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 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

The term ‘temporal discounting’ describes both choice preferences and motivation for delayed rewards. Here we show that neuronal activity in the dorsal part of the primate caudate head (dCDh) signals the temporally discounted value needed to compute the motivation for delayed rewards. Macaque monkeys performed an instrumental task, in which visual cues 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 markedly distorted the normal task performance based on the integration of reward size and delay, but did not affect the task performance for different reward sizes without delay. These results suggest that dCDh is involved in encoding the integrated multidimensional information critical for motivation.

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 is, the greater the motivation will be. When animals and humans suppose the reward to be delayed, their behaviors become slower and less accurate. This decline in motivation is conceptualized as discounting of reward value as a function of time, namely temporal discounting

(Minamimoto et al., 2009; Shadmehr et al., 2010; Berret and Jean 2016). Temporal discounting was originally proposed to describe choice preferences for earlier smaller rewards rather than later larger rewards (Mazur, 1984; Mazur, 2001; Green and Myerson, 2004), implying that motivation and decision-making may share common brain processes. Besides temporal discounting, motivational processes also consider internal drive for reward, such as hunger and thirst, integrating these two factors into motivational value (Toates, 1986; Berridge, 2004; Zhang et al., 2009).

One of the major candidates as the neural systems mediating the computation of expected outcome value and transforming it into action is the basal ganglia (Daw and Doya, 2006; Hikosaka et al., 2006). Several lines of evidence based on electrophysiological studies have suggested that the caudate nucleus (CD) plays an important role in motivational processing via signaling an expected outcome, and monitoring action/outcome leading to future behavioral improvement (Kawagoe et al., 1998; Cromwell and Schultz, 2003; Lau and Glimcher, 2008; Hori et al., 2009). Especially, the dorsal part of the head of the CD (dCDh) is best situated to participate in temporal discounting processes because it receives strong convergent inputs from various frontal cortical areas including the dorsolateral prefrontal cortex (DLPFC), the anterior cingulate cortex (ACC) and the supplementary eye field (SEF) (Haber et al., 1995; Haber et al., 2006), where neuronal activity is related to the expected amount or delay/proximity of rewards (Shidara

and Richmond, 2002; Roesch and Olson, 2003, 2005; Tsujimoto and Sawaguchi, 2005; Sohn and Lee, 2007; So and Stuphorn, 2010). Indeed, it has been shown that neurons in this CD sector respond in relation to temporally discounted values during intertemporal choice (Cai et al., 2011). 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 the 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 behavior, pharmacological inactivation (local muscimol injection) and chemogenetic inactivation (designer receptor exclusively activated by designer drugs, DREADDs) (Nagai et al., 2016; Roth, 2016) of dCDh were 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

74 state remained intact.

75

76 **Results**

77 **Temporal discounting accounts for monkeys' behavior**

78 We studied computation of the motivational value using temporal discounting in macaque
79 monkeys induced delaying reward delivery (Figure 1A). In the basic task, the monkey must
80 release a bar when a red spot turns green. A visual cue appears at the beginning of each trial and
81 remains on throughout. Each of the 6 cues is linked to one combination of reward size (1 small
82 drop; 3 or 4 large drops) and delay to reward (one of 0, 3.3, and 6.9 seconds; Figure 1B). In this
83 and similar tasks, the error rate, i.e., the proportion of trials with an incorrect response (either
84 releasing the bar too early or too late), reflects the monkey's motivation for action, which can be
85 interpreted as the motivational value or decision utility for whether to act or not, according to its
86 prediction about the forthcoming reward. In our previous studies, the error rate was inversely
87 related to the motivational value (Minamimoto et al., 2009). In previous behavioral studies, the
88 subjective value of delayed reward was formulated as a hyperbolic discounting model (Mazur,
89 1984; Mazur, 2001; Green and Myerson, 2004),

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

90 where *DV* 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 Equation 2, with error rates (*E*), reward size (*R*), delay (*D*), and a monkey-specific free-fitting parameter (*a*) (Minamimoto et al., 2009),

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

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 averaged error rates was well described by the inverse relation with hyperbolic delay discounting (Equation 2) (R^2 = 0.96, 0.88, and 0.94 for monkeys BI, FG and ST, respectively; Figure 1C, solid lines). The exponential discounting model (Equation 3) also explained the majority of the cases (7/10 monkeys, $R^2 > 0.9$; e.g., Figure 1C, dotted curves for monkeys BI and ST) well. Consistent with previous results (Minamimoto et al., 2009), leave-one-out cross-validation analysis confirmed that the hyperbolic model fitted the error rates significantly better than exponential function for all three monkeys as well as for seven additional monkeys ($p < 0.05$; see Materials and methods).

Figure 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 Equations 2 and 3, respectively. Note that since two straight lines were simultaneously fitted to the averaged data, the fitting was worse for the data of trials with larger rewards. (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 dCDh. CD, caudate nucleus; Put, putamen.

Figure supplement 1. Error type and timing, and reaction time.

Figure supplement 2. Eye position during cue period.

Source data 1.

The proportion of early errors differed across monkeys, but was relatively consistent within each monkey (Figure 1—figure supplement 1A). Nine of ten monkeys exhibited a pattern in which early errors increased over time, reaching a peak at about 0.7s or 1.8 s after cue onset, while only one monkey (monkey TM) showed an increase in early errors immediately after cue onset. These results suggest that early errors were not rejection responses, but rather the consequence of insufficient motivation to make the correct response. In addition, the late releases did not always occur immediately after the end of the 1s-response window, suggesting that they were not due to extensions of slow reaction (Figure 1—figure supplement 1B). These results also support the interpretation that errors are caused by insufficient motivation to respond correctly.

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, Figure 1—figure supplement 1C). 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, $\eta^2 = 0.003$) (Figure 1—figure supplement 2).

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 dCDh and recorded the activity of 150 presumed projection neurons (i.e., phasically active neurons; see Materials and methods) (Figure 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) (Figure 2A-C; $p < 0.05$, χ^2 test). The cue response was the most prominent activity of

dCDh neurons during the task (Figure 2A and C); 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 (Figure 2D; 49/150; $p < 0.0001$, χ^2 test).

Figure 2. Task-related responses of dCDh 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 (Equation 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.

Figure 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 (DV, Equation 1) for neuronal activities shown in (A) and (B), respectively.
Figure supplement 1. Error trial analysis.
Figure supplement 2. Error trial analysis.

We related spike discharge rates to DV estimated using the hyperbolic function obtained from individual behavior (Equations 2 and 7). The firing rate during the cue period of example neurons (Figure 3A and B) correlated with DV positively (Figure 3C, $R^2 = 0.86$, $p < 0.01$) and negatively (Figure 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 monkeys BI, FG, and ST, respectively); 18 and 9 exhibited positive and negative correlations, respectively. The result did not seem to depend on the shape of DV function: a similar number of neurons showed a significant DV relation when estimating using the exponential function (Equation 3; n = 25). By contrast, significant DV relation was relatively minor in release-related (5/49) and reward responses (3/49). The DV relation was not likely to be a direct reflection of the eye movement or gaze variables, since the monkeys tended to looked at cue location from cue to go signal regardless of rewarding condition (Figure 1—figure supplement 2).

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 (Equation 8). We found that only 3 and 4 neurons showed a significant exclusive effect of size or delay on their cue response, respectively (Figure 4A and B, blue and green). In contrast, for 19 and 5 neurons, DV and both delay and size had a significant effect on the cue response, respectively (Figure 4A and B, 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 (Figure 4—figure supplement 1).

Figure 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.

Figure supplement 1. Impact of DV and comparison with delay and size on release and reward response.

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 dCDh (Figure 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 (Equation 7) was calculated (200-ms window, 10-ms steps) for each DV coding neuron (Figure 5A and B). On average, the effect size rose from 100 ms after cue onset, reaching a peak at 750 ms after the cue (red curve, Figure 5C). Thereafter, it gradually decreased to the bar release (Figure 5D). The effect size did not become 0, indicating that a few neurons (n = 5) also signaled DV around bar release. 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. We postulated that the activity of DV-coding neurons may be related to the process mediating outcome prediction and further the decision to act or not. If this is the case, the DV coding neurons should behave differentially

between correct and error trials. To test this, we performed linear mixed model (LMM) analysis on 22 of 27 DV-coding neurons recorded in a session in which the monkeys made at least three error trials. We found that the majority of DV-coding neurons (17 of 22) were modulated differentially by DV depending on whether the monkey performed correctly or not (Figure 3—figure supplements 1 and 2), supporting the idea that this population of neurons is involved in motivational processes.

Figure 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, n = 27) and non-DV coding neurons (black, n = 73) 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). (E-F) Time course of normalized activity for DV coding (red, n = 27) and non-DV coding neurons (black, n = 73) aligned by cue onset and bar release, respectively. Conventions are the same as C-D.
Source data 1.

Non-DV coding neurons, on the other hand, did not change the effect size from 0 during the cue period, whereas it increased after the bar release (black curve, Figure 5C and D). Comparing the normalized activity of these two populations, whereas DV coding neurons showed an increase in activity toward the bar release, non-DV coding neurons showed a marked transient response to the cue (Figure 5E and F). Given that the monkeys tended to look at the cue location during cue

period (Figure 1—figure supplement 2B), the activity of non-DV coding neurons appeared to largely reflect visual response, but was unlikely to be evoked by eye movement. This suggests that non-DV coding neurons might have a role in detecting cue appearance.

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 (Figure 6A). When we looked at the data from one quarter (e.g., Figure 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 (Toates, 1986; Berridge, 2004; Zhang et al., 2009) (Equation 6; see Materials and methods). The error rates were well explained by Equation 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, Figure 6A and B). The satiation effect, $F(R_{cum})$ (Figure 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.

Figure 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 Equation 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 and average across 9 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.
Figure supplement 1. Impact of discounted value and satiation on cue response.

Although satiation significantly influenced behavior, satiation did not influence dCDh activity, not even when coding DV. When we compared the cue responses between the 1st and 2nd halves of 120 successful trials, the activity patterns were indistinguishable between the 1st and 2nd halves of the recording period in a single neuron (Figure 6D). Similarly, the normalized mean discharge rate of cue responses for each reward condition did not significantly change between the 1st and 2nd halves 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; Figure 6E). Additional neuron-by-neuron analysis using a multiple linear regression model (Equation 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 (Figure 6—figure supplement 1). 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 (Nagai et al., 2016). 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 (Figure 7A and B; see Figure 1D for comparison). Another monkey received injections of a viral vector expressing an inhibitory DREADD, hM4Di, into the dCDh bilaterally. A positron emission tomography (PET) scan with a DREADD-selective radioligand, [^{11}C]deschloroclozapine (DCZ) (Nagai et al., 2020), confirmed that hM4Di expression covered

316 the dCDh (Figure 7C). Chemogenetic silencing was achieved by systemic administration of the
317 selective DREADD agonist DCZ (Nagai et al., 2020). Both pharmacological and chemogenetic
318 inactivation resulted in a significant shift in error rate patterns with respect to reward size and
319 delay in all three monkeys (Figure 7D, left); the behavioral patterns were idiosyncratic across the
320 monkeys, but they were generally not in accordance with the temporal discounting model (i.e.,
321 Equation 2; R^2 = 0.41, 0.19 and 0.76, for monkeys BI, RI and ST, respectively). By contrast, the
322 error rate pattern following vehicle injection remained well explained by the model (Figure 7D,
323 right; $R^2 > 0.86$).

324

325 **Figure 7.** Bilateral inactivation of dCDh disrupted normal motivational performance based on
326 size and delay. (A) CT-based localization of muscimol injection sites. CT image visualizing
327 injection cannulae targeting CD bilaterally (hot color) overlaid on MR image (gray scale) in
328 monkey BI. Scale bar, 5 mm. (B) Muscimol (magenta) and saline injection sites (blue) are
329 mapped by estimating diffusion (4 mm in diameter) from the tip of the cannula. The data of two
330 subjects are overlaid and are separately mapped 3 mm anterior and 3 mm posterior to the anterior
331 commissure (AC). (C) [^{11}C]DCZ-PET visualizing hM4Di expression *in vivo* in monkey ST.
332 Parametric image of specific binding (BP_{ND}) of [^{11}C]DCZ-PET overlaying MR image. Scale bar,
333 5 mm. (D) Error rates (mean ± sem) as function of delay duration under inactivation (left) and
334 control condition (right). Black and red symbols are low and high reward trials, respectively.
335 Dotted lines represent best-fit function of hyperbolic temporal discounting (Equation 2). Number
336 in parentheses indicates number of sessions tested. (E) Distribution of sum of squared residuals
337 (SSR) of best-fit function (Equation 2) to averaged resample data obtained by bootstrap method
338 (n=20,000). Blue and red lines indicate SSR of best-fit of Equation 2 to mean error rates in control
339 and inactivation sessions, respectively.

Figure 7—figure supplement 1. No significant effects of dCDh inactivation on reaction time in delayed reward task.

Figure 7—figure supplement 2. No effect of dCDh inactivation on eye position.

Figure 7—figure supplement 3. Normalized error rates in baseline, control and inactivation session of delayed reward task.

Figure 7—figure supplement 4. Effect of dCDh inactivation on satiation.

Source data 1.

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 in each monkey (inactivation vs. control; Brunner-Munzel test, $p > 0.18$). There was no significant effect of treatment on reaction time in two monkeys (two-way ANOVA, effect size of treatment: $\eta^2 < 0.01$, monkeys BI and RI) but a moderate effect of treatment in monkey ST ($\eta^2 = 0.06$) (Figure 7—figure supplement 1). Type of error (i.e., releasing too early or too late) was unaffected by inactivation (main effect of treatment, $F_{(1,26)} = 1.07$, $p = 0.31$). The monkeys touched and released the bar several times during the delay period, even though the delay time was not shortened. The number of releases depended on reward condition (main effect of reward condition, $F_{(5, 143)} = 25.22$, $p < 0.001$), but there was no significant main effect of treatment (two-way ANOVA, treatment, $F_{(1, 143)} = 2.90$, $p = 0.09$) or

interaction ($F_{(5, 143)}$ = 0.42 , p = 0.83). The duration of gazing at the cue was slightly but not significantly longer during muscimol inactivation (t-test, p = 0.063, Figure 7—figure supplement 2). Together, the bilateral inactivation of dCDh did not cause impairments in overall motivation, motor, or anticipatory behavior.

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 (Figure 7—figure supplement 3). 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 ($p < 0.05$, Figure 7E, red line), suggesting that dCDh silencing distorted normal motivational value processing based on the integration between reward size and delay.

To examine the effect of dCDh inactivation on satiation, we plotted error rates along with the normalized cumulative reward (R_{cum}). Like the results shown in Figure 6A, the error rate in each combination of reward size and delay increased as R_{cum} increased in baseline and vehicle control sessions (Figure 7—figure supplement 4). Satiation-dependent increase in error rates was also observed in two of three monkeys in dCDh inactivation, while monkey ST failed to show this tendency (Figure 7—figure supplement 4). We also examined trial initiation time (duration

between the time the reward was received and the start of the next trial), reflecting satiation effects as a measure of motivation to start time in a previous study (Fujimoto et al., 2019). In both control and inactivation sessions, the trial initiation time was significantly longer in the second half of the session (two-way ANOVA, main effect of 1st vs 2nd, $F_{(1,54)} = 4.32$, $p = 0.042$), where no significant interactive effect of dCDh inactivation was observed (1st vs 2nd × treatment, $F_{(1,54)} = 0.32$, $p = 0.57$). These results suggest that dCDh inactivation does not have a strong effect on satiation.

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 (Figure 8A). It has been repeatedly shown that the error rates of this task will be well explained by the joint function of reward size and satiation (Equation 6) (Minamimoto et al., 2009; Minamimoto et al., 2012b; Fujimoto et al., 2019). 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 ($R^2 > 0.7$)(Figure 8C) and were equally well compared with the baseline data ($p > 0.15$, bootstrap significance test; Figure 8D). 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, $F_{(3, 243)} = 1.69$, $p = 0.17$). The lack of change in the error rate pattern in the reward-size task could be attributed to the relative ease of associating cues with outcome compared to the delayed reward task. However, no clear difference was evident between the two tasks in establishing the cue-outcome relationship as judged by the behavior during the training period (Figure 8—figure supplement 1). Overall, these results suggest that dCDh activity is specifically involved in computing the motivational value based on delay discounting, rather than general motivational processes based on the integration of incentives and drive.

Figure 8. 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) *top:* Error rates (mean ± sem) as function of reward size in muscimol treatment (magenta) and non-treatment control session (black) for monkey RI, respectively. *bottom:* 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. Number in parentheses indicates numbers of sessions tested. (D) Distribution of sum of squared residuals (SSR) of best-fit function (Equation 6) to averaged resample data obtained by bootstrap method (n=20,000).

421 Blue and red lines indicate SSR of best-fit of Equation 6 to the mean error rates in control and
422 inactivation sessions, respectively.
423 **Figure supplement 1.** Comparison of learning in reward size and delayed reward task.
424 **Source data 1.**
425

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.

Previous studies have suggested that the neuronal activity in the CD is involved in translating value into action by signaling multi-dimensional aspects of reward-related information, including presence/absence (Kawagoe et al., 1998; Cromwell and Schultz, 2003), probability (Lau and Glimcher, 2008; Oyama et al., 2010; White and Monosov, 2016), and size of reward (Nakamura et al., 2012; Fujimoto et al., 2019). Neurons in dCDh reflect the action values of a specific

movement (Samejima et al., 2005; Lau and Glimcher, 2008) and might contribute to selecting an action that maximizes future rewards. In the present study, we found that the cue responses of a subpopulation of dCDh neurons reflected temporally discounted values that were inferred from the individual behaviors. It could not be a simple reflection of physical features of a visual cue, since the neuronal signal was observed irrespective of the 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). It has also been suggested that the basal ganglia are involved in assessing information processing for the duration of events or actions. Neuronal signals reflecting the duration of past events related to temporal discrimination were found in the anterior striatum including CD (Chiba et al., 2008). The CD neurons also showed ramping-up activity in response to stimuli that predict timing of action initiation (Suzuki and Tanaka, 2019). It might not be surprising that the neuronal signal reported here was related not only to the forthcoming reward timing, but also to the reward size, hence representing DV. Although it has been reported eye movement-related activity of CD neurons modulated by forthcoming rewarding conditions (Watanabe et al., 2003), the DV signal observed here could not be a direct reflection of eye movements or gaze variables, since the monkeys constantly looked at the cue during the cue period regardless of the rewarding condition (Figure 1—figure supplement 2).

The DV signal emerged just after cue onset, gradually increased, and then disappeared

462 before execution of the action (Figure 5). This time course suggests that the neuronal signal does
463 not simply convey the Pavlovian value of the cue, but can be related to the cognitive process
464 mediating the outcome prediction underlying the decision of whether to act or not. This was
465 supported by the results of the error trial analysis, which showed that most of the DV-coding
466 neurons behaved differently between correct and error (Figure 3—figure supplement 1).
467 Compared with DV, the effect of the reward size or delay duration on cue responses was
468 relatively weak (Figure 4), indicating that the signal integration may take place at least partially in
469 some upstream brain area(s). The first plausible source of temporal discounting is a
470 prefronto-striatal projection. Our recordings were carried out from dCDh, the region receiving
471 direct input from the frontal cortical areas including DLPFC, ACC, and SEF (Selemon and
472 Goldman-Rakic, 1985; Calzavara et al., 2007; Averbeck et al., 2014). DLPFC neurons encode
473 DV as well as reward, delay duration, and target position during an intertemporal choice task
474 (Kalenscher et al., 2005; Kim et al., 2008), exhibiting strong modulations in response to the delay
475 combined with the amount of reward (Tsujimoto and Sawaguchi, 2005; Sohn and Lee, 2007;
476 Hosokawa et al., 2013). The activity in ACC reflects the expected amount of reward (Knutson et
477 al., 2005; Amiez et al., 2006) and the delay/proximity of rewards (Shidara and Richmond, 2002),
478 as well as delay discounting for reward (McClure et al., 2007). Neurons in SEF are also
479 modulated by the amount of reward and delay duration (Roesch and Olson, 2003, 2005; So and

Stuphorn, 2010). 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 (Kobayashi and Schultz, 2008). The third possible source is a thalamostriatal input arising from thalamic nuclei, including the centromedian-parafascicular (CM-Pf) complex (Smith et al., 2004). Neuronal activity in CM reflects the predicted outcome value (Minamimoto et al., 2005), 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 (Figure 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 (Toates, 1986; Berridge, 2004). 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 (Figure 6 and Figure 6—figure supplement 1). A similar insensitivity to satiation has been reported in terms of cue-related activity in the ventral striatum that was correlated with reward value (Roesch et al., 2009). 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 (Rolls et al., 1989; Critchley and Rolls, 1996; Bouret and Richmond, 2010), rostromedial caudate nucleus (rmCD) and ventral pallidum (VP) (Fujimoto et al., 2019). Perhaps satiety-related signals would be represented in a network believed to be critical for guiding a choice of food based on internal drive (Izquierdo and Murray, 2010; Murray and Rudebeck, 2013). 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

516 silencing primate striatal activity (Nagai et al., 2016). Muscimol inactivation is a standard
517 procedure that has repeatedly been used in monkey studies. It has, however, major drawbacks: (1)
518 the extent of an effective area is difficult to be controlled or identified (although we monitored the
519 location of injection sites by computed tomography (CT)), and (2) when the experiments are
520 repeated, mechanical damage to tissue would accumulate. The chemogenetic tool DREADDs, on
521 the other hand, overcomes these problems; once a silencing DREADDs, hM4Di, is delivered,
522 substantially the same neuronal population can be inactivated non-invasively and the effective
523 region can be confirmed by PET imaging, as demonstrated here, and by traditional post-mortem
524 histochemistry. We found that inactivation of dCDh by either method produced consistent
525 behavioral impairments; inactivation abolished the normal pattern of error rates derived from the
526 integration of reward size and delay duration (Figure 7). This impairment cannot be explained
527 simply by changes in the temporal discounting rate or alterations in the evaluation of single
528 incentive factors. Our results are consistent with previous findings that both lesioning and
529 inactivation of the dorsomedial striatum in rats, a homologue of dCDh in primates, reduced the
530 sensitivity of instrumental performance to shifts in the outcome value (Yin et al., 2005; Yin et al.,
531 2008). In contrast, dCDh inactivation did not impair motivation based on reward size alone or
532 according to the integration of reward size and physiological state (i.e., motivational value; Figure
533 8). Thus, 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 (Simmons et al., 2010). 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 (Haber et al., 2006), since ablation or inactivation of these related areas abolished the normal relationship between reward size and error rate in the reward-size task (Simmons et al., 2010; Nagai et al., 2016; Fujimoto et al., 2019). 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 (Balleine et al., 2007; Haber and Knutson, 2010). 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 (Pulcu et al., 2014; Amlung et al., 2019). Human imaging studies have revealed the structural and functional

552 connectivity between DLPFC and the striatum with the individual differences in temporal
553 discounting (van den Bos et al., 2014, 2015). Since silencing dCDh did not induce impulsivity
554 (steepened temporal discounting or facilitating reaction was not observed), it could be difficult in
555 the present study to address the link between dCDh activity and mechanisms underlying
556 impulsivity. Nevertheless, our findings may provide a framework to elucidate dysregulation of
557 motivational systems in impulsive individuals with psychiatric disorders.

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

562

563 **Materials and methods**

564 **Subjects**

565 Ten male rhesus macaque monkeys (5 -11 kg) were used in this study. Of these, three (BI, FG and
566 ST) were also used for recording, and one (ST) and two (BI and RI) for chemogenetic and
567 pharmacological inactivation experiments, respectively. All surgical and experimental
568 procedures were approved by the National Institutes for Quantum and Radiological Science and
569 Technology (11-1038-11) and by the Animal Care and Use Committee of the National Institute of

Mental Health (Annual Report ZIAMH002619), 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) (Hays et al., 1982). Neurobehavioral Systems Presentation software was used to display visual stimuli (Neurobehavioral Systems).

All monkeys were trained and tested with the delayed reward task (Figure 1A and B) (Minamimoto et al., 2009). 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 (Blanchard et al., 2013)], because monkeys are insensitive to post-reward delays in our tasks (please see Figure 3 in Minamimoto et al., 2009). 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) (Figure 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 followed by learning of each task for 1-3 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) were 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)

606 varied and was assigned by unique cue (Figure. 8A) (Minamimoto et al., 2009). The sequence and
607 timing of events were the same as those in the delayed reward task.

608 **Surgery**

609 After behavioral training, magnetic resonance (MR) images at 1.5T (monkey FG) and 7T
610 (monkeys BI, RI and ST) were obtained under anesthesia (intravenous infusion of propofol
611 0.2-0.6 mg/kg/min, or pentobarbital sodium 15-30 mg/kg) to determine the position of the
612 recording or local injection. After obtaining each MR image, a surgical procedure was carried out
613 under general isoflurane anesthesia (1~2%) to implant chambers for unit recording and/or
614 chemical inactivation. For monkeys BI and FG, we implanted a rectangle chamber (22 x 22 mm
615 ID; KDS Ltd.) from vertical in the coronal plane aiming for the bilateral CD. We implanted one or
616 two cylinder chambers (19 mm ID; Crist Instrument Co., Inc.) angled 10° or 20° from vertical in
617 the coronal plane targeting the right or bilateral CD for monkeys ST and RI, respectively. Based
618 on measurements made from the MR images, the centers of the chambers were placed to target
619 the CD near the anterior commissure. A post for head fixation during data collection was also
620 implanted.

621 **Recording neuronal activity and mapping recording location**

622 Single-unit activity was recorded (51, 31, and 68 from monkeys BI, FG and ST, respectively)
623 while monkeys performed the delayed reward task in a block usually consisting of 120 trials.

624 Action potentials of single neurons were recorded from the left CD using epoxy-coated 1.1-1.5
625 MΩ tungsten microelectrodes (Microprobes for Life Science; 1.1-1.5 MΩ at 1 kHz) or
626 glass-coated 1.0 MΩ tungsten microelectrodes (Alpha Omega Engineering Ltd). A guide tube
627 was inserted through the grid hole in the implanted recording chamber into the brain, and the
628 electrodes were advanced through the guide tube by means of a micromanipulator (Narishige
629 MO-97A or Alpha Omega EPS). Spike sorting to isolate single neuron discharges was performed
630 with a time-window algorithm (TDT-RZ2, Tucker Davis Technologies) or custom-made software
631 written in LabVIEW (National Instruments). Striatal neuronal activities were classified into two
632 subtypes: presumed projection neurons and tonically active neurons (TANs, presumed
633 cholinergic interneurons) based on their spontaneous discharge rates and action potential
634 waveforms, as previously described (Yamada et al., 2016). We exclusively examined the activity
635 of the presumed projection neurons, which are characterized as having a low spontaneous
636 discharge rate (< 2 spikes/s) outside the task context and exhibiting phasic discharges in relation
637 to one or more behavioral task events. The activity of TANs recorded from the CD of monkeys
638 performing a similar task was reported in a previous study (Falcone et al., 2019). The timing of
639 action potentials was recorded together with all task events at millisecond precision. In the
640 inactivation study, eye movements were monitored for corneal reflection of an infrared light
641 beam through a video camera at a sampling rate of 120Hz (i_rec,

642 http://staff.aist.go.jp/k.matsuda/eye/).

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

648 **Chemogenetic inactivation**

649 One monkey (ST) received bilateral injections of an adeno-associated virus vector
650 (AAV1-hSyn-hM4Di-IRES-AcGFP; 3 μL/site; 4.7 x $10e^{13}$ particles/mL) at two locations into
651 each side of the CD. The injection procedure was as described previously (Nagai et al., 2016). 49
652 days post vector injection, the monkey underwent a PET scan with [^{11}C]DCZ to visualize *in vivo*
653 hM4Di expression. Chemogenetic silencing was achieved by intramuscular injection (i.m.) with a
654 DREADD selective agonist, DCZ (HY-42110, MedChemExpress; 0.1 mg/kg). DCZ was
655 dissolved in 2% dimethyl sulfoxide (DMSO) in saline to a final volume of 0.65 ml. DCZ solution
656 or vehicle (as control) was administered intramuscularly. Five to ten min following
657 administration, the animal was allowed to start performing the tasks, which continued for 100 min.
658 Based on a previous study, chemogenetic silencing would be effective for 15-120 min after DCZ
659 administration. We performed at most one inactivation study per week. Note that we verified that

660 the DCZ administration (0.1 mg/kg, i.m.) does not cause any significant motivational/motor
661 impairments or alteration of the incentive effect of the performance of reward-size task in
662 monkeys without expressing DREADDs (n = 3) (Nagai et al., 2020). Detailed protocols for PET
663 imaging were described elsewhere (Nagai et al., 2020).

664 **Pharmacological inactivation**

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

injection sites (Figure 1D and Figure 7A). We plotted the injection sites based on the estimate of the liquid diffusion range (4 mm diameter) reported previously (Yoshida et al., 1991; Martin and Ghez, 1999).

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 (Equation 2) or that with exponential temporal discounting (Minamimoto et al., 2009) 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 (Minamimoto et al., 2009).

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 the delayed reward task obtained from each

695 monkey and the average data across monkeys to the following model:

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

696 where the satiation effect, $F(R_{cum})$, as the reward value was exponentially decaying in R_{cum} at a

697 constant λ (Minamimoto et al., 2012a):

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

698 For modeling satiation effects of the error rates in reward-size task, we used an inverse model

699 integrating satiation effect (Equation 5), as follows:

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

700 We also applied conventional ANOVA modeling to the behavioral data. The proportional

701 behavioral data were transformed using the variance stabilizing arcsine transformation before

702 hypothesis testing (Zar, 2010).

703 The trial initiation time was defined as the duration from the reward of previous trial to the

704 time of lever grip to begin a trial, as a measure of motivation to start a trial. We compared the

705 average trial initiation time in the first and second halves of the daily session.

706 Significance of deviation from baseline data was examined by means of the parametric

707 bootstrapping method (n = 20,000). We first constructed distribution of the sum of squared

708 residuals (SSR) of the best fit of the model to the averaged resampled error rates (n = 5 or 4

709 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).

The Brunner-Munzel test was used as non-parametric analysis for median value with Bonferroni correction (Hui et al., 2008).

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 (Ravel and Richmond, 2006). 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:

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

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 (Equation 1) (Mazur, 1984; Mazur, 2001; Green and Myerson, 2004). The effect of DV was compared with that of delay and reward size information on the response by the following multiple linear regression model:

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

where D and R are delay duration and reward size, respectively, β_{delay}, β_{size}, and β_{DV} are the regression coefficients, and β_0 is the intercept. Another linear regression analysis was performed

744 to quantify the influence of temporal discounting of reward value and satiation on the response, as
745 follows:

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

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

748 To examine whether DV-coding neurons differentially behave between correct and error
749 trials, we performed linear mixed models (LMMs) (Bates et al., 2015), in which there is mixed
750 effect of trial completion (correct/error) on slope and/or intercept. Four models were nested to
751 consider the presence or absence of random effects (Fig. 4 – figure supplement 1). We applied the
752 LMM analysis on DV-coding neurons recorded in a session in which the monkeys made at least
753 three error trials. The best model was selected based on BIC.

754

755 **Acknowledgements**

756 We thank George Dold, Keiji Matsuda, Risa Suma, Tomoni Kokufuta, Anzu Maruyama, Jun
757 Kamei, Yuichi Matsuda, Ryuji Yamaguchi, Yoshio Sugii, Maki Fujiwara, and Mayuko Nakano
758 for their technical assistance.

759

760 **Competing interests**

761 The authors declare no competing interests.

762

References

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

Amlung M, Marsden E, Holshausen K, Morris V, Patel H, Vedelago L, Naish KR, Reed DD, McCabe RE (2019) Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. JAMA Psychiatry 76:1176-1186.

Averbeck BB, Lehman J, Jacobson M, Haber SN (2014) Estimates of projection overlap and zones of convergence within frontal-striatal circuits. J Neurosci 34:9497-9505.

Balleine BW, Delgado MR, Hikosaka O (2007) The role of the dorsal striatum in reward and decision-making. J Neurosci 27:8161-8165.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1-48.

Berret B, Jean F (2016) Why don't we move slower? The value of time in the neural control of action. J Neurosci 27: 1056-1070.

Berridge KC (2004) Motivation concepts in behavioral neuroscience. Physiol Behav 81:179-209.

Blanchard TC, Pearson JM, Hayden BY (2013) Postreward delays and systematic biases in measures of animal temporal discounting. Proc Natl Acad Sci U S A 110:15491-15496.

781 Bouret S, Richmond BJ (2010) Ventromedial and orbital prefrontal neurons differentially
782 encode internally and externally driven motivational values in monkeys. J Neurosci
783 30:8591-8601.
784 Cai X, Kim S, Lee D (2011) Heterogeneous coding of temporally discounted values in the
785 dorsal and ventral striatum during intertemporal choice. Neuron 69:170-182.
786 Calzavara R, Mailly P, Haber SN (2007) Relationship between the corticostriatal terminals from
787 areas 9 and 46, and those from area 8A, dorsal and rostral premotor cortex and area 24c:
788 an anatomical substrate for cognition to action. Eur J Neurosci 26:2005-2024.
789 Chiba A, Oshio K, Inase M (2008) Striatal neurons encoded temporal information in duration
790 discrimination task. Exp Brain Res 186:671-676.
791 Critchley HD, Rolls ET (1996) Hunger and satiety modify the responses of olfactory and visual
792 neurons in the primate orbitofrontal cortex. J Neurophysiol 75:1673-1686.
793 Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on
794 neuronal activity in primate striatum. J Neurophysiol 89:2823-2838.
795 Daw ND, Doya K (2006) The computational neurobiology of learning and reward. Curr Opin
796 Neurobiol 16:199-204.

797 Falcone R, Weintraub DB, Setogawa T, Wittig JH, Jr., Chen G, Richmond BJ (2019) Temporal
798 Coding of Reward Value in Monkey Ventral Striatal Tonically Active Neurons. J
799 Neurosci 39:7539-7550.
800 Fujimoto A, Hori Y, Nagai Y, Kikuchi E, Oyama K, Suhara T, Minamimoto T (2019) Signaling
801 Incentive and Drive in the Primate Ventral Pallidum for Motivational Control of
802 Goal-Directed Action. J Neurosci 39:1793-1804.
803 Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic
804 rewards. Psychol Bull 130:769-792.
805 Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging.
806 Neuropsychopharmacology 35:4-26.
807 Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal
808 circuit through the primate basal ganglia. J Neurosci 15:4851-4867.
809 Haber SN, Kim KS, Mailly P, Calzavara R (2006) Reward-related cortical inputs define a large
810 striatal region in primates that interface with associative cortical connections, providing
811 a substrate for incentive-based learning. J Neurosci 26:8368-8376.
812 Hays A, Richmond B, Optican L (1982) Unix-based multiple process system for real-time data
813 aquisition and control. WESCON 2:1-10.

814 Hikosaka O, Nakamura K, Nakahara H (2006) Basal ganglia orient eyes to reward. J
815 Neurophysiol 95:567-584.
816 Hori Y, Minamimoto T, Kimura M (2009) Neuronal encoding of reward value and direction of
817 actions in the primate putamen. J Neurophysiol 102:3530-3543.
818 Hosokawa T, Kennerley SW, Sloan J, Wallis JD (2013) Single-neuron mechanisms underlying
819 cost-benefit analysis in frontal cortex. J Neurosci 33:17385-17397.
820 Hui W, Gel YR, Gastwirth J, L. (2008) lawstat: An R Package for Law, Public Policy and
821 Biostatistics. Journal of Statistical Software 28:6906-6932.
822 Izquierdo A, Murray EA (2010) Functional interaction of medial mediodorsal thalamic nucleus
823 but not nucleus accumbens with amygdala and orbital prefrontal cortex is essential for
824 adaptive response selection after reinforcer devaluation. J Neurosci 30:661-669.
825 Kalenscher T, Windmann S, Diekamp B, Rose J, Gunturkun O, Colombo M (2005) Single units
826 in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice
827 task. Curr Biol 15:594-602.
828 Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive
829 signals in the basal ganglia. Nat Neurosci 1:411-416.
830 Kim S, Hwang J, Lee D (2008) Prefrontal coding of temporally discounted values during
831 intertemporal choice. Neuron 59:161-172.

832 Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005) Distributed neural
833 representation of expected value. J Neurosci 25:4806-4812.
834 Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. J
835 Neurosci 28:7837-7846.
836 Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching
837 behavior. Neuron 58:451-463.
838 Martin JH, Ghez C (1999) Pharmacological inactivation in the analysis of the central control of
839 movement. J Neurosci Methods 86:145-159.
840 Mazur JE (1984) Tests of an equivalence rule for fixed and variable reinforcer delays. Journal of
841 Experimental Psychology: Animal Behavior Processes 10:426-436.
842 Mazur JE (2001) Hyperbolic value addition and general models of animal choice. Psychol Rev
843 108:96-112.
844 McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD (2007) Time discounting for
845 primary rewards. J Neurosci 27:5796-5804.
846 Minamimoto T, Hori Y, Kimura M (2005) Complementary process to response bias in the
847 centromedian nucleus of the thalamus. Science 308:1798-1801.

848 Minamimoto T, La Camera G, Richmond BJ (2009) Measuring and modeling the interaction
849 among reward size, delay to reward, and satiation level on motivation in monkeys. J
850 Neurophysiol 101:437-447.
851 Minamimoto T, Hori Y, Richmond BJ (2012a) Is working more costly than waiting in
852 monkeys? PLoS One 7:e48434.
853 Minamimoto T, Yamada H, Hori Y, Suhara T (2012b) Hydration level is an internal variable for
854 computing motivation to obtain water rewards in monkeys. Exp Brain Res 218:609-618.
855 Murray EA, Rudebeck PH (2013) The drive to strive: goal generation based on current needs.
856 Front Neurosci 7:112.
857 Nagai Y et al. (2016) PET imaging-guided chemogenetic silencing reveals a critical role of
858 primate rostromedial caudate in reward evaluation. Nat Commun 7:13605.
859 Nagai Y et al. (2020) Deschloroclozapine, a potent and selective chemogenetic actuator enables
860 rapid neuronal and behavioral modulations in mice and monkeys. Nat Neurosci
861 23:1157-1167.
862 Nakamura K, Santos GS, Matsuzaki R, Nakahara H (2012) Differential reward coding in the
863 subdivisions of the primate caudate during an oculomotor task. J Neurosci
864 32:15963-15982.

865 Oyama K, Hernadi I, Iijima T, Tsutsui K (2010) Reward prediction error coding in dorsal
866 striatal neurons. J Neurosci 30:11447-11457.
867 Pulcu E, Trotter PD, Thomas EJ, McFarquhar M, Juhasz G, Sahakian BJ, Deakin JF, Zahn R,
868 Anderson IM, Elliott R (2014) Temporal discounting in major depressive disorder.
869 Psychol Med 44:1825-1834.
870 Ravel S, Richmond BJ (2006) Dopamine neuronal responses in monkeys performing visually
871 cued reward schedules. Eur J Neurosci 24:277-290.
872 Roesch MR, Olson CR (2003) Impact of expected reward on neuronal activity in prefrontal
873 cortex, frontal and supplementary eye fields and premotor cortex. J Neurophysiol
874 90:1766-1789.
875 Roesch MR, Olson CR (2005) Neuronal activity dependent on anticipated and elapsed delay in
876 macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. J
877 Neurophysiol 94:1469-1497.
878 Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G (2009) Ventral striatal neurons
879 encode the value of the chosen action in rats deciding between differently delayed or
880 sized rewards. J Neurosci 29:13365-13376.

881 Rolls ET, Sienkiewicz ZJ, Yaxley S (1989) Hunger Modulates the Responses to Gustatory
882 Stimuli of Single Neurons in the Caudolateral Orbitofrontal Cortex of the Macaque
883 Monkey. Eur J Neurosci 1:53-60.
884 Roth BL (2016) DREADDs for Neuroscientists. Neuron 89:683-694.
885 Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward
886 values in the striatum. Science 310:1337-1340.
887 Selemon LD, Goldman-Rakic PS (1985) Longitudinal topography and interdigitation of
888 corticostriatal projections in the rhesus monkey. J Neurosci 5:776-794.
889 Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY (2010) Temporal discounting of
890 reward and the cost of time in motor control. J Neurosci 30:10507-10516.
891 Shidara M, Richmond BJ (2002) Anterior cingulate: single neuronal signals related to degree of
892 reward expectancy. Science 296:1709-1711.
893 Simmons JM, Minamimoto T, Murray EA, Richmond BJ (2010) Selective ablations reveal that
894 orbital and lateral prefrontal cortex play different roles in estimating predicted reward
895 value. J Neurosci 30:15878-15887.
896 Smith Y, Raju DV, Pare JF, Sidibe M (2004) The thalamostriatal system: a highly specific
897 network of the basal ganglia circuitry. Trends Neurosci 27:520-527.

So NY, Stuphorn V (2010) Supplementary eye field encodes option and action value for saccades with variable reward. J Neurophysiol 104:2634-2653.

Sohn JW, Lee D (2007) Order-dependent modulation of directional signals in the supplementary and presupplementary motor areas. J Neurosci 27:13655-13666.

Suzuki TW, Tanaka M (2019) Neural oscillations in the primate caudate nucleus correlate with different preparatory states for temporal production. Commun Biol 2:102.

Toates F (1986) Motivational Systems. Cambridge: Cambridge University Press.

Tsujimoto S, Sawaguchi T (2005) Neuronal activity representing temporal prediction of reward in the primate prefrontal cortex. J Neurophysiol 93:3687-3692.

van den Bos W, Rodriguez CA, Schweitzer JB, McClure SM (2014) Connectivity strength of dissociable striatal tracts predict individual differences in temporal discounting. J Neurosci 34:10298-10310.

van den Bos W, Rodriguez CA, Schweitzer JB, McClure SM (2015) Adolescent impatience decreases with increased frontostriatal connectivity. Proc Natl Acad Sci U S A 112:E3765-3774.

Watanabe K, Lauwereyns J, Hikosaka O (2003) Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. J Neurosci 23:10052-10057.

White JK, Monosov IE (2016) Neurons in the primate dorsal striatum signal the uncertainty of object-reward associations. Nat Commun 7:12735.

Yamada H, Inokawa H, Hori Y, Pan X, Matsuzaki R, Nakamura K, Samejima K, Shidara M, Kimura M, Sakagami M, Minamimoto T (2016) Characteristics of fast-spiking neurons in the striatum of behaving monkeys. Neurosci Res 105:2-18.

Yin HH, Ostlund SB, Balleine BW (2008) Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. Eur J Neurosci 28:1437-1448.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513-523.

Yoshida M, Nagatsuka Y, Muramatsu S, Niijima K (1991) Differential roles of the caudate nucleus and putamen in motor behavior of the cat as investigated by local injection of GABA antagonists. Neurosci Res 10:34-51.

Zar JH (2010) Biostatistical Analysis: Prentice Hall.

Zhang J, Berridge KC, Tindell AJ, Smith KS, Aldridge JW (2009) A neural computational model of incentive salience. Plos Comput Biol 5:e1000437.

Figure Supplements

Figure 1—figure supplement 1. Error type and timing, and reaction time and eye position.
(A) Proportion of early error for each monkey. Thick and thin dots indicate mean and data of each session, respectively. (B) Distribution of timing for early and late bar release for each monkey. Red arrows indicate the timing of go. (C) 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.

Figure 1—figure supplement 2. Eye position during cue period.
(A) 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. (B) Time course of the proportion of eye position within the cue area aligned by CUE (left) and GO onset (right). Thick curves and shaded areas represent mean and SD, respectively. Colors represent rewarding condition.

Figure 1—source data 1.

Figure 3—figure supplement 1. Error trial analysis.
Table shows that the number of neurons whose activity is explained best by models 1–4. Note that linear mixed model (LMM) analysis was applied to 22 of 27 DV-coding neurons recorded in a session in which the monkeys made at least three error trials. *fr*, firing rate; *dv*, discounted value; *trial*, trial type (correct or error). Seventeen neurons were differently modulated by DV depending on whether the monkey perform correct or not, while remaining 5 were similarly modulated regardless of performance.

Figure 3—figure supplement 2. Error trial analysis.
Example of differential activity between error and correct trials of a DV-coding neuron. Thin and thick dots indicate relationship between firing rate and temporally discounted value (Equation 1) in individual trials and mean values for each rewarding condition, respectively. Color indicate correct (red) and error (green) trials, respectively. Thick lines indicate best-fit of LMM (model 4 in Figure 3–figure supplement 1).

964

965 **Figure 4—figure supplement 1.** Impact of DV and comparison with delay and size on release
966 and reward response.
967 (A) Scatterplot of standardized partial regression coefficients (SPRC) of DV (ordinate) against
968 those of size and delay on release response, respectively (abscissa). (B) Same as A, but for reward
969 response. Colored dots indicate neurons with significant ($p < 0.05$) coefficient, while gray dots
970 correspond to neurons without any significant effect (NA). DV/DV & Other, neurons with
971 significant coefficient of DV; Size & Delay, those with both size and delay; Size, those with size
972 exclusively; Delay, those with delay exclusively. Numbers in parentheses indicate number of
973 neurons.

974

975 **Figure 5—source data 1.**

976

977 **Figure 6—figure supplement 1.** Impact of discounted value and satiation on cue response.
978 (A) Scatterplot of standardized regression coefficients (SRC) of discharge rates during cue period
979 for DV (ordinate) against those for cumulative reward (abscissa). Red dots indicate DV-coding
980 neurons. Red and blue, and purple circles indicate non-DV coding neurons with significant ($p <$
981 0.05) coefficient for DV and cumulative reward (CR), and both respectively. Black circles
982 correspond to neurons without any significant effect (NS). (B) Representative waveforms (mean
983 ± SD) recorded from a CD neuron (Monkey ST #10) during first (purple) and last quartile
984 (orange) of recording period. Changes in firing rate were not attributable to alteration in action
985 potential isolation.

986

987 **Figure 7—figure supplement 1.** No significant effects of dCDh inactivation on reaction time in
988 delayed reward task.
989 Comparison of reaction time (mean ± SD) between baseline, control and inactivation session in
990 monkeys BI, RI and ST.

991

992 **Figure 7—figure supplement 2.** No effect of dCDh inactivation on eye position.

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.

Figure 7—figure supplement 3. 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.

Figure 7—figure supplement 4. Effect of dCDh inactivation on satiation.

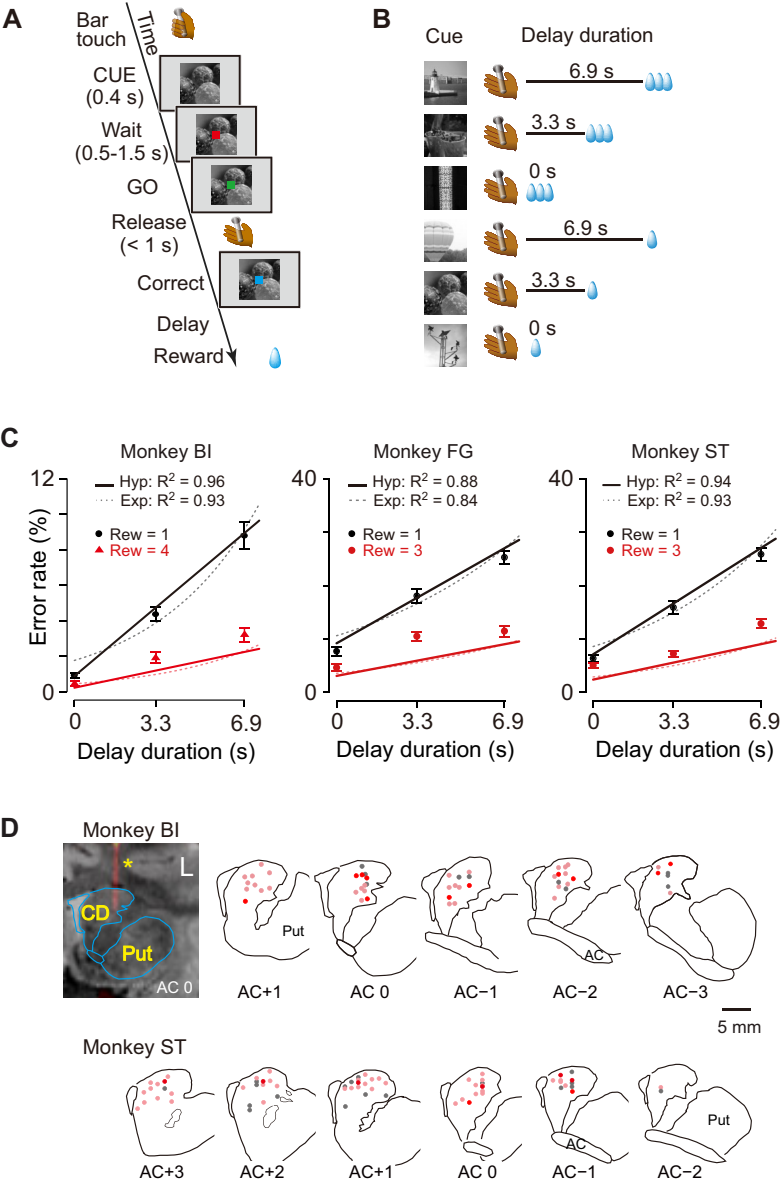
Error rates (mean ± sem) as a function of normalized cumulative reward (R_{cum}) in baseline, control and inactivation session of delayed reward task. Dotted curves are the best fit of Equation 4 to the data. Note that the satiation effect was disrupted in the inactivation session in the monkey ST, but remained normal after inactivation in the reward-size task in the same monkey (see Fig. 8C).

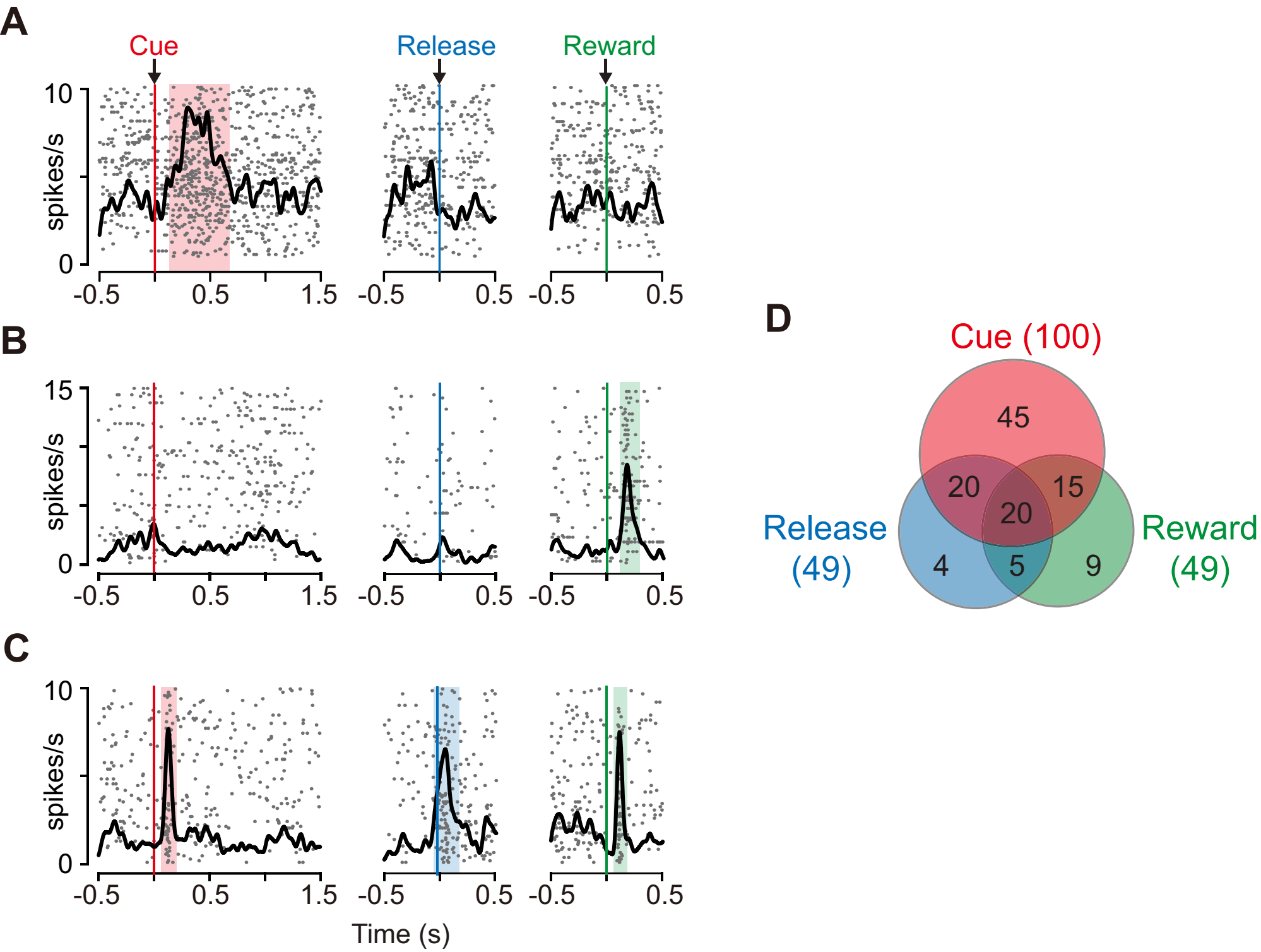
Figure 7—source data 1.

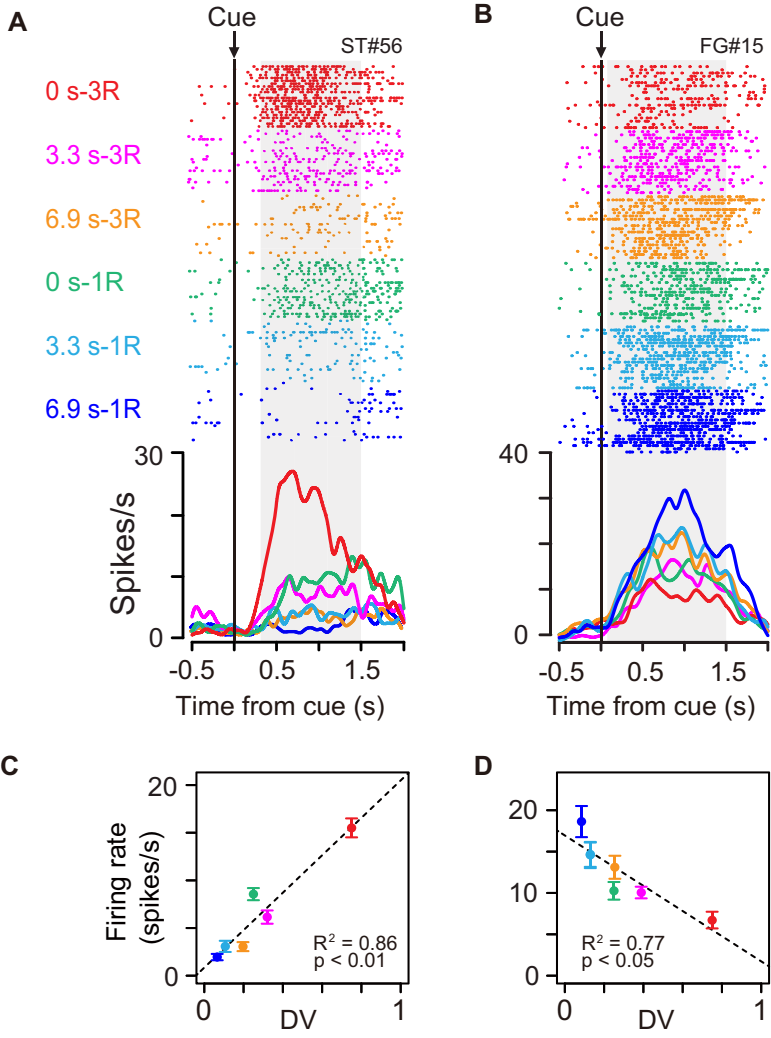
Figure 8—figure supplement 1. Comparison of learning in reward size and delayed reward task.

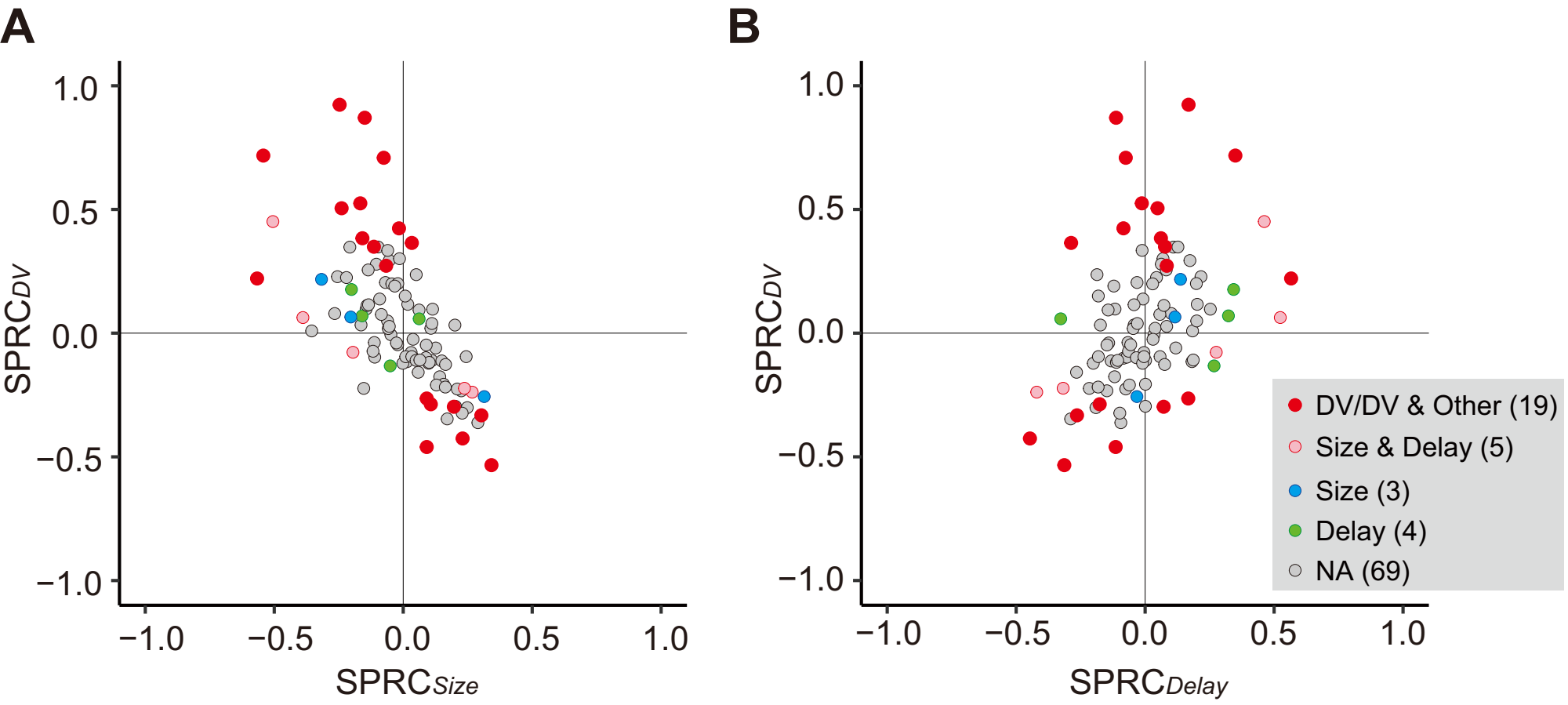
(A) Monkey RI was trained with reward-size task followed by delayed reward task. (top) Error rates as a function of session were plotted for both tasks. (bottom) Error rates as a function of reward size or delay duration during initial and stable phase were shown. (B) Monkey ST was trained with delayed reward task followed by reward-size task. Others are the same as A.

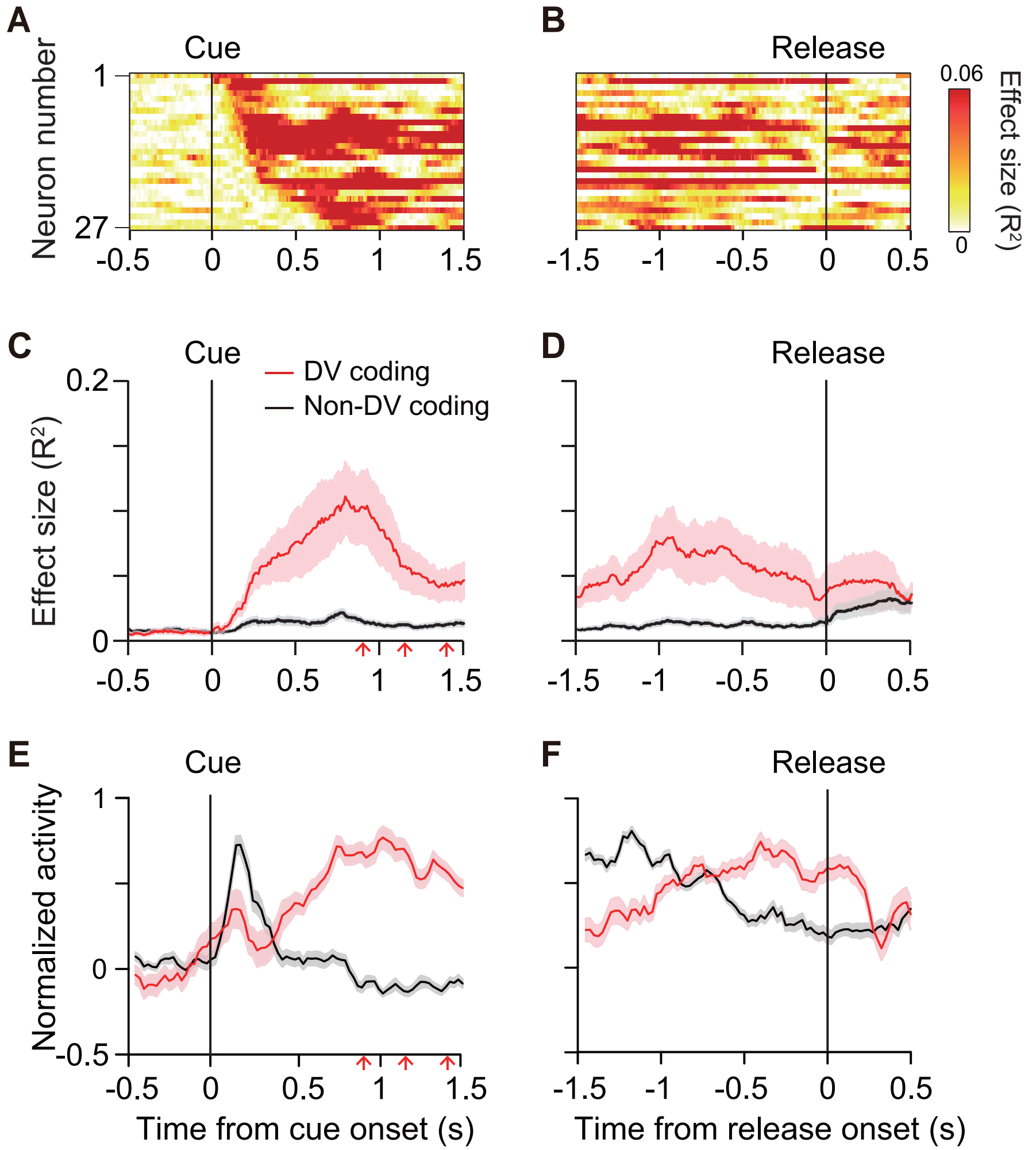
Figure 8—source data 1.

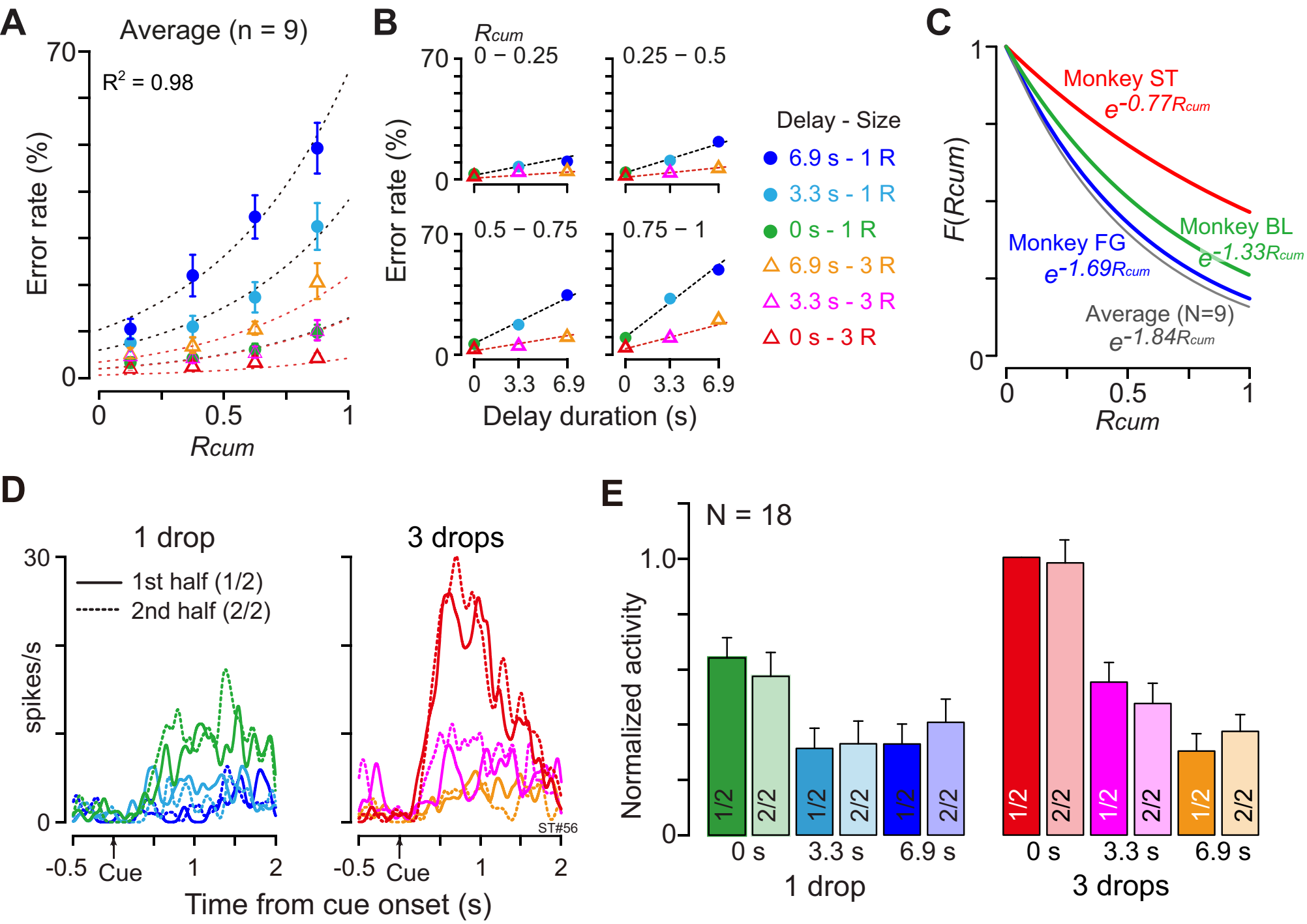

A
Bar touch
Time
CUE (0.4 s)
Wait (0.5-1.5 s)
GO
Release (< 1 s)
Correct
Delay
Reward
B
Cue
Delay duration
6.9 s
3.3 s
0 s
6.9 s
3.3 s
0 s
C
Monkey BI
Hyp: R2 = 0.96
Exp: R2 = 0.93
Rew = 1
Rew = 4
Error rate (%)
12
0
0
3.3
6.9
Delay duration (s)
Monkey FG
Hyp: R2 = 0.88
Exp: R2 = 0.84
Rew = 1
Rew = 3
40
0
0
3.3
6.9
Delay duration (s)
Monkey ST
Hyp: R2 = 0.94
Exp: R2 = 0.93
Rew = 1
Rew = 3
40
0
0
3.3
6.9
Delay duration (s)
D
Monkey BI
L
CD
Put
AC 0
Put
AC
AC+1
AC 0
AC−1
AC−2
AC−3
5 mm
Monkey ST
AC+3
AC+2
AC+1
AC 0
AC
AC−1
Put
AC−2

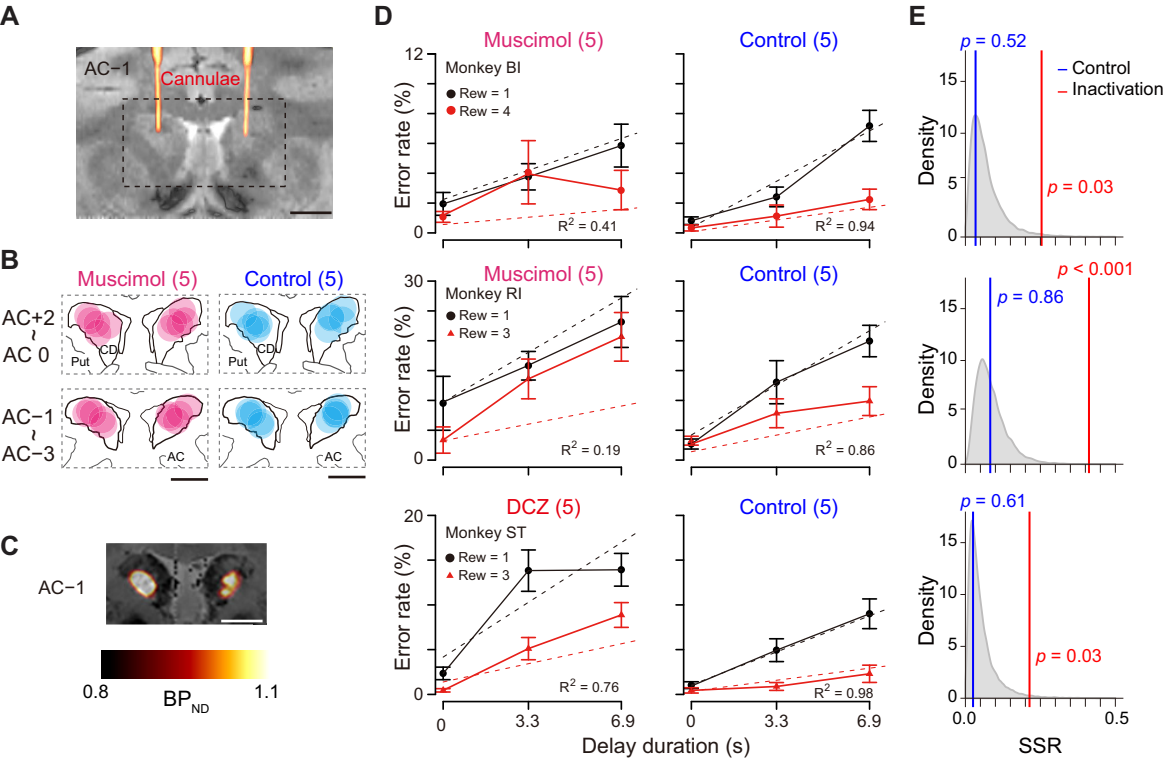

A
Cue
Release
Reward
10
spikes/s
0
-0.5
0.5
1.5
B
15
C
Time (s)
D
Cue (100)
45
20
15
20
Release
(49)
4
5
9
Reward
(49)

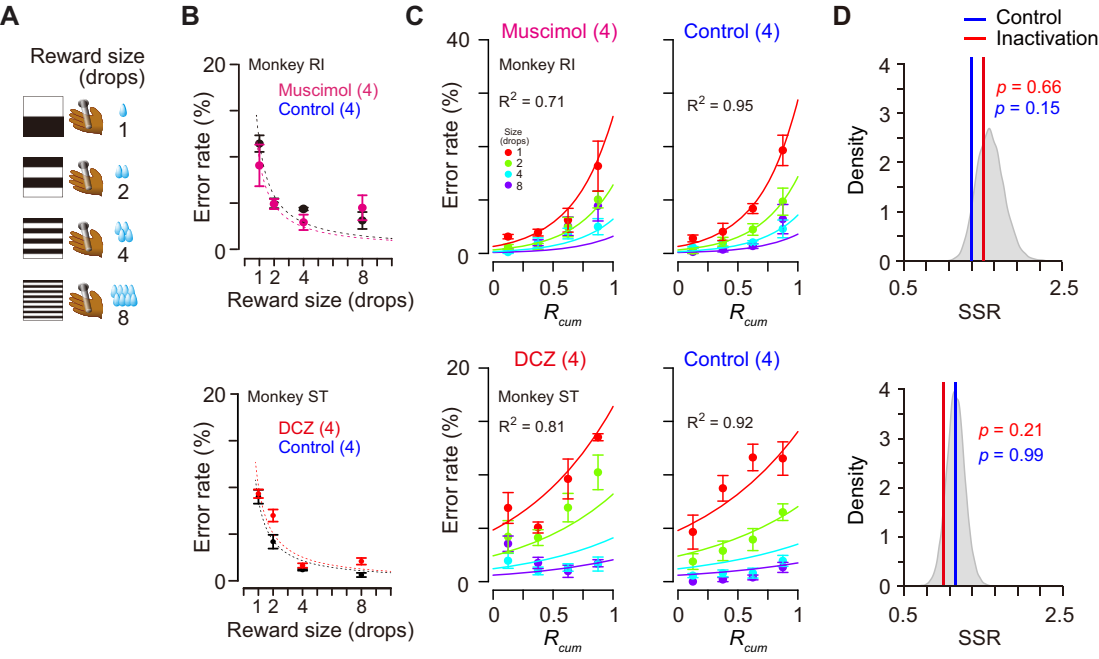

A
Cue
ST#56
0 s-3R
3.3 s-3R
6.9 s-3R
0 s-1R
3.3 s-1R
6.9 s-1R
30
Spikes/s
0
-0.5
0.5
1.5
Time from cue (s)
B
Cue
FG#15
40
0
-0.5
0.5
1.5
Time from cue (s)
C
20
Firing rate (spikes/s)
0
R² = 0.86
p < 0.01
0
DV
1
D
20
10
0
R² = 0.77
p < 0.05
0
DV
1

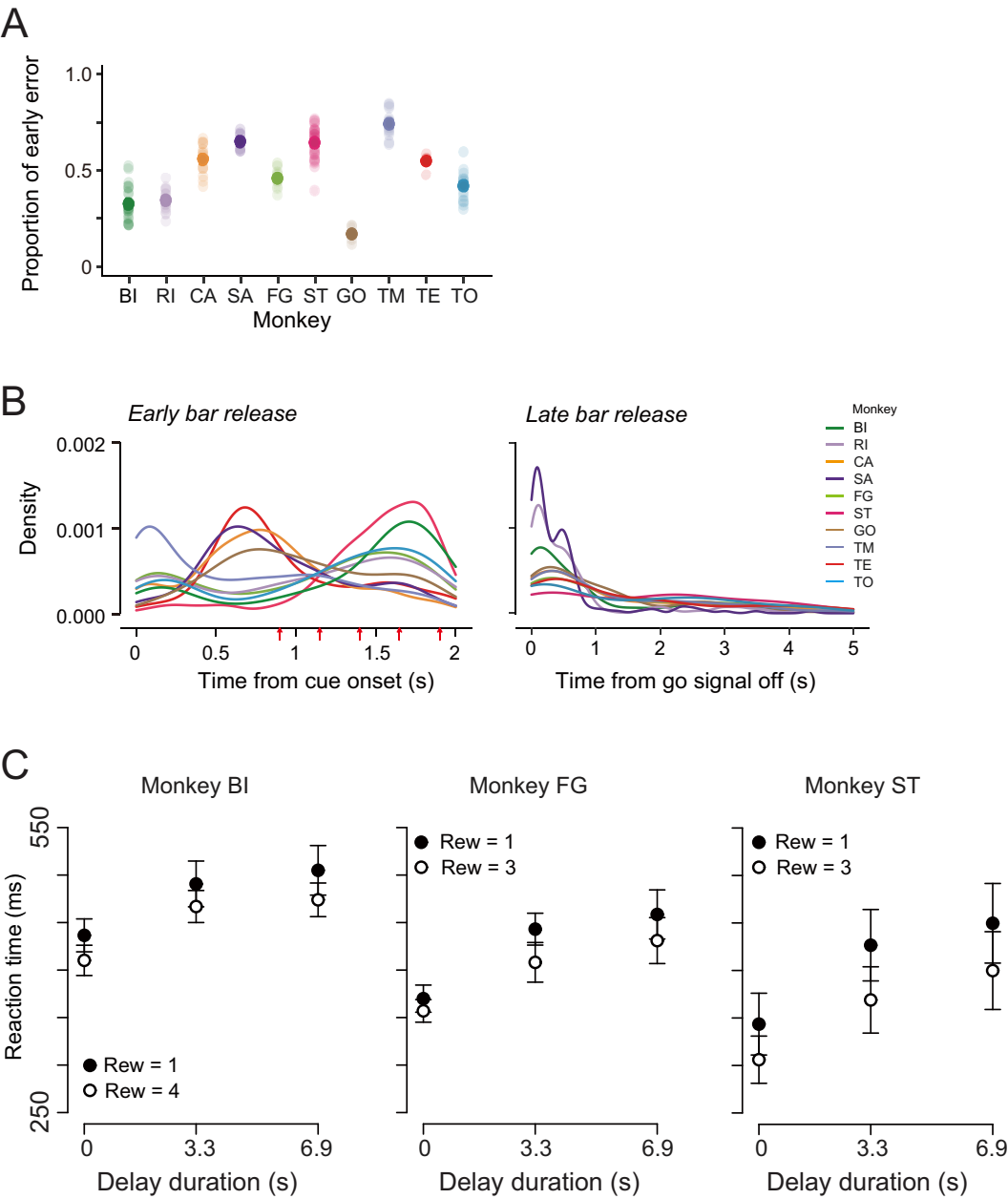

A
B
1.0
0.5
0.0
−0.5
−1.0
SPRC$_{DV}$
SPRC$_{Size}$
SPRC$_{Delay}$
DV/DV & Other (19)
Size & Delay (5)
Size (3)
Delay (4)
NA (69)

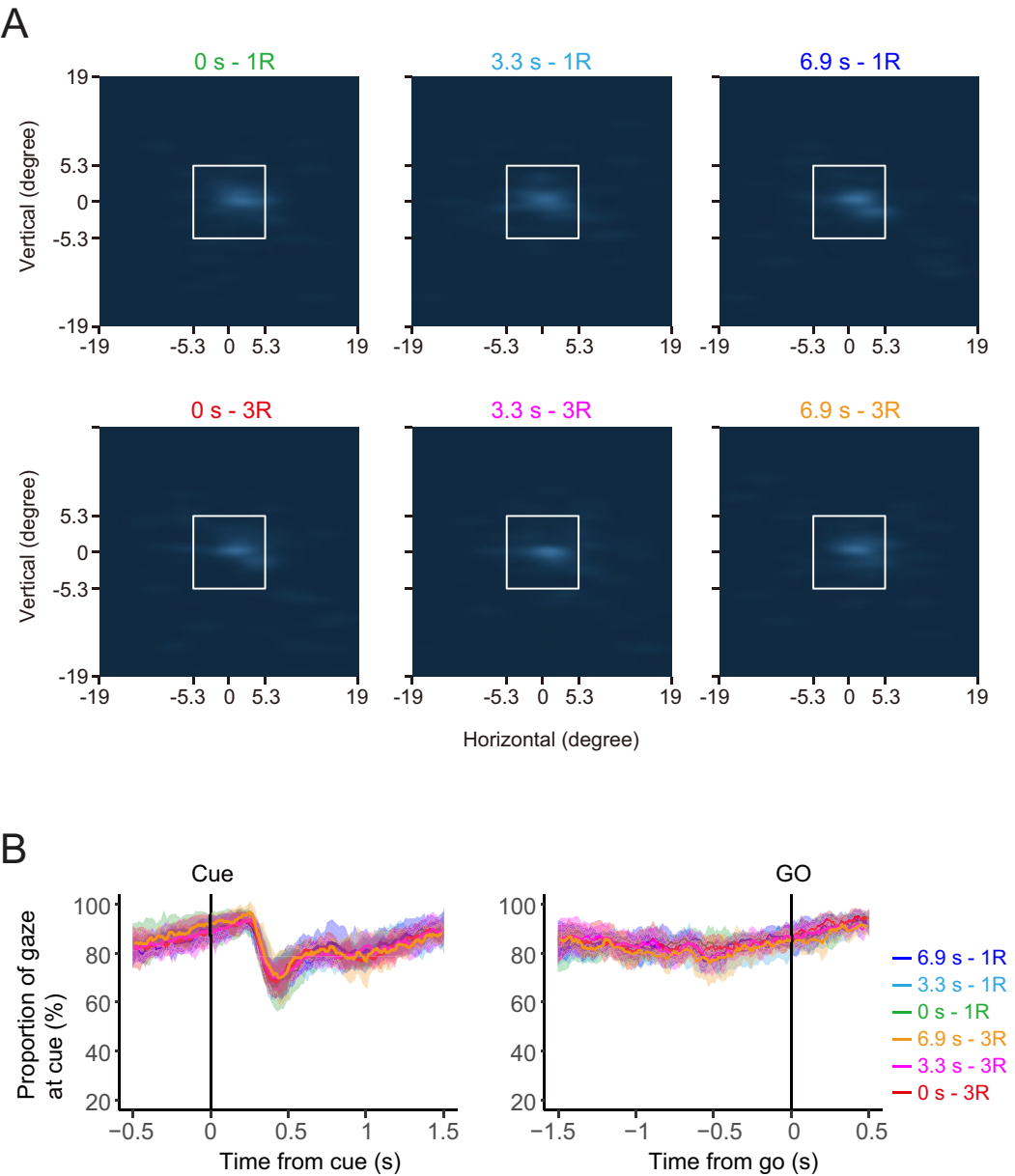

A
Cue
Neuron number
1
27
-0.5
0
0.5
1
1.5
B
Release
0.06
0
Effect size (R²)
-1.5
-1
-0.5
0
0.5
C
Cue
DV coding
Non-DV coding
Effect size (R²)
0.2
0
D
Release
E
Cue
Normalized activity
1
0
-0.5
Time from cue onset (s)
F
Release
Time from release onset (s)


A
Average (n = 9)
R² = 0.98
Error rate (%)
70
0
0
0.5
1
Rcum
B
Rcum
0 − 0.25
0.25 − 0.5
0.5 − 0.75
0.75 − 1
Error rate (%)
70
0
0 3.3 6.9
Delay duration (s)
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
C
F(Rcum)
1
0
Monkey ST
e-0.77Rcum
Monkey BL
e-1.33Rcum
Monkey FG
e-1.69Rcum
Average (N=9)
e-1.84Rcum
0
0.5
1
Rcum
D
1 drop
3 drops
30
0
spikes/s
1st half (1/2)
2nd half (2/2)
-0.5
Cue
1
2
ST#56
Time from cue onset (s)
E
N = 18
Normalized activity
1.0
0
1/2
2/2
0 s
3.3 s
6.9 s
1 drop
3 drops

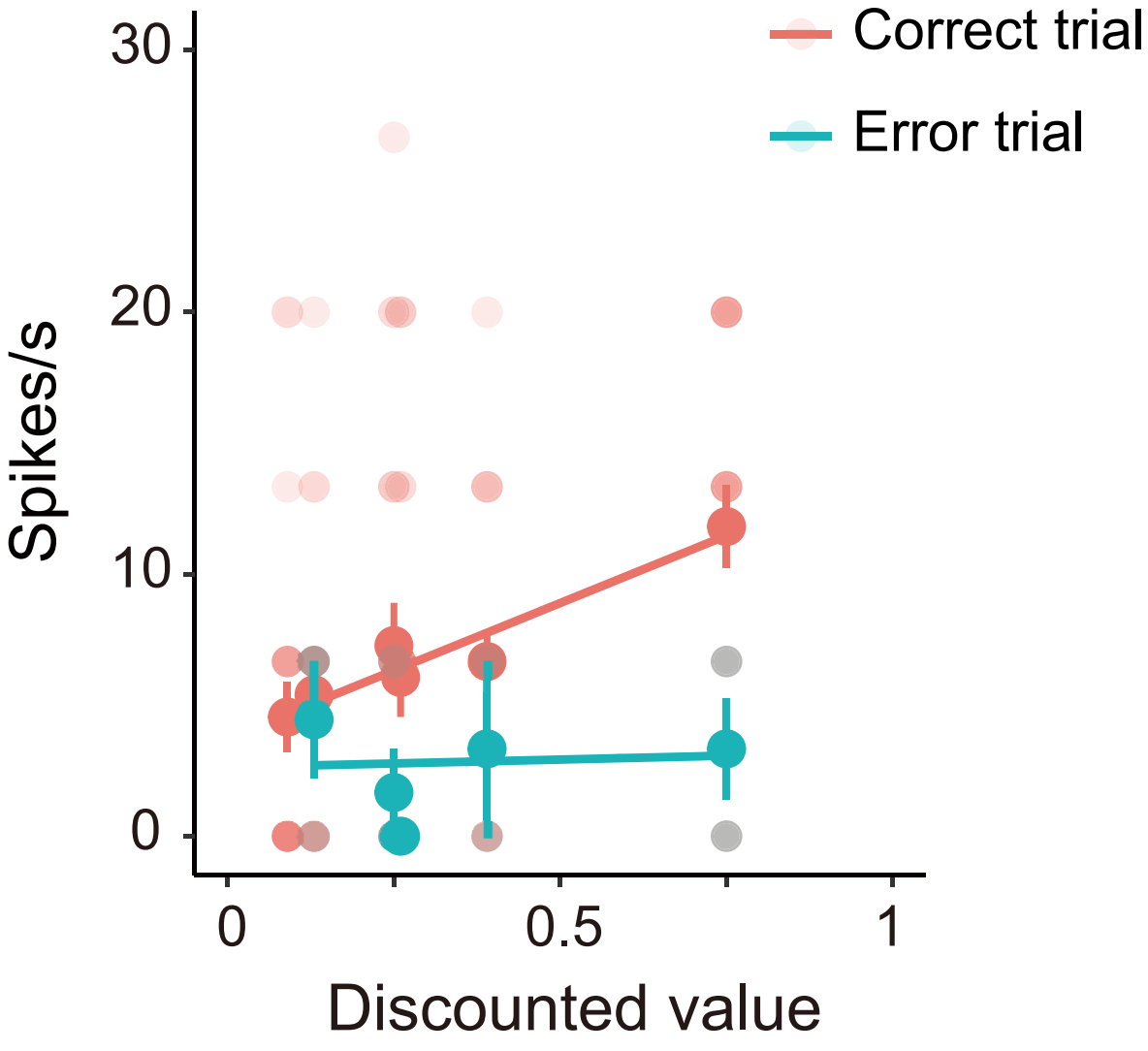

A
AC−1
Cannulae
B
Muscimol (5)
Control (5)
AC+2
~
AC 0
Put
CD
AC−1
~
AC−3
AC
C
AC−1
0.8
BP ND
1.1
D
Muscimol (5)
Monkey BI
Rew = 1
Rew = 4
Error rate (%)
R2 = 0.41
Control (5)
R2 = 0.94
Muscimol (5)
Monkey RI
Rew = 1
Rew = 3
R2 = 0.19
Control (5)
R2 = 0.86
DCZ (5)
Monkey ST
Rew = 1
Rew = 3
R2 = 0.76
Control (5)
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

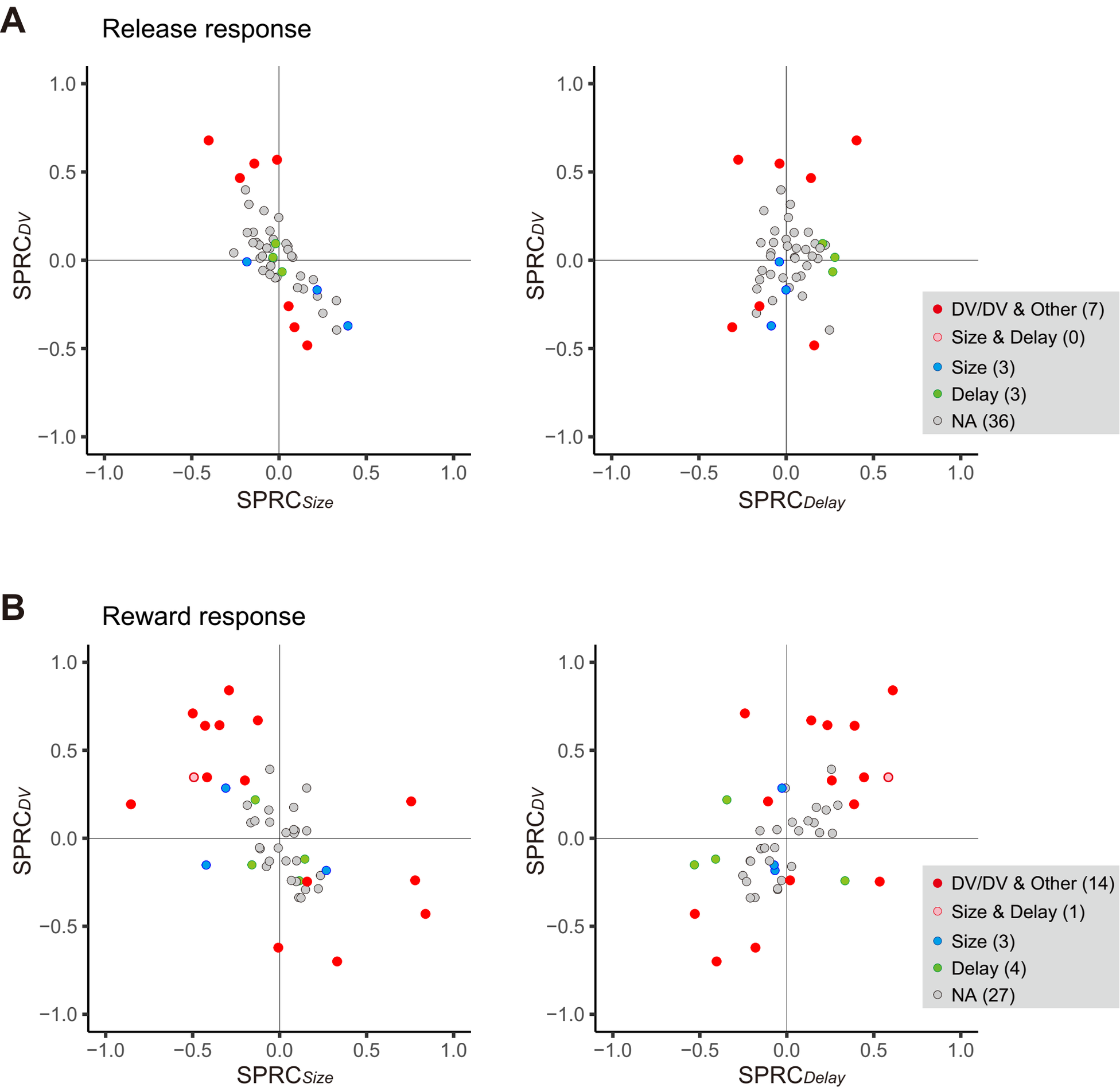

A
Reward size (drops)
1
2
4
8
B
Monkey RI
Muscimol (4)
Control (4)
Error rate (%)
Reward size (drops)
Monkey ST
DCZ (4)
Control (4)
C
Muscimol (4)
Monkey RI
R2 = 0.71
Control (4)
R2 = 0.95
Size (drops)
DCZ (4)
Monkey ST
R2 = 0.81
R2 = 0.92
Rcum
D
Control
Inactivation
p = 0.66
p = 0.15
p = 0.21
p = 0.99
Density
SSR

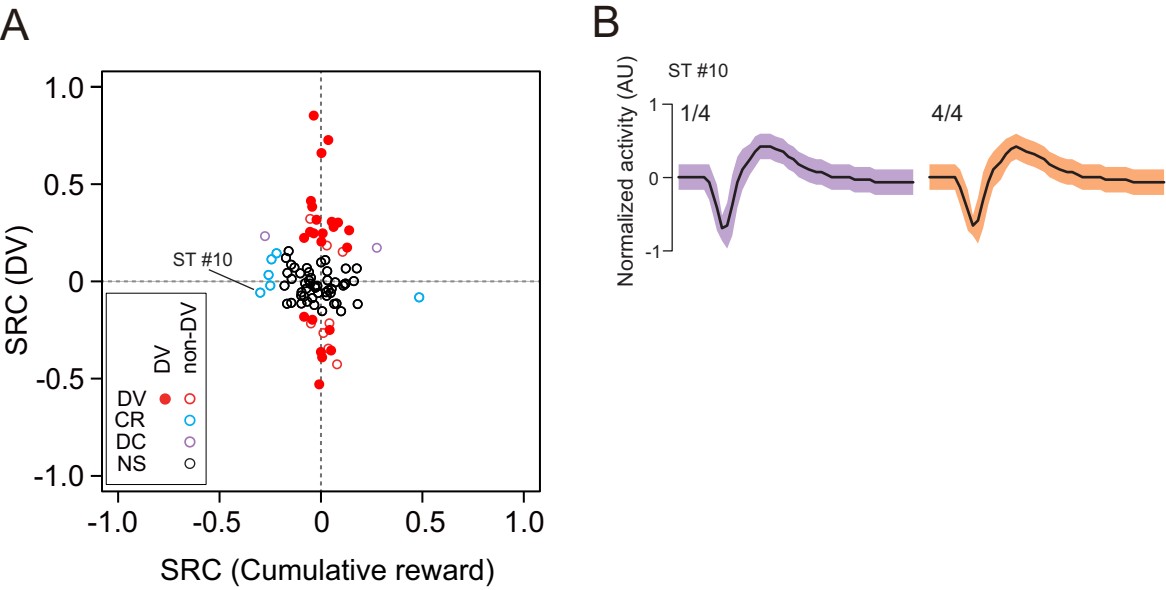

A
Proportion of early error
1.0
0.5
0
BI RI CA SA FG ST GO TM TE TO
Monkey
B
Early bar release
Late bar release
Density
0.002
0.001
0.000
0 0.5 1 1.5 2
Time from cue onset (s)
0 1 2 3 4 5
Time from go signal off (s)
Monkey
BI
RI
CA
SA
FG
ST
GO
TM
TE
TO
C
Monkey BI
Monkey FG
Monkey ST
Reaction time (ms)
550
250
Rew = 1
Rew = 4
Rew = 3
0 3.3 6.9
Delay duration (s)

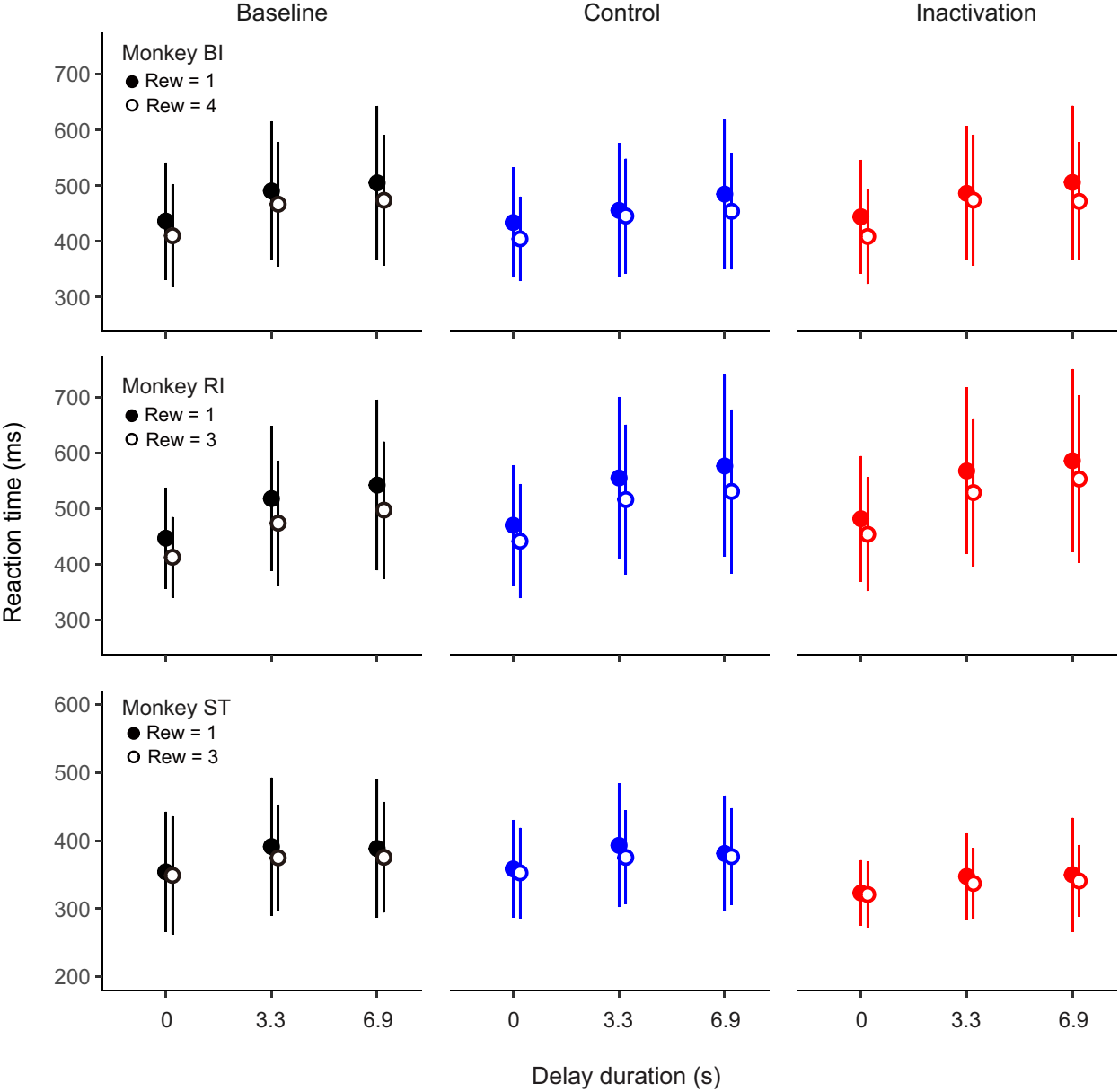

A
0 s - 1R
3.3 s - 1R
6.9 s - 1R
0 s - 3R
3.3 s - 3R
6.9 s - 3R
Vertical (degree)
Horizontal (degree)
19
5.3
0
-5.3
-19
Density
0.8
0.4
0.0
B
Cue
GO
Proportion of gaze at cue (%)
100
80
60
40
20
−0.5
0
0.5
1
1.5
−1.5
−1
Time from cue (s)
Time from go (s)
6.9 s - 1R
3.3 s - 1R
0 s - 1R
6.9 s - 3R
3.3 s - 3R
0 s - 3R

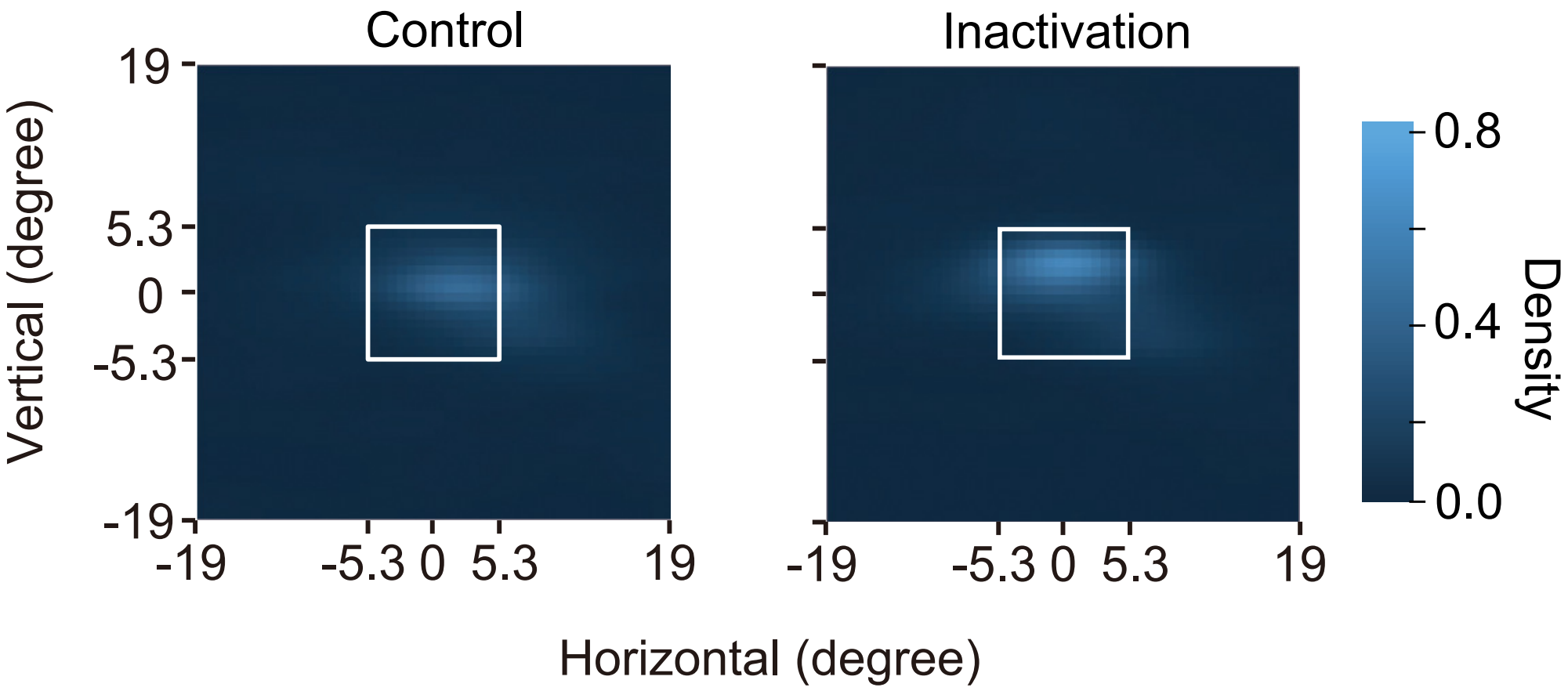
Model		Formula	n
Model #1		$fr \sim dv$	5
Model #2	mixed effect on both slope and intercept	$fr \sim dv + (dv \mid trial)$	1
Model #3	mixed effect on intercept	$fr \sim dv + (1 \mid trial)$	5
Model #4	mixed effect on slope	$fr \sim dv + (0 + dv \mid trial)$	11
Total			22

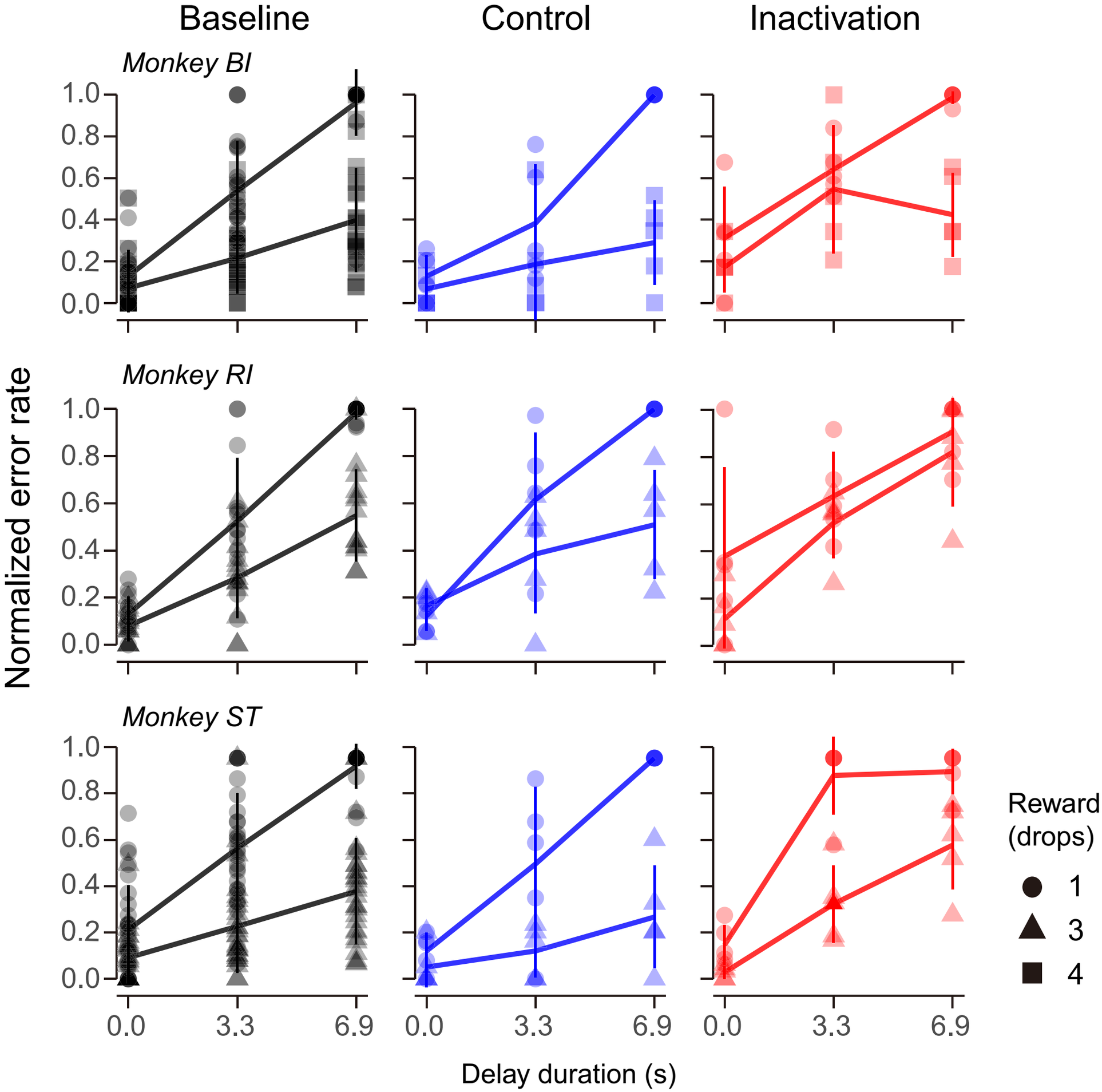

Correct trial
Error trial
Spikes/s
30
20
10
0
0
0.5
1
Discounted value

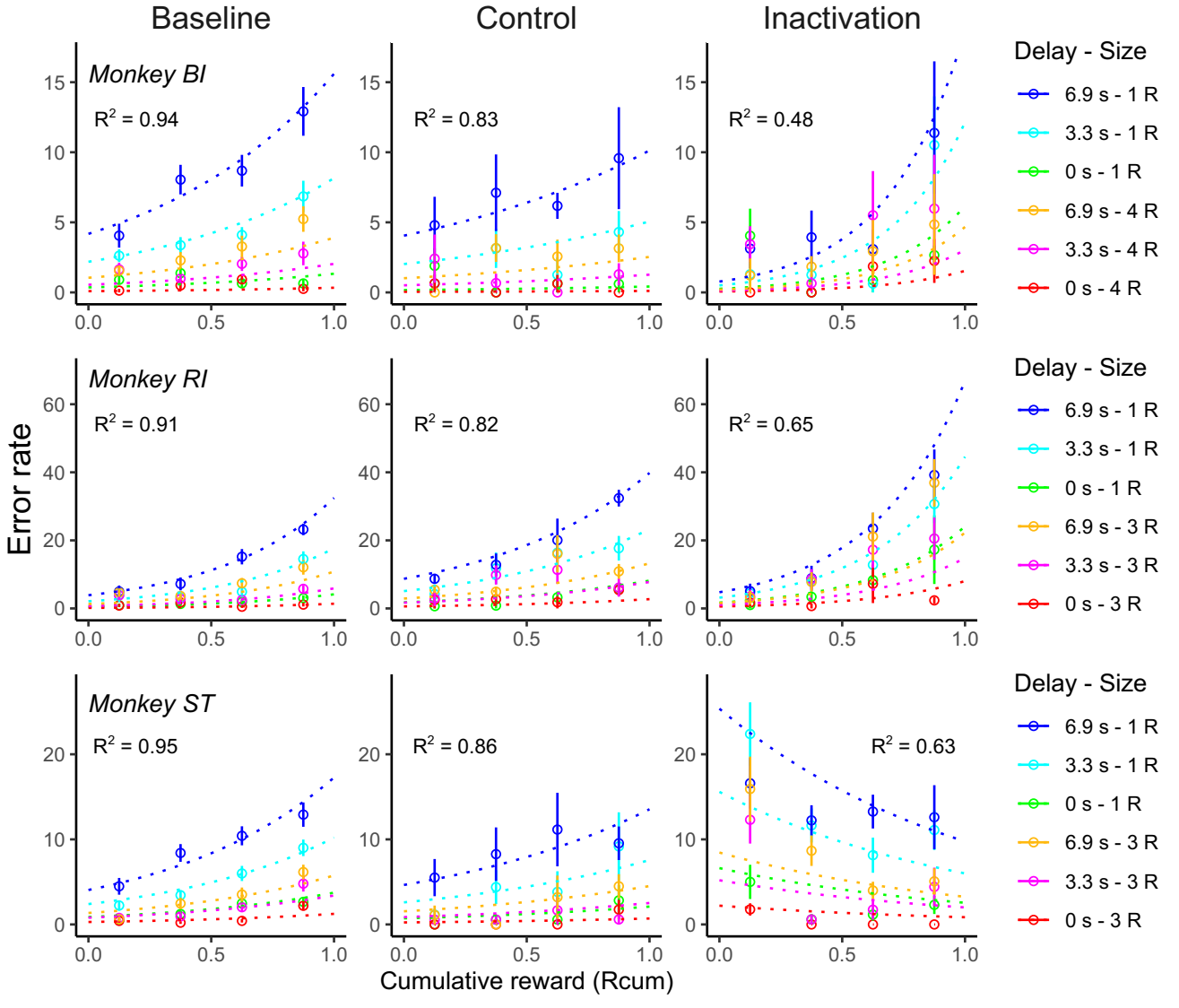

A
Release response
B
Reward response
SPRC_DV
SPRC_Size
SPRC_Delay
1.0
0.5
0.0
−0.5
−1.0
DV/DV & Other (7)
Size & Delay (0)
Size (3)
Delay (3)
NA (36)
DV/DV & Other (14)
Size & Delay (1)
Size (3)
Delay (4)
NA (27)

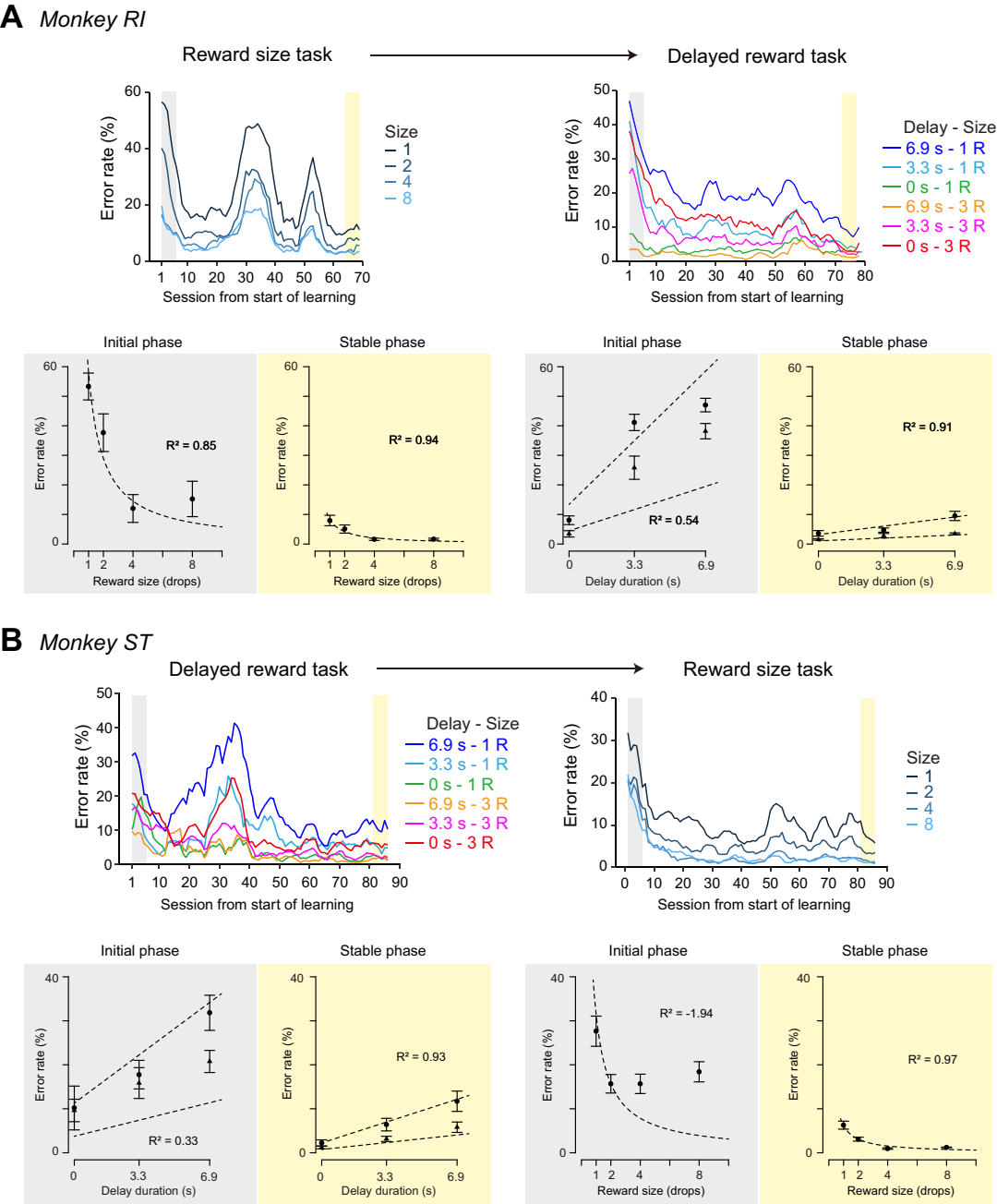

A
1.0
0.5
0
-0.5
-1.0
SRC (DV)
ST #10
DV
non-DV
DV
CR
DC
NS
-1.0
-0.5
0
0.5
1.0
SRC (Cumulative reward)
B
ST #10
Normalized activity (AU)
1
0
-1
1/4
4/4


Baseline
Control
Inactivation
Monkey BI
Rew = 1
Rew = 4
Monkey RI
Rew = 1
Rew = 3
Monkey ST
Rew = 1
Rew = 3
Reaction time (ms)
Delay duration (s)
0
3.3
6.9
200
300
400
500
600
700


Control
Inactivation
19
5.3
0
-5.3
-19
Vertical (degree)
-19
-5.3
0
5.3
19
Horizontal (degree)
0.8
0.4
0.0
Density


Baseline
Control
Inactivation
Monkey BI
Monkey RI
Monkey ST
Normalized error rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
3.3
6.9
Delay duration (s)
Reward (drops)
1
3
4


Baseline
Control
Inactivation
Monkey BI
Monkey RI
Monkey ST
R² = 0.94
R² = 0.83
R² = 0.48
R² = 0.91
R² = 0.82
R² = 0.65
R² = 0.95
R² = 0.86
R² = 0.63
Error rate
Cumulative reward (Rcum)
Delay - Size
6.9 s - 1 R
3.3 s - 1 R
0 s - 1 R
6.9 s - 4 R
3.3 s - 4 R
0 s - 4 R
6.9 s - 3 R
3.3 s - 3 R
0 s - 3 R


A
Monkey RI
Reward size task
Delayed reward task
Error rate (%)
Session from start of learning
Size
1
2
4
8
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
Initial phase
Stable phase
R² = 0.85
R² = 0.94
R² = 0.54
R² = 0.91
Reward size (drops)
Delay duration (s)
B
Monkey ST
Delayed reward task
Reward size task
R² = 0.33
R² = 0.93
R² = -1.94
R² = 0.97