HD, Rolls ET. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. J Neurophysiol. 1996;75(4):1673-86.

41. Rolls ET, Sienkiewicz ZJ, Yaxley S. Hunger modulates the responses to gustatory stimuli of single neurons in the caudoIateraI orbitofrontal cortex of the macaque monkey. Eur J Neurosci. 1989;1(1):53-60.

42. Izquierdo A, Murray EA. Functional interaction of medial mediodorsal thalamic nucleus but not nucleus accumbens with amygdala and orbital prefrontal cortex is essential for adaptive response selection after reinforcer devaluation. J Neurosci. 2010;30(2):661-9.

43. Murray EA, Rudebeck PH. The drive to strive: goal generation based on current needs. Front Neurosci. 2013;7:112. doi: 10.3389/fnins.2013.00112.

44. Yin HH, Ostlund SB, Balleine BW. Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. Eur J Neurosci. 2008;28(8):1437-48.

45. Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci. 2005;22(2):513-23.

46. Balleine BW, Delgado MR, Hikosaka O. The role of the dorsal striatum in reward and decision-making. J Neurosci. 2007;27(31):8161-5.

47. Haber SN, Knutson B. The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology. 2010;35(1):4-26.

48. Amlung M, Marsden E, Holshausen K, Morris V, Patel H, Vedelago L, et al. Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. JAMA Psychiatry. 2019. doi: 10.1001/jamapsychiatry.2019.2102.

49. Pulcu E, Trotter PD, Thomas EJ, McFarquhar M, Juhasz G, Sahakian BJ, et al. Temporal discounting in major depressive disorder. Psychol Med. 2014;44(9):1825-34.

50. Hays A, Richmond B, Optican L. Unix-based multiple process system for real-time data aquisition and control. WESCON 1982;2:1-10.

51. Blanchard TC, Pearson JM, Hayden BY. Postreward delays and systematic biases in measures of animal temporal discounting. Proc Natl Acad Sci U S A. 2013;110(38):15491-6

52. Yamada H, Inokawa H, Hori Y, Pan X, Matsuzaki R, Nakamura K, et al. Characteristics of fast-spiking neurons in the striatum of behaving monkeys. Neurosci Res. 2015. doi: 10.1016/j.neures.2015.10.003.

53. Falcone R, Weintraub DB, Setogawa T, Wittig JH, Jr., Chen G, Richmond BJ. Temporal Coding of Reward Value in Monkey Ventral Striatal Tonically Active Neurons. J Neurosci. 2019;39(38):7539-50.

54. Minamimoto T, Hori Y, Richmond BJ. Is working more costly than waiting in monkeys? PLoS One. 2012;7(11):e48434. doi: 10.1371/journal.pone.0048434.

55. Zar JH. Biostatistical Analysis: Prentice Hall. 2010.

56. Ravel S, Richmond BJ. Dopamine neuronal responses in monkeys performing visually cued reward schedules. Eur J Neurosci. 2006;24(1):277-90.

Supporting information

S1 Fig. Reaction time and eye position. (A) Reaction time (mean ± SD) of delayed reward task as a function of delay duration in monkeys BI, FG, and ST. Black and white symbols indicate small (1 drop) and large reward (3 or 4 drops), respectively. (B) Density plots of eye position during cue period of delayed reward task obtained from monkey RI. Colors indicate normalized looking-time. White squares indicate the frame of cue stimulus.

S2 Fig. Identification of subregion of caudate nucleus receiving DLPFC projection. (A) Top: Location of viral vector (AAV2-CMV-mKO) injection into dorsal and ventral banks of principal sulcus (PS). Photograph of fluorescence stereomicroscope presents position of 9 injection sites along with PS. Bottom: Photo of typical coronal Nissl and DAB staining sections (A1), which visualizes mKO fluorescence positive neurons. Scale bar, 50 μm. (B) Distribution patterns of corticostriatal terminals from DLPFC. Microphotograph and charts represent DAB-staining mKO-positive sites. Focal projection areas were localized from 2 mm anterior to anterior commissure (AC) to 2 mm posterior to AC. Scale bar, 2 mm. AS, arcuate sulcus; CD, caudate nucleus; Put, putamen.

S3 Fig. Impact of DV and comparison with delay and size on cue on release and reward response. (A) Scatterplot of standardized partial regression coefficients (SPRC) of DV (ordinate) against those of size and delay on release response, respectively (abscissa). (B) Same as A, but for reward response. Colored dots indicate neurons with significant ($p < 0.05$) coefficient, while gray dots correspond to neurons without any significant effect (NA). DV/DV & Other, neurons with significant coefficient of DV; Size

& Delay, those with both size and delay; Size, those with size exclusively; Delay, those with delay exclusively. Numbers in parentheses indicate number of neurons.

S4 Fig. Impact of discounted value and satiation on cue response. (A) Scatterplot of standardized regression coefficients (SRC) of discharge rates during cue period for DV (ordinate) against those for cumulative reward (abscissa). Red and blue dots indicate neurons with significant ($p < 0.05$) coefficient for DV and cumulative reward (CR), respectively. Black circles correspond to neurons without any significant effect (NS). (B) Representative waveforms (mean ± SD) recorded from a CD neuron (Monkey ST #10) during first (purple) and last quartile (orange) of recording period. Changes in firing rate were not attributable to alteration in action potential isolation.

S5 Fig. No significant effects of CD inactivation on behavioral parameters in delayed reward task performance. (A) Comparison of total reward amount in a session between control and inactivation session in monkey BI. Bars and dots represent average and individual sessions. (B) Comparison of reaction time between control and inactivation session in monkey BI. Density plots and histogram represent pooled reaction times for each reward condition for control (blue) and inactivation session (red). Dots and lines indicate mean ± SD. (C) Density plots of eye position during cue period of delayed reward task obtained from monkey RI. Colors indicate normalized looking-time. Left and right panels for control and inactivation sessions, respectively. White squares indicate frame of cue stimulus.

S6 Fig. Normalized error rates in baseline, control and inactivation session of delayed reward task. Symbols represent normalized error rates for each reward condition by maximum error rates in each session. Thick lines connect average error rates for 3 delay conditions in each reward size. Vertical lines indicate sem.

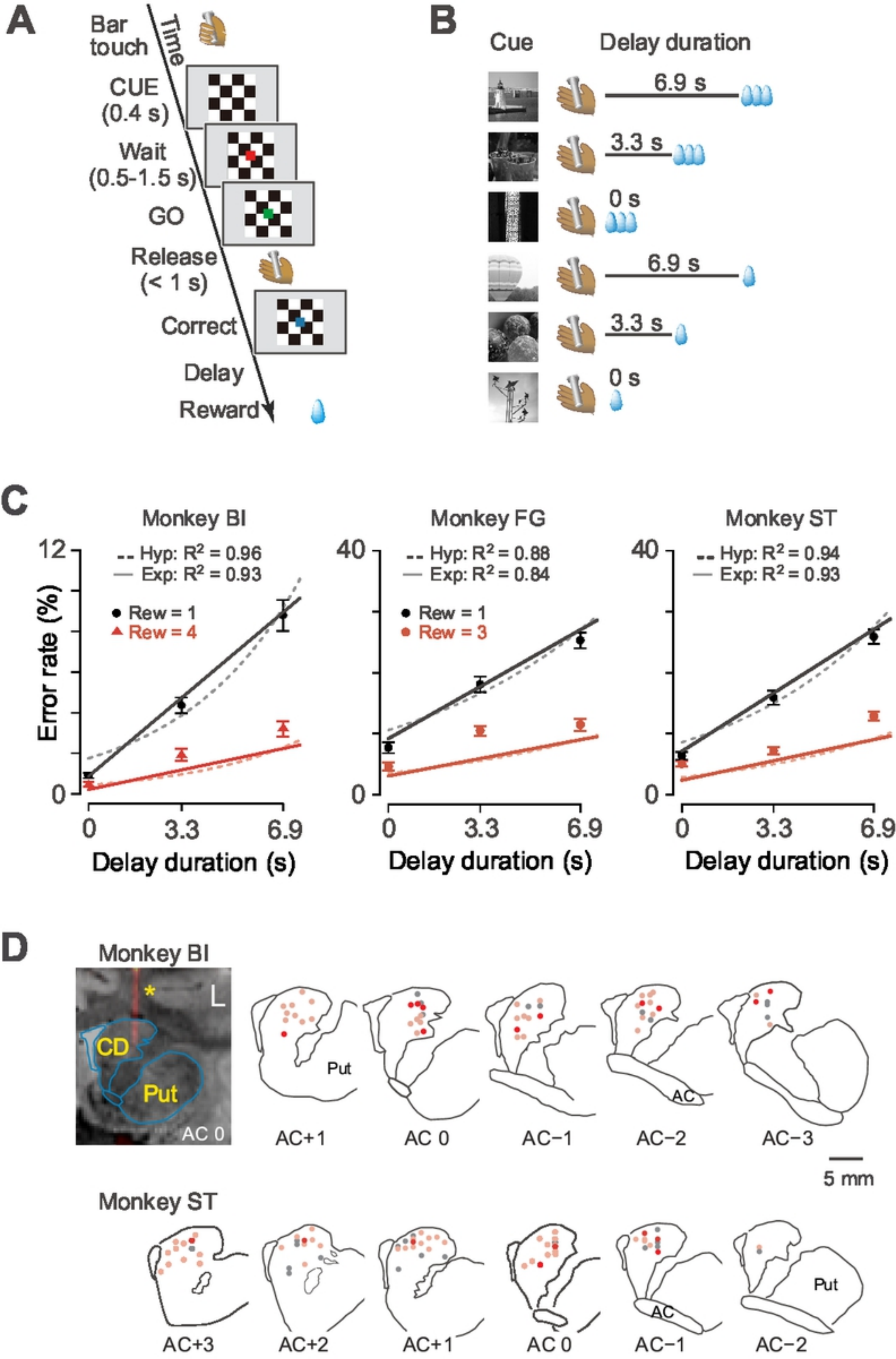
A
Bar touch
Time
CUE (0.4 s)
Wait (0.5-1.5 s)
GO
Release (< 1 s)
Correct
Delay
Reward
B
Cue
Delay duration
6.9 s
3.3 s
0 s
6.9 s
3.3 s
0 s
C
Monkey BI
Monkey FG
Monkey ST
Hyp: R² = 0.96
Exp: R² = 0.93
Hyp: R² = 0.88
Exp: R² = 0.84
Hyp: R² = 0.94
Exp: R² = 0.93
Rew = 1
Rew = 4
Rew = 3
Error rate (%)
Delay duration (s)
D
Monkey BI
CD
Put
AC 0
L
AC+1
AC 0
AC−1
AC−2
AC−3
5 mm
Monkey ST
AC+3
AC+2
AC+1
AC 0
AC−1
AC−2

Figure 1

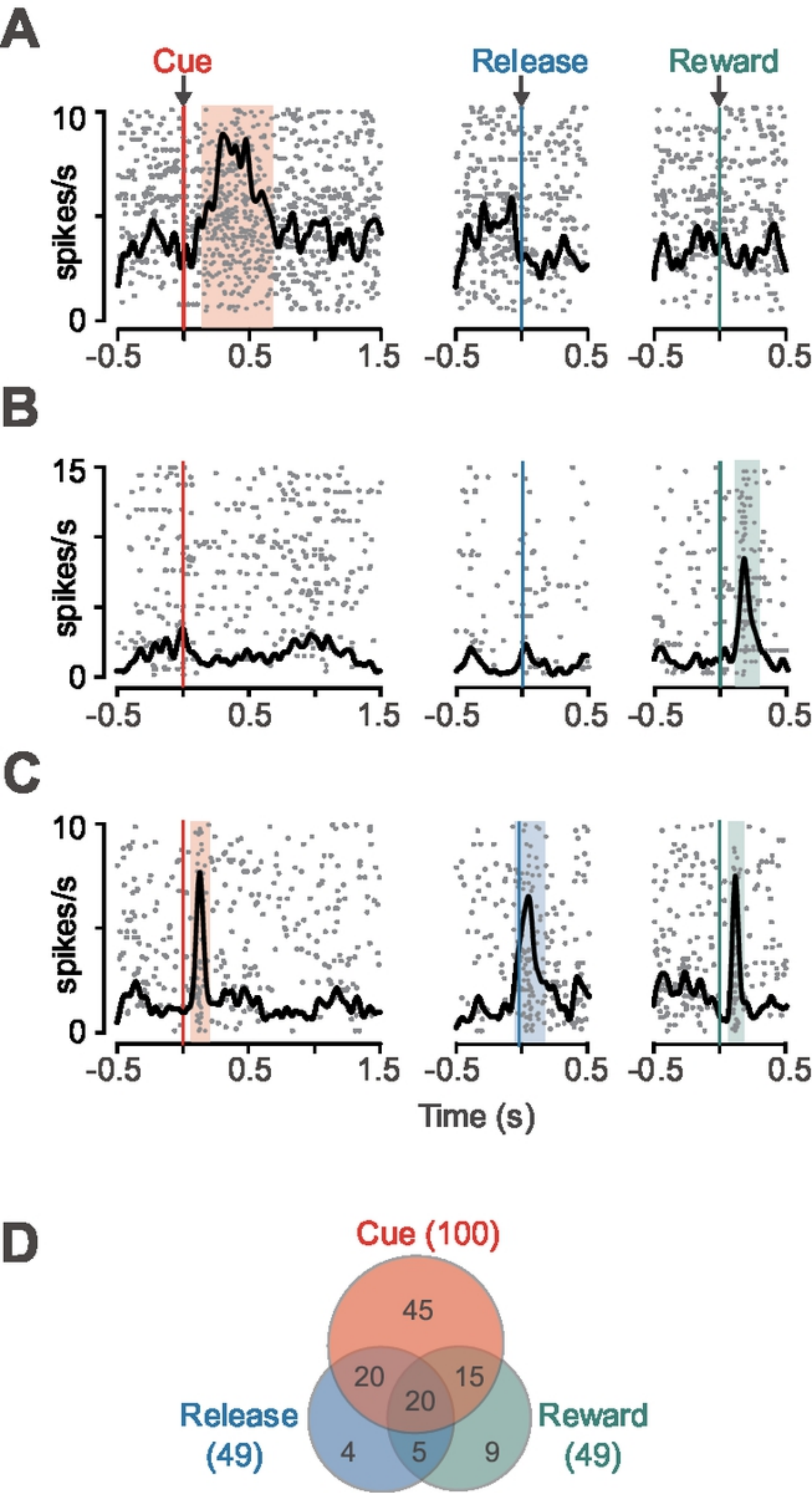
A
Cue
Release
Reward
10
spikes/s
0
-0.5
0.5
1.5
-0.5
0.5
-0.5
0.5
B
15
spikes/s
0
-0.5
0.5
1.5
-0.5
0.5
-0.5
0.5
C
10
spikes/s
0
-0.5
0.5
1.5
-0.5
0.5
-0.5
0.5
Time (s)
D
Cue (100)
45
20
15
20
Release
(49)
Reward
(49)
4
5
9

Figure 2

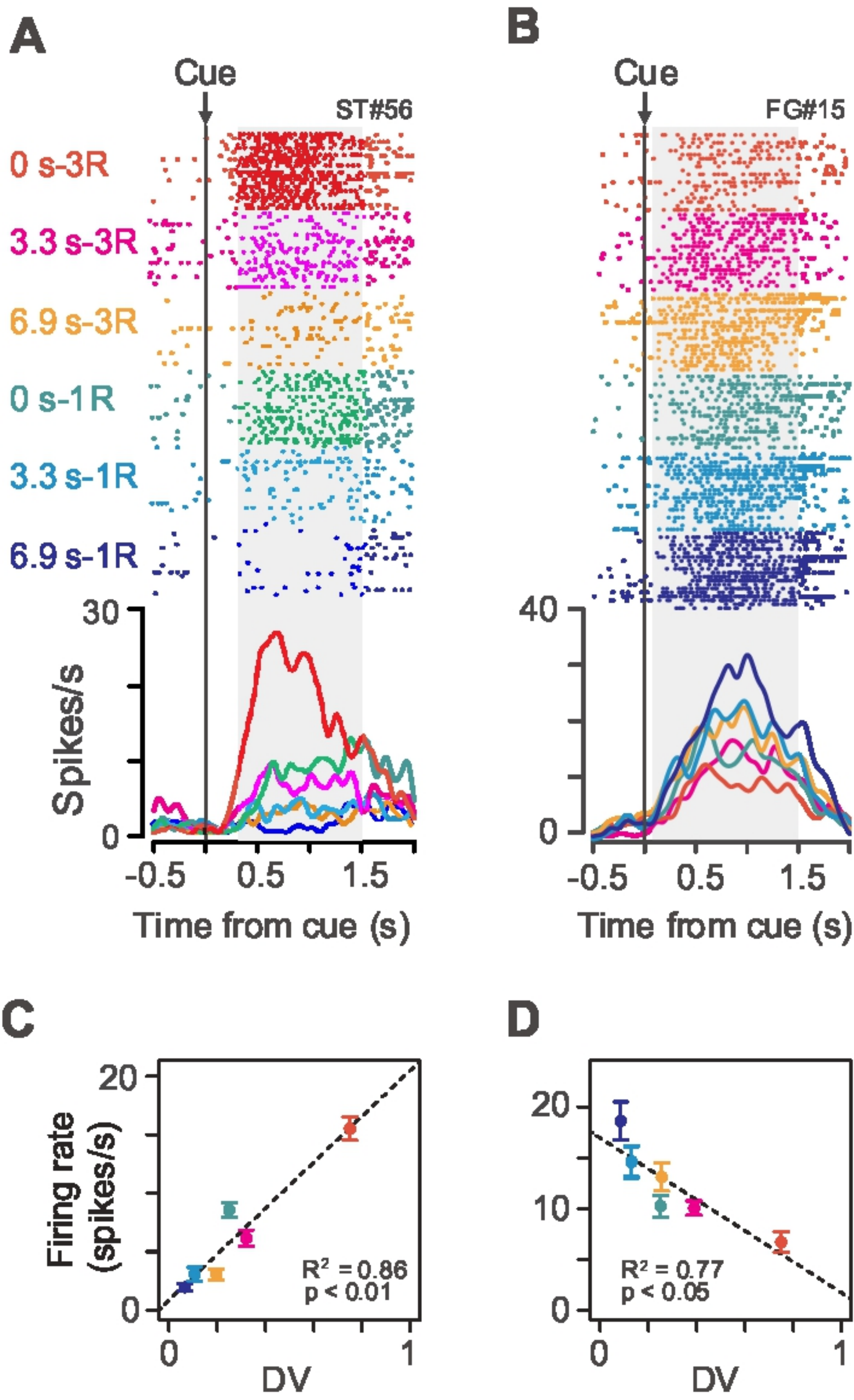
A
Cue
ST#56
0 s-3R
3.3 s-3R
6.9 s-3R
0 s-1R
3.3 s-1R
6.9 s-1R
30
Spikes/s
0
-0.5
0.5
1.5
Time from cue (s)
B
Cue
FG#15
40
0
-0.5
0.5
1.5
Time from cue (s)
C
20
Firing rate (spikes/s)
0
R² = 0.86
p < 0.01
0
1
DV
D
20
10
0
R² = 0.77
p < 0.05
0
1
DV

Figure 3

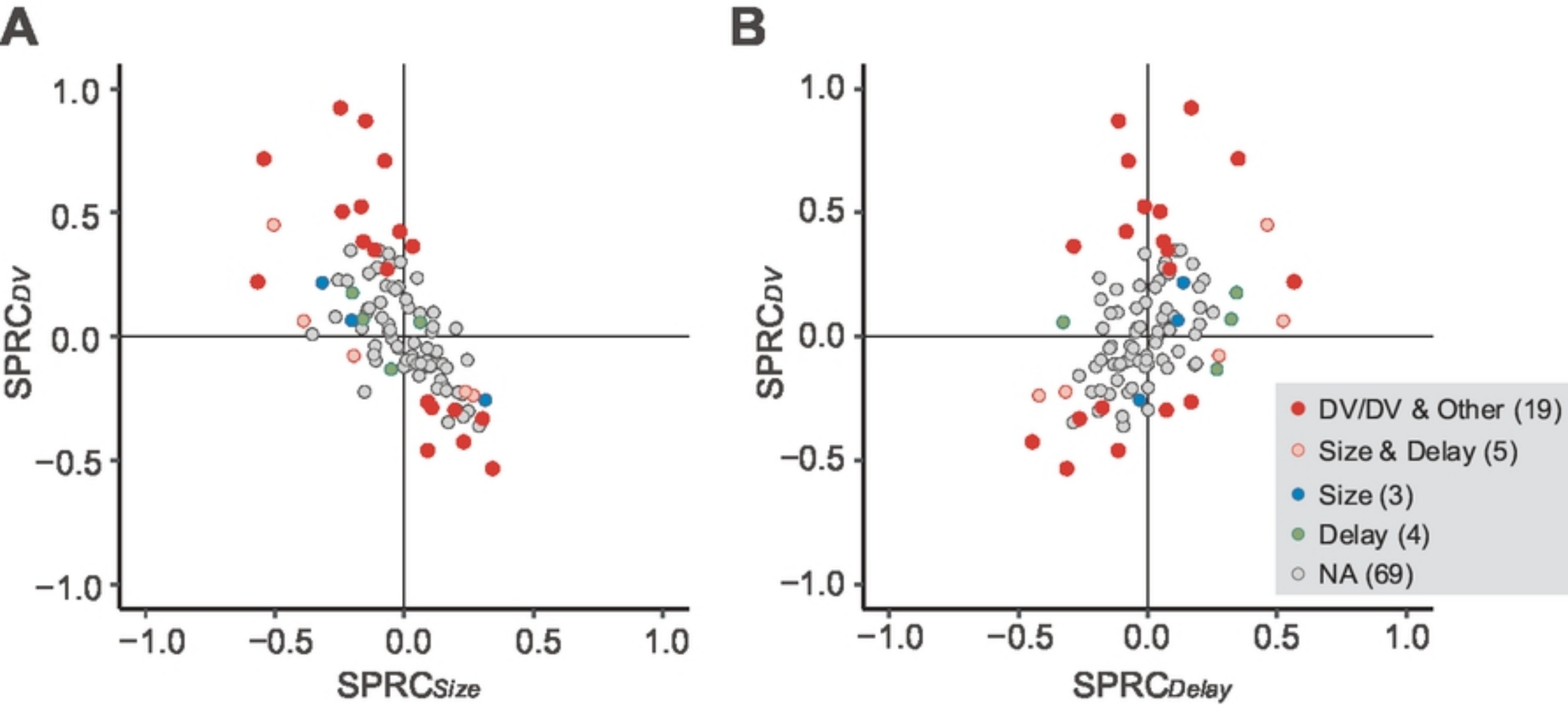
A
B
1.0
0.5
0.0
−0.5
−1.0
SPRCDV
SPRCSize
SPRCDelay
DV/DV & Other (19)
Size & Delay (5)
Size (3)
Delay (4)
NA (69)

Figure 4

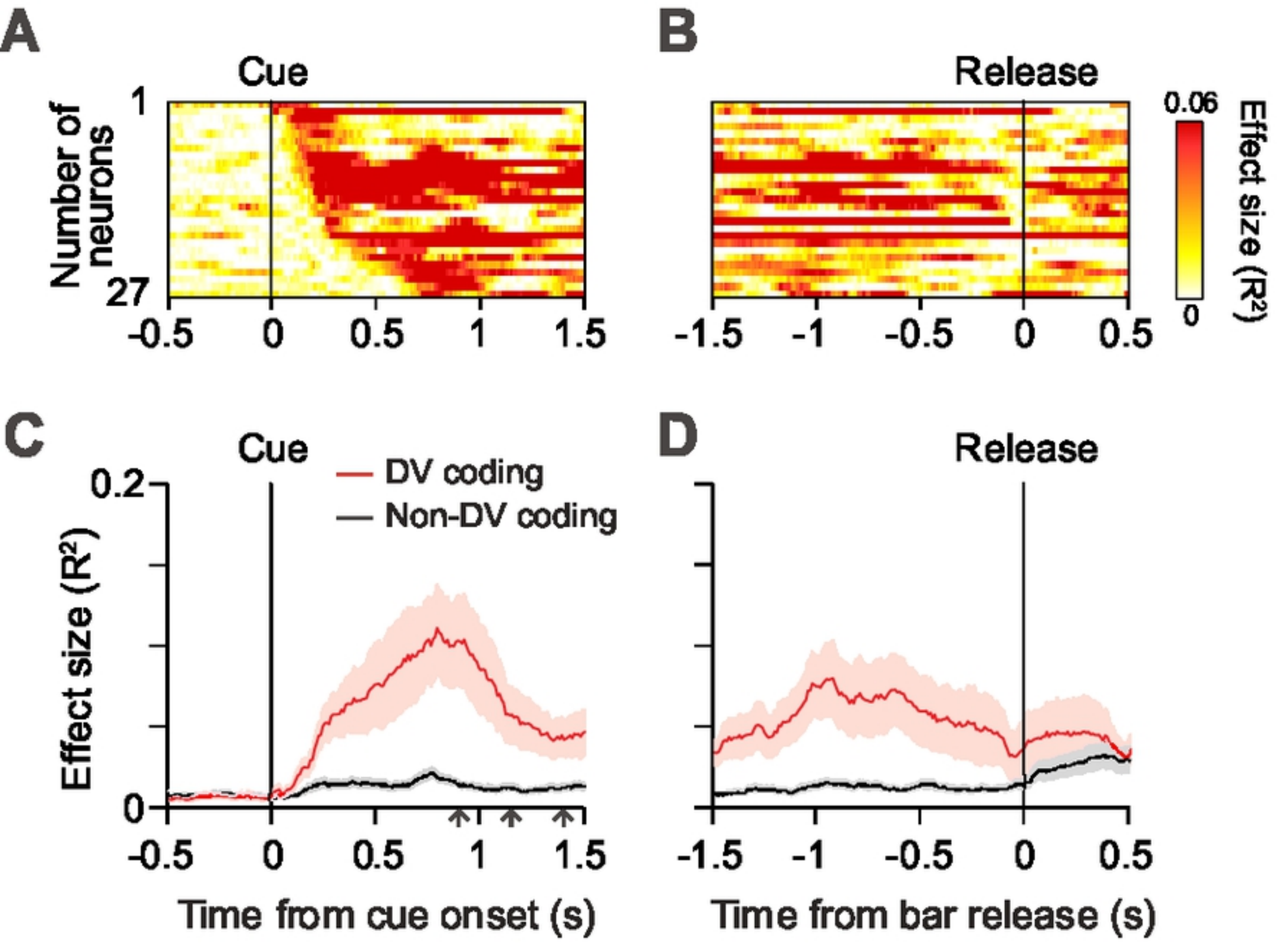
A
Cue
Number of neurons
1
27
-0.5
0
0.5
1
1.5
B
Release
0.06
0
Effect size (R²)
-1.5
-1
-0.5
0
0.5
C
Cue
DV coding
Non-DV coding
0.2
0
Effect size (R²)
-0.5
0
0.5
1
1.5
Time from cue onset (s)
D
Release
-1.5
-1
-0.5
0
0.5
Time from bar release (s)

Figure 5

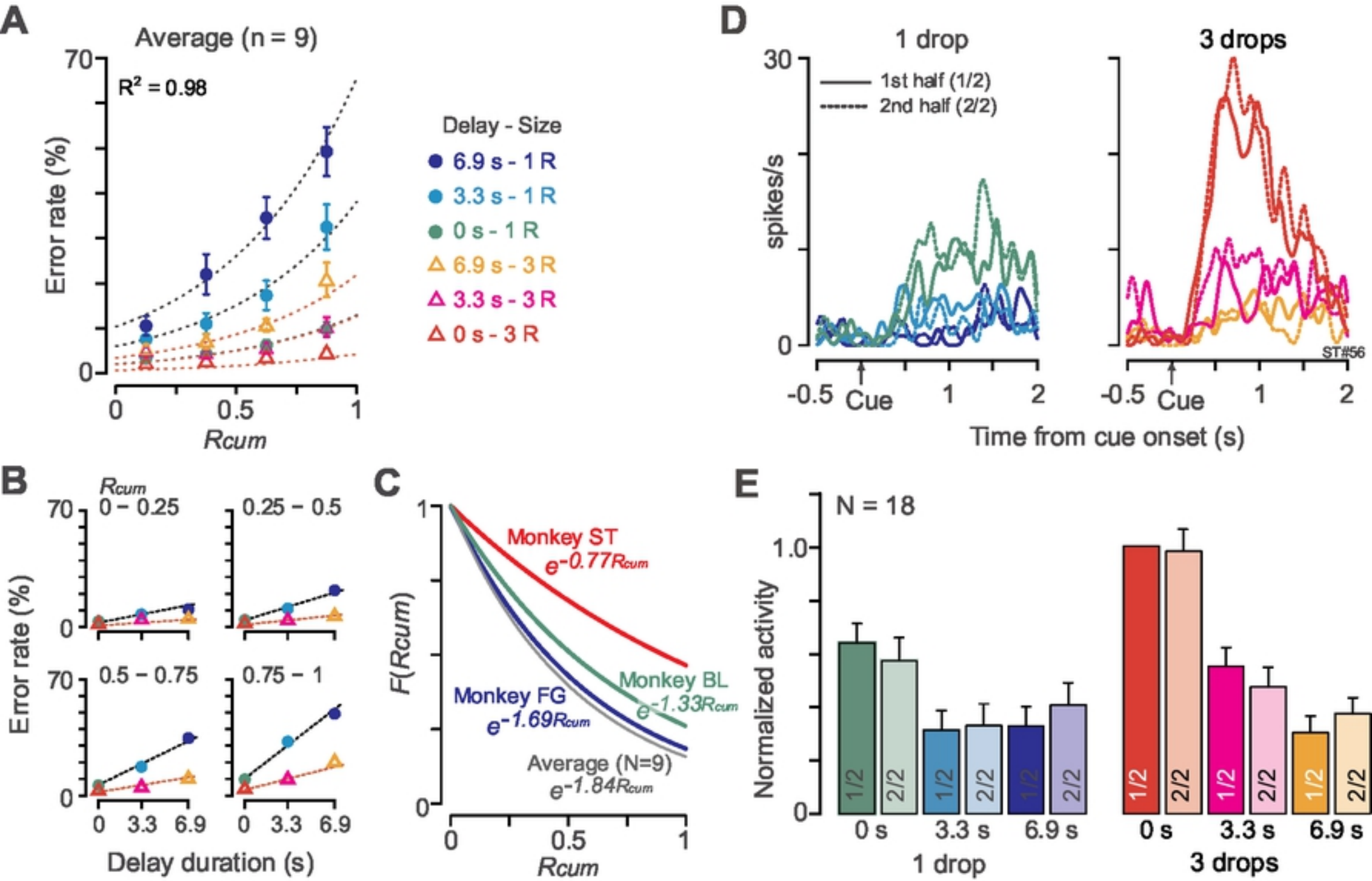
A
Average (n = 9)
R² = 0.98
Error rate (%)
Rcum
Delay - Size
6.9 s - 1 R
3.3 s - 1 R
0 s - 1 R
6.9 s - 3 R
3.3 s - 3 R
0 s - 3 R
B
Rcum
0 – 0.25
0.25 – 0.5
0.5 – 0.75
0.75 – 1
Error rate (%)
Delay duration (s)
C
Monkey ST
e-0.77Rcum
Monkey BL
e-1.33Rcum
Monkey FG
e-1.69Rcum
Average (N=9)
e-1.84Rcum
F(Rcum)
Rcum
D
1 drop
3 drops
1st half (1/2)
2nd half (2/2)
spikes/s
Cue
ST#56
Time from cue onset (s)
E
N = 18
Normalized activity
1/2
2/2
0 s
3.3 s
6.9 s
1 drop
3 drops

Figure 6

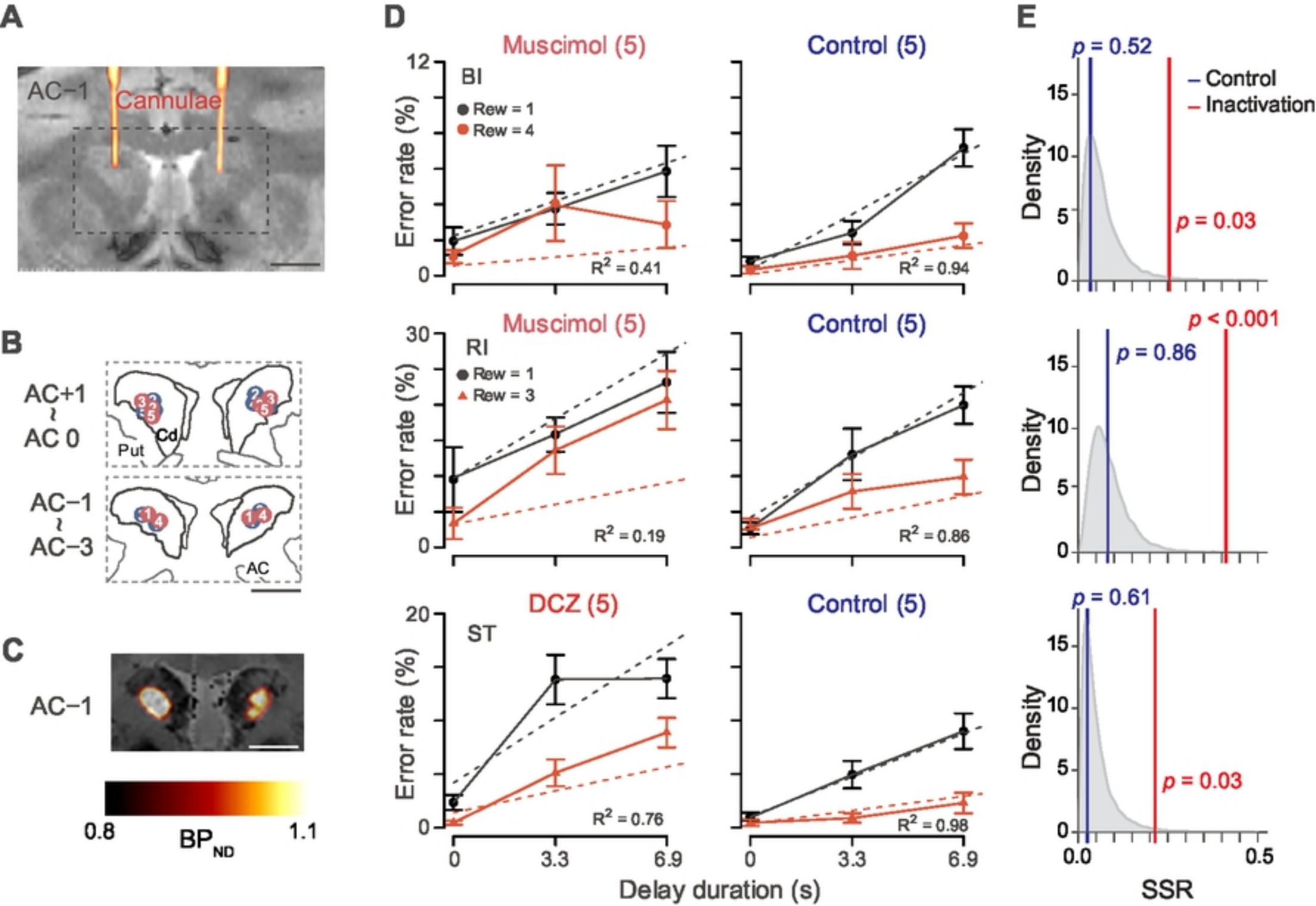
A
AC−1
Cannulae
B
AC+1
~
AC 0
Cd
Put
AC−1
~
AC−3
AC
C
AC−1
0.8
BP ND
1.1
D
Muscimol (5)
Control (5)
BI
Rew = 1
Rew = 4
Error rate (%)
R2 = 0.41
R2 = 0.94
RI
Rew = 3
R2 = 0.19
R2 = 0.86
DCZ (5)
ST
R2 = 0.76
R2 = 0.98
0
3.3
6.9
Delay duration (s)
E
p = 0.52
Control
Inactivation
Density
p = 0.03
p < 0.001
p = 0.86
p = 0.61
p = 0.03
0.0
0.5
SSR

Figure 7

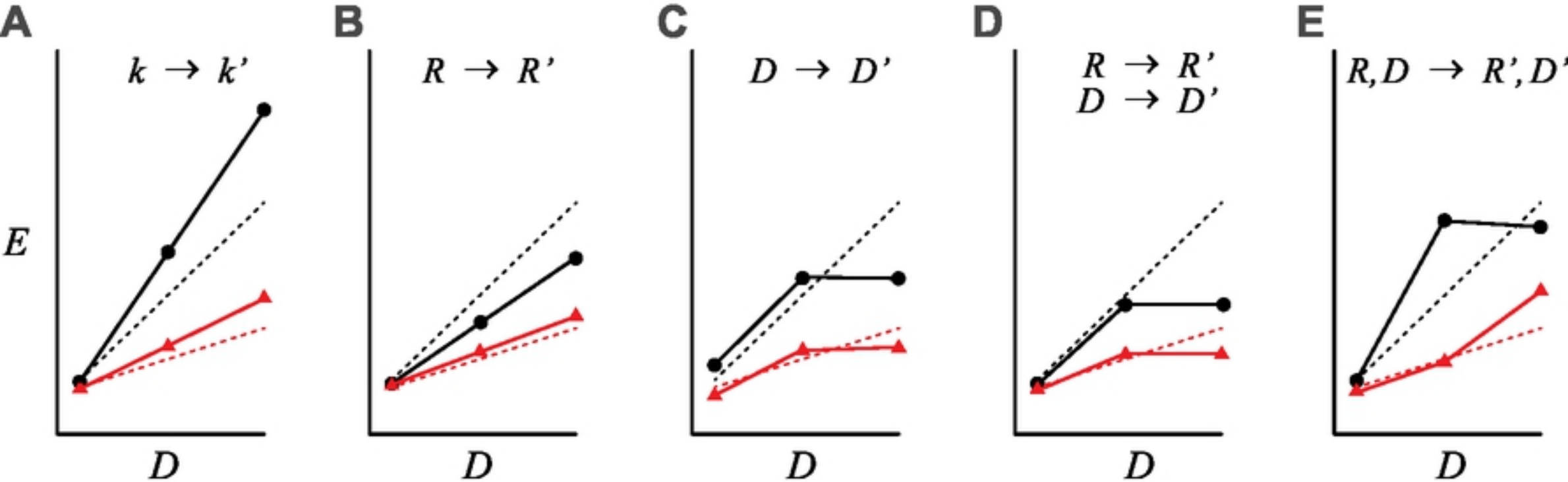

Figure 8

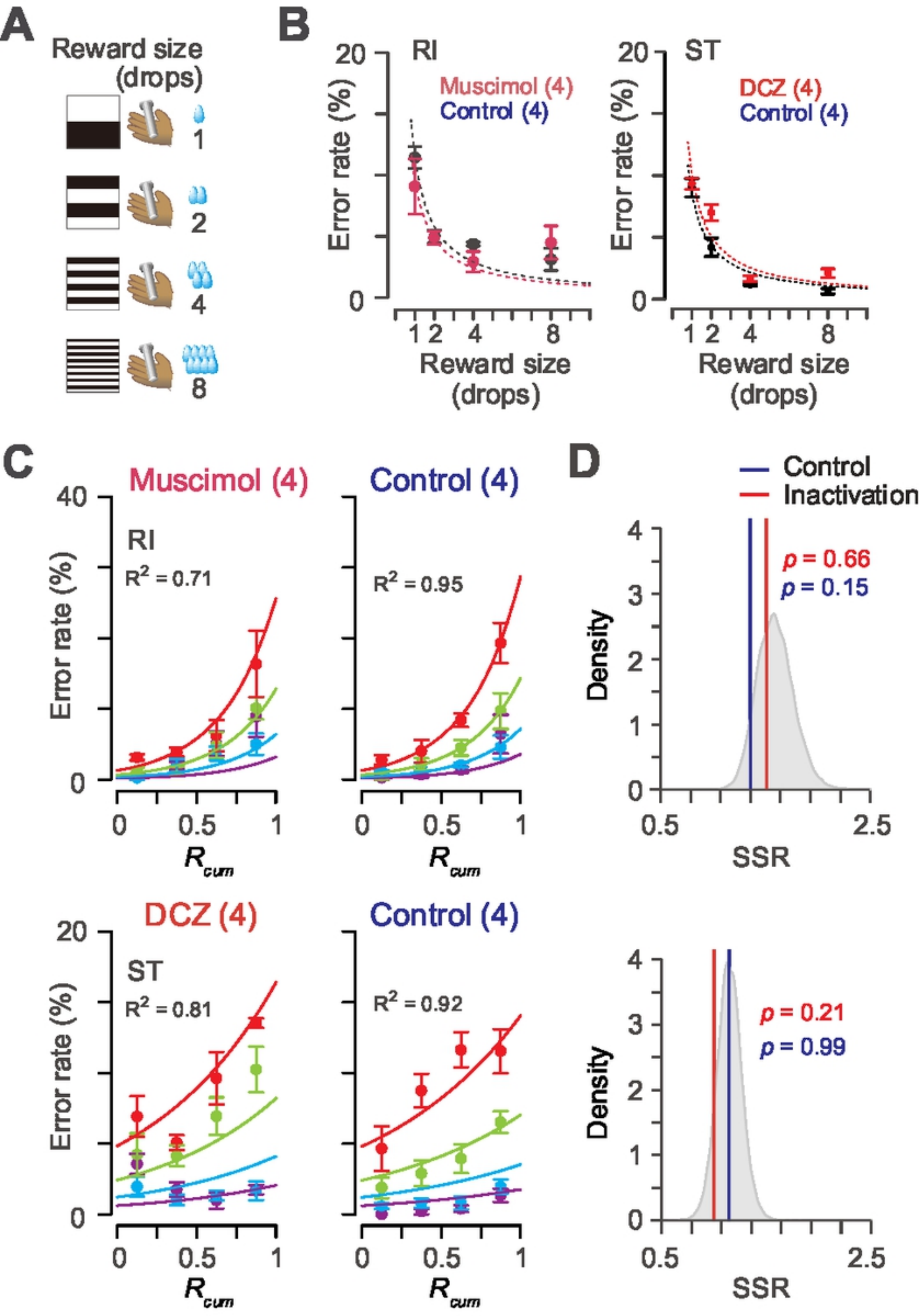
A
Reward size (drops)
1
2
4
8
B
RI
Muscimol (4)
Control (4)
Error rate (%)
Reward size (drops)
ST
DCZ (4)
Control (4)
C
Muscimol (4)
Control (4)
RI
$R^2 = 0.71$
$R^2 = 0.95$
DCZ (4)
Control (4)
ST
$R^2 = 0.81$
$R^2 = 0.92$
R_{cum}
D
Control
Inactivation
$p = 0.66$
$p = 0.15$
$p = 0.21$
$p = 0.99$
Density
SSR

Figure 9