Signaling incentive and drive in the primate ventral pallidum for motivational control of goal-directed action

Atsushi Fujimoto[1], Yukiko Hori[1], Yuji Nagai[1], Erika Kikuchi[1], Kei Oyama[1], Tetsuya Suhara[1] and Takafumi Minamimoto[1*]

[1] Department of Functional Brain Imaging, National Institute of Radiological Sciences, National Institutes for Quantum and Radiological Science and Technology, Chiba, Japan, 263-8555

Running Title: Central role of VP for goal-directed behavior

Corresponding Author: Takafumi Minamimoto, Ph.D.

Department of Functional Brain Imaging,

National Institute of Radiological Sciences,

National Institutes for Quantum and Radiological Science and Technology

4-9-1, Anagawa, Inage-ku, Chiba, Japan, 263-8555

E-mail : minamimoto.takafumi@qst.go.jp

Number of pages: 40 pages

Number of figures: 8 figures

Number of words for Abstract: 217 words

Number of words for Introduction: 553 words

Number of words for Discussion: 1217 words

Conflict of Interest

The authors declare no competing financial interests.

Acknowledgments

We thank J. Kamei, Y. Matsuda, R. Yamaguchi, Y. Sugii and R. Suma for their technical assistance, and Dr. I. Monosov for his invaluable technical advice and discussion. This work was supported by JSPS KAKENHI [JP15H05917] (to T. M.), [JP15H06872, JP17K13275] (to A. F.) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT), and by the Strategic Research Program for Brain Sciences from the Japan Agency for Medical Research and Development (AMED) JP18dm0107146 (to T. M.).

Author Contributions

A. F. and T. M. designed the research; A. F., Y. H., Y. N. and E. K. performed the research; A. F. analyzed the data; all authors wrote the manuscript.

Abstract

Processing incentive and drive is essential for control of goal-directed behavior. The limbic part of the basal ganglia has been emphasized in these processes, yet the exact neuronal mechanism has remained elusive. In this study, we examined the neuronal activity of the ventral pallidum (VP) and its upstream area, the rostromedial caudate (rmCD), while two male macaque monkeys performed an instrumental lever-release task, in which a visual cue indicated the forthcoming reward size. We found that the activity of some neurons in VP and rmCD reflected the expected reward-size transiently following the cue. Reward-size coding appeared earlier and stronger in VP than in rmCD. We also found that the activity in these areas was modulated by the satiation level of monkeys, which also occurred more frequently in VP than in rmCD. The information regarding reward-size and satiation-level was independently signaled in the neuronal populations of these areas. The data thus highlighted the neuronal coding of key variables for goal-directed behavior in VP. Furthermore, pharmacological inactivation of VP induced more severe deficit of goal-directed behavior than inactivation of rmCD, which was indicated by abnormal error repetition and diminished satiation effect on the performance. These results suggest that VP encodes incentive value and internal drive,

and plays a pivotal role in the control of motivation to promote goal-directed behavior.

Significance Statement

The limbic part of the basal ganglia has been emphasized in the motivational control of goal-directed action. Here, we investigated how the ventral pallidum (VP) and the rostromedial caudate (rmCD) encode incentive value and internal drive, and control goal-directed behavior. Neuronal recording and subsequent pharmacological inactivation revealed that the VP had stronger coding of reward size and satiation level than rmCD. Reward size and satiation level were independently encoded in the neuronal population of these areas. Furthermore, VP inactivation impaired goal-directed behavior more severely than rmCD inactivation. These results highlighted the central role of VP in the motivational control of goal-directed action.

Introduction

Motivational control over the purposeful action, or goal-directed behavior, is essential for gaining reward from an environment through knowledge about the association between action and its consequence (Dickinson, 1985; Dickinson and Balleine, 1994). Impairment of motivational control of goal-directed behavior promotes autonomic (habitual) control of actions, and this is evident in addictive disorders (Everitt and Robbins, 2005; Ersche et al., 2016). Goal-directed behavior is

73 regulated by two factors — the incentive value of the goal (reward) and the internal drive
74 (physiological state) of an agent (Berridge, 2004; Zhang et al., 2009). Accordingly, motivational
75 processes would be governed by the signals related to these two factors, although their neural
76 mechanism has remained largely unknown.

77 The ventral pallidum (VP), an output nucleus of ventral basal ganglia, is posited in the heart of the
78 limbic system (Haber et al., 1985; Groenewegen et al., 1993; Ray and Price, 1993; Mai and Paxinos,
79 2011) and has been strongly implicated in reward processing (Smith et al., 2009; Castro et al., 2015;
80 Root et al., 2015). Neuronal activity in the VP has been shown to reflect the incentive value of
81 reward cue in rodents (Tindell et al., 2004; Ahrens et al., 2016) and monkeys (Tachibana and
82 Hikosaka, 2012; Saga et al., 2017). Pharmacological manipulation of VP disrupted normal
83 reward-based behavior in monkeys (Tachibana and Hikosaka, 2012; Saga et al., 2017).
84 Dysregulation of neuronal activity induces addiction-like behavior in mice (Mahler et al., 2014;
85 Faget et al., 2018). Collectively, these results suggest a significant contribution of VP to
86 goal-directed behavior.

87 Other studies have also focused on the rostromedial part of the caudate nucleus (rmCD), one of
88 the upstream structures of the VP (Haber et al., 1990), as making a significant contribution to
89 goal-directed behavior. The rmCD receives projections from the lateral orbitofrontal cortex (OFC)
90 (Haber and Knutson, 2010; Averbeck et al., 2014), and attenuation of OFC-striatal activity promotes

habitual control of action over the goal-directed action in rodents (Yin et al., 2005; Gremel and Costa, 2013; Gremel et al., 2016). In monkeys, neuronal activity in the middle caudate including rmCD reflected reward size (Nakamura et al., 2012), and silencing of rmCD neurons induced a loss of reward-size sensitivity and disrupted goal-directed performance (Nagai et al., 2016). Taken together, these results raise the question of how rmCD and VP signal incentive and drive, and contribute to the dynamic control of goal-directed behavior.

In the present study, we aimed to elucidate the contribution of VP and rmCD to the control of goal-directed behavior. To address this, we analyzed the single-unit activities of these two areas while macaque monkeys performed an instrumental lever-release task, in which a visual cue indicated the forthcoming reward size (Minamimoto et al., 2009). As this task design permits us to infer the impact of incentive value (i.e., reward size) and internal drive (i.e., satiation level of monkeys) on performance, we assessed the neuronal correlate of the two factors, and compared neuronal coding between the two areas. With a population-level comparison, we found that the coding of reward size and satiation level in VP was greater than that in rmCD. Pharmacological inactivation of VP further examined the causal contribution of the neuronal activity to goal-directed action. Our results suggest a central role of VP in motivational control of goal-directed behavior and may provide implication for the neural mechanism of addictive disorders.

Materials and Methods

Subjects

Four male rhesus monkeys (*Macaca mulatta*, 5.7-7.2 kg) were used in this study. Two were used for neuronal recording (monkeys TA and AP), and the other two were used for local inactivation experiments (monkeys RI and BI). Monkey RI was also used in the previous rmCD inactivation study (Nagai et al., 2016). All surgical and experimental procedures were approved by the Animal Care and Use Committee of the National Institutes for Quantum and Radiological Science and Technology and were in accordance with the guidelines published in the NIH Guide for the Care and Use of Laboratory Animals.

Behavioral task

The monkeys squatted on a primate chair inside a dark, sound-attenuated, and electrically shielded room. A touch-sensitive lever 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 Jr et al., 1982). Presentation software was used to display visual stimuli (Neurobehavioral Systems Inc., Berkeley, CA).

All four monkeys were trained to perform the reward-size task (Minamimoto et al., 2009) (Fig. 1a). In each of the trials, the monkey had the same requirement to obtain one of four sizes of liquid

127 rewards (1, 2, 4, or 8 drops, 1 drop = ca. 0.1 mL). A trial began when a monkey gripped a lever. A
128 visual cue and a red spot appeared sequentially, with a 0.4 s interval, at the center of the monitor.
129 After a variable interval (0.5 -1.5 s), the central spot turned to green ('go' signal), and the monkey
130 had to release the lever within the reaction time window (0.2-1.0 s). If the monkey released the lever
131 correctly, the spot turned to blue (0.2-0.4 s), and then a reward was delivered. The next trial began
132 following an inter-trial interval (ITI, 1.5 s). When trials were performed incorrectly, they were
133 terminated immediately (all visual stimuli disappeared), and the next trial began with the same
134 reward condition following the ITI. There were two types of errors: premature lever releases (lever
135 releases before or no later than 0.2 s after the appearance of the go signal, named "early errors") and
136 failures to release the lever within 1.0 s after the appearance of the go signal (named "late errors").
137 The size of the reward was chosen randomly and was indicated by visual cues at the beginning of
138 each trial. Two sets of cues were used: a stripe set (for monkeys TA, RI, and BI) and an image set
139 (for monkey AP) (Fig. 1b). The monkeys used for electrophysiology (monkeys TA and AP) were
140 trained with different cue sets so that we could interpret the reward-related neuronal signal
141 irrespective of the visual features of the cue stimuli.
142 Prior to the experiment with the reward-size task, all monkeys had been trained to perform color
143 discrimination trials in a cued multi-trial reward schedule task for more than 3 months.
144

145 ***Surgery***

146 After behavioral training, a surgical procedure was carried out to implant one or two recording
147 chambers and a head fixation device under general isoflurane anesthesia (1-2%). The angles of the
148 chamber(s) were vertical (monkeys TA, AP, and BI) or 20° tilted from the vertical line (monkey RI)
149 in the coronal plane. Prior to surgery, overlay magnetic resonance (MR) and X-ray computed
150 tomography (CT) images were created using PMOD image analysis software (PMOD Technologies
151 Ltd, Zurich, Switzerland) to estimate the stereotaxic coordinates of the target brain structures. MR
152 images at 7T (Bruker Corp., Billerica, MA) and CT images (3D Accuitomo170: J. Morita Corp.,
153 Osaka, Japan) were obtained under anesthesia (propofol 0.2-0.6 mg/kg/min, i.v.).

154

155 ***Neuronal recordings***

156 Single-unit activity was recorded from monkeys TA and AP while they performed the reward-size
157 task. We analyzed all successfully isolated activities and held at least 10 trials for each reward
158 condition. Action potentials of single neurons were recorded from VP and rmCD using a
159 glass-coated 1.0 MΩ tungsten microelectrode (Alpha Omega Engineering Ltd., Nazareth, Israel). A
160 guide tube was inserted through the grid hole in the implanted recording chamber into the brain, and
161 the electrodes were advanced through the guide tube by means of a micromanipulator (MO-97A:
162 Narishige Co., Ltd., Tokyo, Japan). Spike sorting to isolate single neuron discharges was performed

with a time-window algorithm (TDT-RZ2: Tucker Davis Technologies Inc., Alachua, FL). The timing of action potentials was recorded together with all task events at millisecond precision.

For the VP recordings, we targeted the region just below the anterior commissure (AC) in the +0-1 mm coronal plane (Fig. 2a). VP neuron was characterized by high spontaneous firing rate with phasic discharge to the task events (Tachibana and Hikosaka, 2012). For the rmCD recordings, we targeted the area within 2-4 mm laterally and 2-7 mm ventrally from the medial and upper edge of the caudate nucleus, in the +4-5 mm coronal plane (Fig. 2b). The rmCD neurons were classified into three subtypes based on the electrophysiological criteria (Aosaki et al., 1995; Yamada et al., 2016). The presumed medium-spiny projection neurons (PANs: phasically-active neurons) were characterized by low spontaneous firing and phasic discharge to the task events, while the presumed cholinergic interneurons (TANs: tonically-active neurons) were characterized by broad spike width (valley-to-peak width) and tonic firing around 3.0-8.0 Hz. The presumed parvalbumin-containing GABAergic interneurons (FSNs: fast-spiking neurons) was characterized by narrow spike width and relatively higher spontaneous firing than other types of caudate neurons. A spike-width analysis was performed using the Off-line sorter (Plexon, Dallas, TX).

To reconstruct the recording location, electrodes were visualized using CT scans after each recording session, and the positions of the tip were mapped onto the MR image using PMOD.

181 ***Muscimol microinjection***

182 To achieve neuronal silencing, $GABA_A$ agonist muscimol (Sigma-Aldrich Co., St. Louis, MO) was
183 injected bilaterally into the VP (monkeys RI and BI) using the same procedures as reported
184 previously (Nagai et al., 2016). We used two stainless steel injection cannulae inserted into the
185 caudate (O.D. 300 µm: Muromachi-Kikai Co. Ltd., Tokyo, Japan), one in each hemisphere. Each
186 cannula was connected to a 10-µL microsyringe (#7105KH: Hamilton Company, Reno, NV) via
187 polyethylene tubing. These cannulae were advanced through the guide tube by means of an oil-drive
188 micromanipulator. Muscimol (3 µg/1 µL saline) was injected at a rate of 0.2 µL/min by auto-injector
189 (Legato210: KD Scientific Inc., Holliston, MA) for a total volume of 2 µL in each side. The
190 behavioral session (100 min) was started soon after the injection was finished. We performed at most
191 one inactivation study per week. For a control, we performed sham experiments at other times, in
192 which the time-course and mechanical settings were set identical to the muscimol session. At the end
193 of each session, a CT image was obtained to visualize the injection cannulae in relation to the
194 chambers and skull. The CT image was overlaid on an MR image by using PMOD to assist in
195 identifying the injections sites.

196

197 ***Experimental design and statistical analysis***

198 All statistical analyses were performed with R Statistical Package. For the behavioral analysis, the

199 data obtained from the two monkeys were analyzed. The dependent variables of interest were the
200 error rate, reaction time (RT), and lever grip time. The error rate was calculated by dividing the total
201 number of errors (early and later errors) by the total number of trials. RT was defined as the duration
202 from a 'go' signal to the time point of lever release in a correct trial. Average error rate and RT were
203 computed for each reward condition. The lever-grip time was defined as the duration from the end of
204 ITI to the time when the monkey gripped the lever to initiate a trial (i.e., the latency to start a trial).
205 Error rate and RT were analyzed using two-way repeated-measures ANOVAs with reward size (1, 2,
206 4, 8 drops) and satiation level (proportion of cumulative reward in a session: 0.125, 0.375, 0.625,
207 0.875) as within-subjects factors. The lever-grip time was analyzed using one-way repeated
208 measures ANOVAs with satiation level (cumulative reward: 0.125, 0.375, 0.625, 0.875) as a
209 within-subjects factor. The proportional behavioral data were transformed using the variance
210 stabilizing arcsine transformation before hypothesis testing (Zar, 2013). The error rate was also
211 analyzed by a model fitting as described previously (Minamimoto et al., 2009; Minamimoto et al.,
212 2012). To assess the effects of reward size and satiation level on the error rate, following model was
213 used:

214 $$E = c/(R \times f(R_{cum})) \qquad (1),$$

215 where E and R denote error rate and reward size, respectively. Parameter c is a monkey-specific
216 parameter that represents reward-size sensitivity. $f(R_{cum})$ denotes the reward discounting function,

which was modeled as follows:

$$f(R_{cum}) = e^{-\lambda \times R_{cum}} \quad (2),$$

where R_{cum} is a normalized cumulative reward in a session (0-1), and λ is a monkey-specific parameter that represents the steepness of reward discounting.

For neuronal data analysis, three task periods (cue period: 100-700 ms after cue on, pre-release period: 0-300 ms before lever release, reward period: 0-300 ms after reward delivery) and the baseline period (ITI: 0-500 ms before cue on) were defined. A neuron was classified as reward-size coding neuron when the firing rate during the task period was significantly modulated by reward size (main effect of reward size $p < 0.05$, one-way ANOVA) and linearly reflected reward size ($p < 0.05$, linear regression analysis). Neurons that showed positive or negative correlation were classified as positive or negative reward-size coding neurons, respectively.

To quantify the time course of reward-size coding, the effect size (R squared) in a linear regression analysis with reward-size was calculated for every 100-ms window shift in 10-ms steps. Coding latency was defined as the duration between the cue onset and the time at which the first of three following consecutive 100-ms test intervals showed a significant reward-size effect ($p < 0.05$). Peak effect size was defined as the maximum effect size of individual neurons. Average coding latency and peak effect size were compared between VP and rmCD by Wilcoxon rank-sum test.

The effect of reward size and satiation level on firing rate during the task periods was also

assessed by the following multiple linear regression model:

$$Y = \beta_1 \times R + \beta_2 \times S + r \quad (3),$$

where Y is the firing rate, R is the reward size, S is the satiation level, β_1 and β_2 are the regression coefficients, and r is a constant. Satiation level was inferred using equation (2) with the individual parameter λ derived from behavioral analysis. Neurons were classified into reward-size and satiation-level coding neurons if they had a significant correlation coefficient ($p < 0.05$) with each variable. A neuron was classified as a motivational-value coding neuron, when a neuron had a significant positive reward-size coefficient and negative satiation-level coefficient, or vice versa. The proportion of each type of neuron in the neuronal population (VP and rmCD) was calculated for each of the task periods, and compared between VP and rmCD using Wilcoxon rank-sum test with a threshold of statistical significance set by Bonferroni correction (alpha = 0.05/4). The proportion of motivational-value coding neurons was further compared to that of pseudo-motivational-value coding neurons, which were calculated by multiplying the proportion of neurons that coded satiation level and reward size orthogonally; this calculated the dual coding of two items of information by chance (i.e., joint probability).

For behavioral analysis of muscimol microinjection effects, the data obtained from two monkeys (monkeys RI and BI) were used for VP inactivation, while the data obtained from two monkeys (monkeys RI and RO) for rmCD inactivation (Nagai et al., 2016) were used for comparison. The

dependent variables of interest were the change of error rate and early-error rate from the control session. The early-error rate was calculated by dividing the number of early error trials by that of total error trials (i.e., the sum of early and late error trials). RTs and lever-grip time in the first and latter halves of sessions were also compared to assess the effects of satiation in control and muscimol sessions. Because VP inactivation induced similar effects in the two monkeys (monkeys BI and RI) regarding the elevation of error rate, the data were pooled across monkeys and compared to rmCD inactivation by Wilcoxon rank-sum test.

Results

Behavioral performance reflected reward size and satiation level

Two monkeys (TA and AP) learned to perform the reward-size task, in which unique visual cues provided information of the upcoming reward size (1, 2, 4, 8 drops of liquid; Fig. 1a and b). For both monkeys, error rate reflected reward size, such that the monkeys made more error responses (premature release or too-late release of the lever) when small rewards were assigned (Fig 1c). Error rate also reflected the satiation level, such that the error rate increased according to reward accumulation (Fig. 1d). Two-way repeated-measures ANOVAs (reward size: 1, 2, 4, 8 drops × cumulative reward: 0.125, 0.375, 0.625, 0.875) confirmed the significant effects of reward prediction and satiation on the behavioral performance (main effects of reward size and cumulative reward, and

271 their interaction, $p < 0.01$, $F > 11$). As reported previously, the error rates were explained by a model
272 in which the expected reward size was multiplied by an exponential decay function according to
273 reward accumulation (Minamimoto et al., 2009; Minamimoto et al., 2012) (Fig. 1d, see Materials
274 and Methods).

275 The reaction time (RT) also changed in association with reward size and satiation level, such that
276 the monkeys had a slower reaction for a smaller reward and when they had accumulated large
277 amounts of rewards. Two-way repeated-measures ANOVAs revealed significant main effects of
278 reward size and cumulative reward, and their interaction on RT (all $p < 0.01$, $F > 3.5$). The lever-grip
279 time also changed according to the satiation level, such that the monkey tended to slowly grip the
280 lever to initiate a trial in the later stage of a session. One-way repeated-measures ANOVAs revealed
281 significant main effect of cumulative reward ($p < 0.01$, $F = 102$). Together, these results suggest that
282 the monkeys adjusted their motivation of the action based on the incentive value (i.e., expected
283 reward size) and the internal drive (i.e., current satiation level).

284

285 **Task-related activity of VP and rmCD neurons**

286 While monkeys TA and AP performed the reward-size task, we recorded the activity of 102 neurons
287 (50 from TA and 52 from AP) in VP (Fig. 2a) and 106 neurons (68 from TA and 38 from AP) in
288 rmCD (Fig. 2b). rmCD neurons were further classified into phasically-active neurons (PANs, n = 56),

tonically-active neurons (TANs, n = 44) and fast-spiking neurons (FSNs, n = 6), based on the criteria that were established in earlier studies (Aosaki et al., 1995; Yamada et al., 2016) (see Materials and Methods). The baseline firing rate (0-500 ms before cue, mean ± SEM) was high in the VP (34.8 ± 1.8 spk/s) and low in the rmCD (PANs 3.6 ± 0.4 spk/s, TANs 7.3 ± 0.4 spk/s, FSNs 14.3 ± 1.6 spk/s). Because the characteristics of neuronal activity among the three subtypes in rmCD were not significantly different in terms of value coding (i.e., reward-size and satiation-level coding) ($p > 0.53$, two-sample Kolmogorov-Smirnov test), we decided to treat them as a single population for the subsequent analyses. We will also report the results from PANs to ensure that the same conclusions would be reached.

Neuronal activity in VP and rmCD reflected reward-size

We first examined how the incentive value is represented in VP and rmCD neurons. Fig. 3a-d illustrates four examples of neuronal activity showing reward-size modulation during the cue period. The first VP neuron example increased its activity after the largest reward (8 drops) cue, but decreased after smaller (1, 2, 4 drops) ones (Fig. 3a). The firing rate was positively correlated with the reward size ($p < 0.01$, $r = 0.49$, linear regression analysis, Fig. 3a, right), and therefore this neuron exhibited positive reward-size coding in this period. In another VP neuron example, the firing rate during the cue period became lower as a larger reward was expected ($p < 0.01$, $r = -0.52$, linear

regression analysis), and thus this neuron exhibited negative reward-size coding (Fig. 3b). Similarly, we found that some rmCD neurons linearly encoded reward size during the cue period (Fig. 3c and d).

VP had stronger reward-size coding than rmCD

We found significant linear reward-size modulation on the activity of at least one task period of the majority of VP neurons (68/102) ($p < 0.05$, linear regression analysis). The proportion was significantly larger than the proportion of reward-size coding neurons in rmCD (46/106; $p < 0.01$, $\chi^2 = 10$, chi-square test). In rmCD, a large part of reward-size coding was observed in PANs (31/46). Reward-size coding was mainly observed during the cue period in both areas (VP 63/68, rmCD 40/46). On population activity, both VP and rmCD clearly showed both types of linear reward-size coding during the cue period (Fig. 3e-g). During release or reward periods, however, linear reward-size coding was less clear.

To quantify reward-size coding, we computed the effect size (R squared) of activity in the sliding window (100 ms bin, 10 ms step) for each of the recorded neurons (Fig. 4a and b). Fig. 4 c and d show the average effect size of positive coding neurons (top) and negative coding neurons (bottom) in VP and rmCD, respectively. In both areas, the effect size rapidly and transiently increased after cue presentation, for both positive and negative coding neurons (Fig. 4c and d). The peak effect size

of VP neurons (0.23 ± 0.016, median ± SEM) was significantly larger than that of rmCD neurons (0.14 ± 0.018) ($p < 0.01$, df = 101, rank-sum test) and that of PANs (0.14 ± 0.027, n = 26, $p = 0.018$, df = 87). Taken together, both VP and rmCD neurons exhibited reward-size modulation mainly after cue presentation, in which the former showed a stronger modulation in terms of the proportion of neuron and effect size.

Reward-size coding emerged earlier in VP than in rmCD

We compared the time course of reward-size coding after cue in two populations (VP, n = 63; rmCD, n = 40). The latency of reward-size coding of VP neurons was significantly shorter than that of rmCD neurons (VP 115 ± 17 ms, rmCD 225 ± 20 ms, median ± SEM; $p < 0.01$, df = 101, rank-sum test). Positive coding occurred earlier in VP than in rmCD (VP 100 ± 19 ms, rmCD 250 ± 27 ms, $p < 0.01$, df = 66, Fig. 5a and b), whereas the difference did not reach significance for negative coding (VP 150 ± 37 ms, rmCD 200 ± 29 ms, $p = 0.48$, df = 33, Fig. 5c and d).

If rmCD is the primary source for providing reward information to VP, the projection neurons (i.e., PANs) in rmCD would encode reward size earlier than VP neurons. However, the latency of reward-size coding of VP neurons was again shorter than that of PANs (VP 115 ± 17 ms, PANs 190 ± 25 ms, $p = 0.049$, df = 87). These results suggest that rmCD cannot be the primary source for the reward-size coding in VP.

Encoding of satiation level in VP and rmCD

As shown above, the monkeys' goal-directed behavior (i.e., error rate) is affected by internal drive (i.e., satiation change) as well as incentive value (Fig. 1d). We found that some neurons changed their firing rate according to the satiation level. For instance, a VP neuron decreased its activity after the cue with reward-size (negative reward-size coding), while the decrease became smaller according to reward accumulation (Fig. 6a-c). This was not due to changes in isolation during a recording session as confirmed by unchanged spike waveforms (Fig. 6d). In another example rmCD neuron, the firing rate in the cue period was positively related to reward size and negatively related to reward accumulation (Fig. 6e-h).

To assess how satiation level and reward size were encoded in VP and rmCD, we performed a multiple linear regression analysis on the activity during each of four task periods (ITI, cue, pre-release, reward). For this analysis, the satiation level was inferred using the model with monkey-specific parameter λ that explained individual behavioral data (Fig. 1d, see Materials and Methods). Fig. 7a shows scatter plots of standardized regression coefficient for satiation level and reward size for the activity during the cue period of individual neurons in VP and rmCD. The proportion of satiation-level coding neurons during the cue period was not significantly different between the two areas (Fig. 7b left; $p = 0.28$, $\chi^2 = 1.2$, chi-square test). In the other task periods,

however, this proportion was significantly larger in VP than in rmCD ($p < 0.05$ with Bonferroni correction, rank-sum test) (Fig. 7b left). In contrast, the proportion of reward-size coding neurons in VP was larger than that in rmCD for the cue period ($p < 0.01$, $\chi^2 = 11$), but not for the other task periods ($p > 0.05$ with Bonferroni correction) (Fig. 7b center). A similar tendency was found when we compared VP and PANs; satiation-level coding was more frequent in VP than in PANs during the ITI, pre-release, and reward periods ($p < 0.05$), while reward-size coding tended to be more frequent in VP than in PANs during the cue period ($p = 0.10$, $\chi^2 = 2.7$). These results suggest that satiation level and reward size were encoded in different time courses throughout the trial in VP and rmCD, and both were strongly signaled in VP.

By definition, motivational value increases as expected reward size increases, and as satiation level decreases (Berridge, 2004; Zhang et al., 2009). Thus, motivational value coding neuron is defined as a neuron that showed positive reward-size coding and negative satiation-level coding, or vice versa (Fig. 7a, yellow areas; see Fig. 6 for examples). We found some motivational-value coding neurons during cue, pre-release and reward periods in both areas (Fig. 7b right). However, the proportion was not significantly larger than that of neurons by chance coding both satiation level and reward size in the opposite direction (all $P > 0.05$, chi-square test). This result suggests that reward size and satiation level were not systematically integrated, but were independently signaled in VP and rmCD.

379

380 **Inactivation of VP disrupted goal-directed behavior**

381 Previous study demonstrated that bilateral inactivation of rmCD by local injection of the $GABA_A$
382 receptor agonist muscimol diminished reward-size sensitivity (Nagai et al., 2016) as indicated by an
383 increase in the error rate of reward-size task especially in larger reward-size trials (Fig. 8a),
384 supporting that the neuronal activity in rmCD is essential for controlling goal-directed behavior
385 based on the expected reward size.

386 To examine how VP contributes to the control of goal-directed behavior, we injected muscimol
387 into bilateral VP (3 µg/µL, 2 µL/side) of two monkeys (monkeys BI and RI) and tested them with
388 reward-size task. CT images visualizing the injection cannulae confirmed the sites of muscimol
389 injection; they were located in the VP matching the recording sites (Fig. 8b; see Fig. 2a for
390 comparison). Bilateral VP inactivation significantly increased the error rate regardless of reward size
391 (main effect of treatment, $p < 0.01$, $F_{(1, 17)} = 449.7$, repeated measures ANOVA, Fig. 8c). The
392 increase in error rate by VP inactivation was significantly greater than that by rmCD inactivation
393 (data pooled across monkeys, $p < 0.01$, df = 18, rank-sum test). After VP inactivation, the monkeys
394 frequently released the lever before the go signal appeared, or even before red or cue came on, and
395 thereby increased the proportion of early errors ($p < 0.01$, df = 16, rank-sum test; Fig. 8d left) and
396 error repetition ($p < 0.01$, df = 16, Fig. 8d right). The error pattern changes were not observed after

397 rmCD inactivation ($p > 0.47$, $df = 22$, Fig. 8d). Because VP inactivation did not decrease, but rather
398 increased the total number of trials performed (control 922 ± 132 trials; VP inactivation 1464 ± 153
399 trials; mean ± SEM), increases in error rate were not simply due to a decrease of general motivation
400 or arousal level. Unlike error rates, VP inactivation did not change RTs overall (main effect of
401 treatment, $p = 0.15$, $F_{(1, 17)} = 2.2$, repeated measures ANOVA), suggesting that increases in error rate
402 were not simply due to motor deficits. In control condition, both RTs and lever-grip time were
403 extended in the latter half of a session, reflecting satiation-induced decreases in motivation ($p <$
404 0.025, $df = 20$, rank-sum test; Fig. 8e). By contrast, both RTs and lever-grip time were insensitive to
405 satiation after VP inactivation ($p > 0.40$, $df = 16$, rank-sum test, Fig. 8e), suggesting abolished
406 normal behavioral control by internal drive. Taken together, these results suggest that inactivation of
407 VP disrupted normal goal-directed control of behavior including loss of reward-size and satiation
408 effects.
409
410 **Discussion**
411 In the present study, we examined the activity of VP and rmCD neurons during goal-directed
412 behavior controlled by both incentive value (i.e., reward size) and internal drive (i.e., satiation level).
413 We found that reward-size coding after a reward-size cue was stronger and earlier in VP neurons
414 than in rmCD neurons. We also found that satiation-level coding was observed throughout a trial,

and appeared more frequently in VP than in rmCD neurons. In both areas, information regarding reward size and satiation level was not systematically integrated into a single neuron but was independently signaled in the population. Inactivation of the bilateral VP disrupted normal goal-directed control of action, suggesting a causal role of VP in signaling incentive and drive for motivational control of goal-directed behavior.

Past studies demonstrated that neurons in VP and rmCD encode the incentive value of cue during performing the task that offered binary outcomes (e.g., large or small reward) (Tindell et al., 2004; Nakamura et al., 2012; Tachibana and Hikosaka, 2012; Saga et al., 2017). Instead, the reward-size task used in the present study offered four reward sizes that enabled us to quantify the linearity of value coding in the activity of single neurons. With this paradigm, we found that some rmCD neurons exhibited the activity reflecting reward-size mainly during cue periods. The activity is likely to mediate goal-directed behavior based on expected reward size, because inactivation of this brain area impaired the task performance by the loss of reward-size sensitivity (Nagai et al., 2016). Our data are also consistent with previous research, which proposed the projection from the OFC to the striatum as being the critical pathway of carrying incentive information and performing a goal-directed behavior (Yin et al., 2005; Gremel and Costa, 2013; Gremel et al., 2016); we thereby confirmed the role of rmCD for mediating goal-directed behavior.

The present results, however, highlighted the more prominent role of VP in signaling incentive information for goal-directed behavior. We found that neuronal modulation by expected value was stronger and more frequent in VP neurons than in rmCD neurons. Also, the coding latency of VP neurons was significantly shorter than that of rmCD projection neurons (PANs). These results suggest that VP signals incentive value that does not primarily originate from rmCD. This suggestion may also extend to the limbic striatum, given the previous finding of similar earlier incentive signaling in VP than in the nucleus accumbens in rats (Richard et al., 2016).

The proportion of linear incentive-value coding neurons in VP (67%) was comparable to that of dopamine neurons, which was previously reported to cover 50-70% of neurons in monkeys (Schultz, 1998; Satoh et al., 2003; Matsumoto and Hikosaka, 2009). This appears to be prominent among other brain areas; in other studies using choice paradigm for neuronal recording in monkeys, value coding was observed in 30-40% of neurons in the OFC (Padoa-Schioppa and Assad, 2006; Rudebeck et al., 2013), and in 10-20% of neurons in the ventromedial prefrontal cortex (vmPFC) and the ventral striatum (VS) (Rudebeck et al., 2013; Strait et al., 2015). Although direct comparison is not possible due to different experimental conditions, our results suggest that VP is one of the most critical stages for processing incentive value for directing action.

451

452 A remaining question is: where is such rich and rapid value information derived from? One possible
453 source is the basolateral amygdala (BLA), which has a reciprocal connection to the VP (Mitrovic and
454 Napier, 1998; Root et al., 2015) and is known to contain neurons reflecting incentive value of cue
455 with short latency (Paton et al., 2006; Belova et al., 2008; Jenison et al., 2011). Recent studies
456 demonstrated that amygdala lesion impaired reward-based learning more severely than VS lesion in
457 monkeys (Averbeck et al., 2014; Costa et al., 2016), supporting the contribution of BLA-VP
458 projection in value processing. Another candidate is the projection from the subthalamic nucleus
459 (STN); VP has a reciprocal connection with the medial STN that receives projections from limbic
460 cortical areas (Haynes and Haber, 2013), composing the limbic cortico-subthalamo-pallidal
461 'hyperdirect' pathway (Nambu et al., 2002). It has been shown that STN neurons respond to cues
462 predicting rewards in monkeys (Matsumura et al., 1992; Darbaky et al., 2005; Espinosa-Parrilla et al.,
463 2015). Future studies should identify the source of the value information in terms of the latency,
464 strength and linearity of the coding.

465

466 In addition to reward-size coding, VP neurons also encoded the internal drive (satiation level) of
467 monkeys. This satiation-level coding was prominent even in the ITI phase, suggesting that this is
468 indeed a reflection of motivational state rather than task structure *per se*. A similar type of state

coding has been reported in agouti-related peptide (AgRP) producing neurons in the arcuate nucleus of the hypothalamus (ARH); hunger/satiety state modulates the firing rate of ARH-AgRP neurons (Chen et al., 2015), which regulates feeding behavior together with the lateral hypothalamus (LH) (Petrovich, 2018). Given that VP receives direct input from the hypothalamus, the satiation coding in VP might reflect the state-dependent activity originating from the hypothalamus. Although the current results indicate that both incentive value and internal drive are not systematically integrated into a single neuron level, VP may play a pivotal role in representing the two factors, which may be integrated in downstream structures, such as the mediodorsal (MD) thalamus (Haber and Knutson, 2010). The MD is one of the brain regions responsible for 'reinforcer devaluation', i.e., appropriate action selection according to the satiation of specific needs (Mitchell et al., 2007; Izquierdo and Murray, 2010), and therefore motivational value could be formulated in this area.

The causal contribution of value coding in VP was examined by an inactivation study. We found that bilateral inactivation of VP increased premature errors irrespective of incentive conditions and attenuated satiation effects without general motor impairments or decrease of general motivation. The present results together with the previous study support the view that the value coding of VP contributes to the motivational control of goal-directed behavior (Tachibana and Hikosaka, 2012). Another mechanism is also possible, such as, that suppressing general high neuronal activity in VP

(cf. Fig. 3) would promote premature response. Inactivation of VP would activate its efferent target neurons including dopamine neurons by a disinhibition mechanism, and thereby abnormally invigorate current actions (Niv et al., 2007; Tachibana and Hikosaka, 2012). However, this story is not so simple, as injection of the $GABA_A$ receptor antagonist bicuculline into bilateral VP also increased premature responses in monkeys (Saga et al., 2017). Thus, disruption of value coding in VP may be the fundamental mechanism underlying observed abnormal behavior. The loss of information regarding motivational value could promote a shift from goal-directed to habitual control of action (Dickinson, 1985; Dickinson and Balleine, 1994), which is implicated in the hallmark of addictive disorders (Everitt and Robbins, 2005; Ersche et al., 2016). Our results, therefore, emphasize the importance of future investigations into the exact neuronal mechanisms of motivational value formulation and control of actions in both the normal and abnormal state.

In conclusion, our data highlight the critical contribution of VP in goal-directed action. VP neurons independently encode information regarding incentive and drive that are essential for motivational control of goal-directed behavior. Regarding this view, VP may gain access to motor-related processes and adjust the motivation of action based on the expected reward value in accordance with the current needs.

505 **References**

506 Ahrens AM, Meyer PJ, Ferguson LM, Robinson TE, Aldridge JW (2016) Neural Activity in
507 the Ventral Pallidum Encodes Variation in the Incentive Value of a Reward Cue. J
508 Neurosci 36:7957-7970.

509 Aosaki T, Kimura M, Graybiel AM (1995) Temporal and spatial characteristics of tonically
510 active neurons of the primate's striatum. J Neurophysiol 73:1234-1252.

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

513 Belova MA, Paton JJ, Salzman CD (2008) Moment-to-moment tracking of state value in the
514 amygdala. J Neurosci 28:10023-10030.

515 Berridge KC (2004) Motivation concepts in behavioral neuroscience. Physiol Behav
516 81:179-209.

517 Castro DC, Cole SL, Berridge KC (2015) Lateral hypothalamus, nucleus accumbens, and
518 ventral pallidum roles in eating and hunger: interactions between homeostatic and
519 reward circuitry. Front Syst Neurosci 9:90.

520 Chen Y, Lin YC, Kuo TW, Knight ZA (2015) Sensory detection of food rapidly modulates
521 arcuate feeding circuits. Cell 160:829-841.

522 Costa VD, Dal Monte O, Lucas DR, Murray EA, Averbeck BB (2016) Amygdala and Ventral

523 Striatum Make Distinct Contributions to Reinforcement Learning. Neuron
524 92:505-517.

525 Darbaky Y, Baunez C, Arecchi P, Legallet E, Apicella P (2005) Reward-related neuronal
526 activity in the subthalamic nucleus of the monkey. Neuroreport 16:1241-1244.

527 Dickinson A (1985) Actions and habits: the development of behavioural autonomy. Phil
528 Trans R Soc Lond B 308:67-78.

529 Dickinson A, Balleine B (1994) Motivational control of goal-directed action. Animal Learning
530 & Behavior 22:1-18.

531 Ersche KD, Gillan CM, Jones PS, Williams GB, Ward LH, Luijten M, de Wit S, Sahakian BJ,
532 Bullmore ET, Robbins TW (2016) Carrots and sticks fail to change behavior in
533 cocaine addiction. Science 352:1468-1471.

534 Espinosa-Parrilla JF, Baunez C, Apicella P (2015) Modulation of neuronal activity by reward
535 identity in the monkey subthalamic nucleus. Eur J Neurosci 42:1705-1717.

536 Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from
537 actions to habits to compulsion. Nat Neurosci 8:1481-1489.

538 Faget L, Zell V, Souter E, McPherson A, Ressler R, Gutierrez-Reed N, Yoo JH, Dulcis D,
539 Hnasko TS (2018) Opponent control of behavioral reinforcement by inhibitory and
540 excitatory projections from the ventral pallidum. Nat Commun 9:849.

541 Gremel CM, Costa RM (2013) Orbitofrontal and striatal circuits dynamically encode the
542 shift between goal-directed and habitual actions. Nat Commun 4:2264.
543 Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K,
544 Lovinger DM, Costa RM (2016) Endocannabinoid Modulation of Orbitostriatal
545 Circuits Gates Habit Formation. Neuron 90:1312-1324.
546 Groenewegen HJ, Berendse HW, Haber SN (1993) Organization of the output of the ventral
547 striatopallidal system in the rat: ventral pallidal efferents. Neuroscience 57:113-142.
548 Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human
549 imaging. Neuropsychopharmacology 35:4-26.
550 Haber SN, Groenewegen HJ, Grove EA, Nauta WJ (1985) Efferent connections of the
551 ventral pallidum: evidence of a dual striato pallidofugal pathway. J Comp Neurol
552 235:322-335.
553 Haber SN, Lynd E, Klein C, Groenewegen HJ (1990) Topographic organization of the ventral
554 striatal efferent projections in the rhesus monkey: an anterograde tracing study. J
555 Comp Neurol 293:282-298.
556 Haynes WI, Haber SN (2013) The organization of prefrontal-subthalamic inputs in primates
557 provides an anatomical substrate for both functional specificity and integration:
558 implications for Basal Ganglia models and deep brain stimulation. J Neurosci

559 33:4804-4814.

560 Hays Jr A, Richmond B, Optican L (1982) Unix-based multiple-process system, for real-time

561 data acquisition and control.

562 Izquierdo A, Murray EA (2010) Functional interaction of medial mediodorsal thalamic

563 nucleus but not nucleus accumbens with amygdala and orbital prefrontal cortex is

564 essential for adaptive response selection after reinforcer devaluation. J Neurosci

565 30:661-669.

566 Jenison RL, Rangel A, Oya H, Kawasaki H, Howard MA (2011) Value encoding in single

567 neurons in the human amygdala during decision making. J Neurosci 31:331-338.

568 Mahler SV, Vazey EM, Beckley JT, Keistler CR, McGlinchey EM, Kaufling J, Wilson SP,

569 Deisseroth K, Woodward JJ, Aston-Jones G (2014) Designer receptors show role for

570 ventral pallidum input to ventral tegmental area in cocaine seeking. Nat Neurosci

571 17:577-585.

572 Mai JK, Paxinos G (2011) The human nervous system: Academic Press.

573 Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive

574 and negative motivational signals. Nature 459:837-841.

575 Matsumura M, Kojima J, Gardiner TW, Hikosaka O (1992) Visual and oculomotor functions

576 of monkey subthalamic nucleus. J Neurophysiol 67:1615-1632.

577 Minamimoto T, La Camera G, Richmond BJ (2009) Measuring and modeling the interaction
578 among reward size, delay to reward, and satiation level on motivation in monkeys. J
579 Neurophysiol 101:437-447.
580 Minamimoto T, Hori Y, Richmond BJ (2012) Is working more costly than waiting in
581 monkeys? PLoS One 7:e48434.
582 Mitchell AS, Browning PG, Baxter MG (2007) Neurotoxic lesions of the medial mediodorsal
583 nucleus of the thalamus disrupt reinforcer devaluation effects in rhesus monkeys. J
584 Neurosci 27:11289-11295.
585 Mitrovic I, Napier TC (1998) Substance P attenuates and DAMGO potentiates amygdala
586 glutamatergic neurotransmission within the ventral pallidum. Brain Res
587 792:193-206.
588 Nagai Y et al. (2016) PET imaging-guided chemogenetic silencing reveals a critical role of
589 primate rostromedial caudate in reward evaluation. Nat Commun 7:13605.
590 Nakamura K, Santos GS, Matsuzaki R, Nakahara H (2012) Differential reward coding in
591 the subdivisions of the primate caudate during an oculomotor task. J Neurosci
592 32:15963-15982.
593 Nambu A, Tokuno H, Takada M (2002) Functional significance of the
594 cortico-subthalamo-pallidal 'hyperdirect' pathway. Neurosci Res 43:111-117.

595 Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of
596 response vigor. Psychopharmacology (Berl) 191:507-520.
597 Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic
598 value. Nature 441:223-226.
599 Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents
600 the positive and negative value of visual stimuli during learning. Nature
601 439:865-870.
602 Petrovich GD (2018) Lateral Hypothalamus as a Motivation-Cognition Interface in the
603 Control of Feeding Behavior. Front Syst Neurosci 12:14.
604 Ray JP, Price JL (1993) The organization of projections from the mediodorsal nucleus of the
605 thalamus to orbital and medial prefrontal cortex in macaque monkeys. J Comp
606 Neurol 337:1-31.
607 Richard JM, Ambroggi F, Janak PH, Fields HL (2016) Ventral Pallidum Neurons Encode
608 Incentive Value and Promote Cue-Elicited Instrumental Actions. Neuron
609 90:1165-1173.
610 Root DH, Melendez RI, Zaborszky L, Napier TC (2015) The ventral pallidum:
611 Subregion-specific functional anatomy and roles in motivated behaviors. Prog
612 Neurobiol 130:29-70.

613 Rudebeck PH, Mitz AR, Chacko RV, Murray EA (2013) Effects of amygdala lesions on
614 reward-value coding in orbital and medial prefrontal cortex. Neuron 80:1519-1531.
615 Saga Y, Richard A, Sgambato-Faure V, Hoshi E, Tobler PN, Tremblay L (2017) Ventral
616 Pallidum Encodes Contextual Information and Controls Aversive Behaviors. Cereb
617 Cortex 27:2528-2543.
618 Satoh T, Nakai S, Sato T, Kimura M (2003) Correlated coding of motivation and outcome of
619 decision by dopamine neurons. J Neurosci 23:9913-9923.
620 Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1-27.
621 Smith KS, Tindell AJ, Aldridge JW, Berridge KC (2009) Ventral pallidum roles in reward
622 and motivation. Behav Brain Res 196:155-167.
623 Strait CE, Sleezer BJ, Hayden BY (2015) Signatures of Value Comparison in Ventral
624 Striatum Neurons. PLoS Biol 13:e1002173.
625 Tachibana Y, Hikosaka O (2012) The primate ventral pallidum encodes expected reward
626 value and regulates motor action. Neuron 76:826-837.
627 Tindell AJ, Berridge KC, Aldridge JW (2004) Ventral pallidal representation of pavlovian
628 cues and reward: population and rate codes. J Neurosci 24:1058-1069.
629 Yamada H, Inokawa H, Hori Y, Pan X, Matsuzaki R, Nakamura K, Samejima K, Shidara M,
630 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, Knowlton BJ, Balleine BW (2005) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur J Neurosci 22:505-512.

Zar JH (2013) Biostatistical Analysis: Pearson New International Edition: Pearson Higher Ed.

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

Legends

Figure 1. Reward-size task and behavioral performance. ***a.*** Sequence of a trial. ***b.*** Cue stimuli. Either stripe set (left row) or image set (right row) was used to inform the reward size (1, 2, 4, 8 drops of liquid). ***c.*** Error rate (mean ± SEM) as a function of reward size for monkeys TA and AP, respectively. ***d.*** Mean error rate as a function of normalized cumulative reward for the two monkeys. Each color indicates reward size. Curves were best fit of Eq. 1 and 2 with c = 2.1, λ = 1.9, for monkey TA; c = 4.7, λ = 2.5, for monkey AP.

Figure 2. Recording sites in VP and rmCD. ***a-b.*** Recording sites of VP and rmCD, respectively. Left:

649 CT/MR fusion image showing the position of an electrode. Right: Schematic pictures representing
650 the locations of the recorded neurons: positive reward-size coding neurons (red), negative
651 reward-size coding neurons (blue), and non-coding neurons (white). Representative slices from
652 monkey TA were used. Cd: Caudate nucleus, Put: Putamen, GPe: External segment of the globus
653 pallidus, AC: Anterior commissure.
654
655 **Figure 3.** Reward-size coding in VP and rmCD. ***a-b.*** Example activity of VP neurons showing
656 positive (***a***) and negative (***b***) reward-size coding during cue period, respectively. Left: Raster spikes
657 and spike density function (SDF, sigma = 10 ms) were aligned at task events. The colors correspond
658 to the respective reward sizes. Red bars above shadings indicate significant linear correlation at the
659 task period ($p < 0.05$, linear regression analysis), while gray bars indicate no significant correlation
660 ($p > 0.05$). Right: Relationship between firing rate (mean ± SEM) during cue period and reward size.
661 Regression lines are shown in red ($p < 0.01$). ***c-d.***. Examples of rmCD neurons. Schema of the
662 figures are the same as in **a-b**. ***e-f.*** Left: Population activities of VP neurons that were classified into
663 positive (**e**) and negative (**f**) reward-size coding neurons, respectively. Curves and shades indicate
664 mean and SEM of normalized activity to the baseline aligned at task events. Digits in each panel
665 indicate the number of reward-size coding neurons at each task period. Right: Relationship between
666 normalized neuronal activity during cue period (mean ± SEM) and reward size. ***g-h.*** Population

667 activities of rmCD neurons. Schema of the figures are the same as in **e-f**.

668

669 **Figure 4.** Time course of reward-size coding. ***a-b.*** Time-dependent change of effect size (R squared)
670 depicted with heat plots for VP neurons (***a***) and for rmCD neurons (***b***). Each panel shows data of
671 each task event. Neurons are sorted by the coding latency from the cue. Upper rows show positive
672 reward-size coding neurons, and lower rows show negative reward-size coding neurons. ***c-d.***
673 Average effect size of positive (***top***) and negative (***bottom***) reward-size coding neurons around task
674 events for VP neurons and for rmCD neurons. Digits in each panel indicate the number of
675 reward-size coding neurons at each task period.

676

677 **Figure 5.** Coding latency of the expected reward size. ***a, c.*** Effect size histogram aligned to cue for
678 positive (***a***) and negative (***c***) reward-size coding neurons, reconstructed from Figure 4c-d. The data
679 from VP (red and blue) and rmCD (gray) are depicted in the same panel. Vertical lines indicate the
680 median of coding latency. ***b, d.*** Distribution of coding latency for positive (***b***) and negative (***d***)
681 reward-size coding neurons. Asterisk indicates significant difference between VP and rmCD ($p <$
682 0.01, rank-sum test).

683

684 **Figure 6.** Dual coding of reward size and satiation level in single neurons. ***a-d.*** An example VP

neuron showing negative reward-size coding and positive satiation-level coding. ***a-b.*** Raster plots and SDF are shown for each reward condition (***a***) and for each session period (***b***). ***c.*** The relationship between firing rate during the cue period (mean ± SEM) and reward size are plotted for the session period. Linear regressions are shown as colored lines. ***d.*** Waveforms of each spike (orange) and average waveform (black) during the first minute (left) and last minute (right) in the recording session. ***e-h.*** An example rmCD neuron showing positive reward-size coding and negative satiation-level coding. Schema of the figures are the same as in ***a-d***.

Figure 7. Separate coding of reward-size and satiation-level in VP and rmCD. ***a.*** Scatter plot of standardized correlation coefficients (cue period) for reward size (abscissa) against satiation level (ordinate) are shown for VP neurons (***left***) and rmCD neurons (***right***), respectively. Colors indicate significant reward-size coding neurons (orange), satiation-level coding neurons (green), motivational-value coding neurons (yellow), and non-coding neurons (gray). Histograms in the main panels illustrate the distribution of coefficients with significant neurons with satiation level and reward size. ***b.*** Proportion of VP neurons (solid line) and rmCD neurons (dashed line) that showed satiation-level coding (***left***), reward-size coding (***center***) and motivational-value coding (***right***) for each task period (ITI, cue, pre-release, and reward periods). Asterisks indicate significant difference between VP and rmCD ($p < 0.05$ with Bonferroni correction, chi-square test). For motivational-value

703 coding (***right***), proportion of neurons with pseudo-motivational value coding by chance is shown in
704 gray. No significant difference was observed between data and estimation ($p > 0.05$ with Bonferroni
705 correction, chi-square test).
706
707 **Figure 8.** Behavioral change due to inactivation of VP and rmCD. ***a.*** Error rate in control (black) and
708 rmCD inactivation (green) sessions for monkey RO (***left***) and monkey RI (***right***). Digits in panels
709 indicate number of sessions. ***b.*** Injection sites in VP. Top panel shows representative CT/MR fusion
710 image for confirmation of injection sites in monkey BI. Bottom two panels illustrate the location of
711 injection indicated by magenta dots. ***c.*** Error rate in control (black) and VP inactivation (magenta)
712 sessions for monkeys BI (***left***) and RI (***right***), respectively. ***d.*** Change of early-error rate (***left***) and
713 length of repetitive errors (***right***) by rmCD inactivation (green) and VP inactivation (magenta).
714 Asterisks indicate significant differences from control session (* $p < 0.05$, ** $p < 0.01$, rank-sum
715 test). ***e.*** RT (***left***) and lever grip time (***right***) in the first and latter half of control session (black) and
716 of VP inactivation session (magenta) (mean ± SEM).

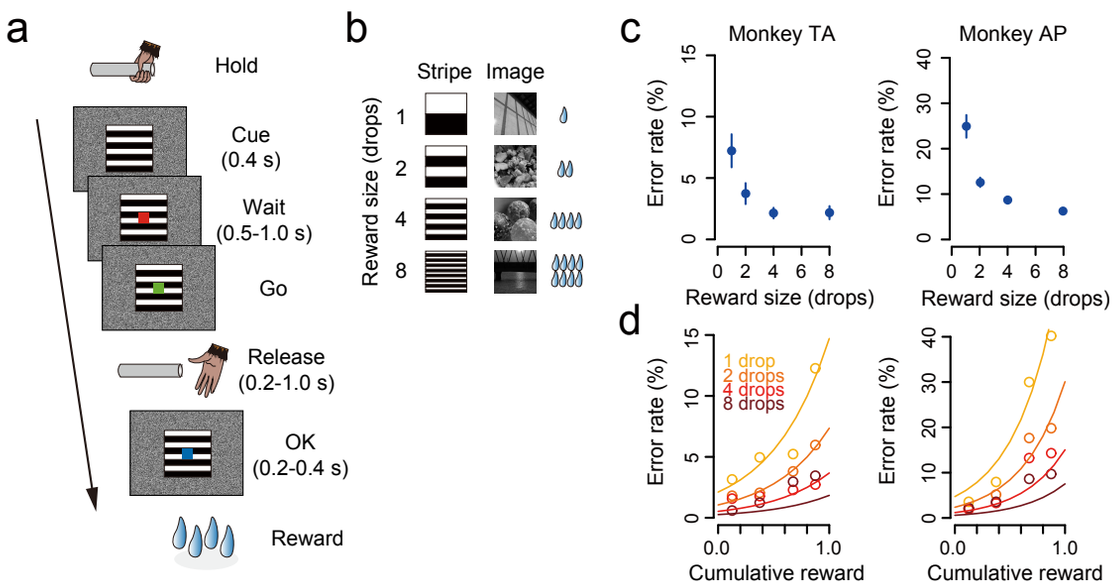

Figure 1

Fujimoto et al.

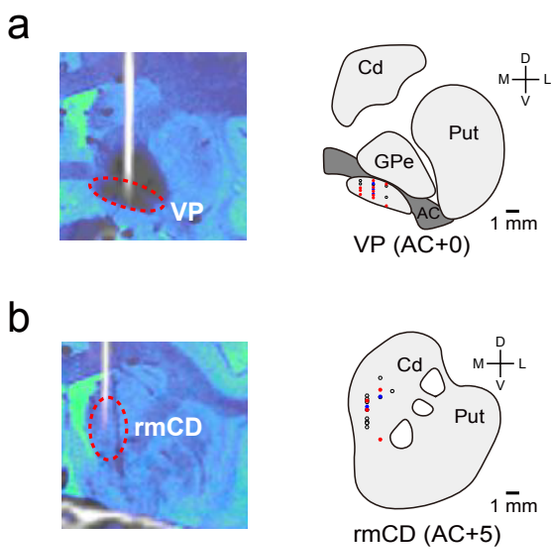

Figure 2

Fujimoto et al.

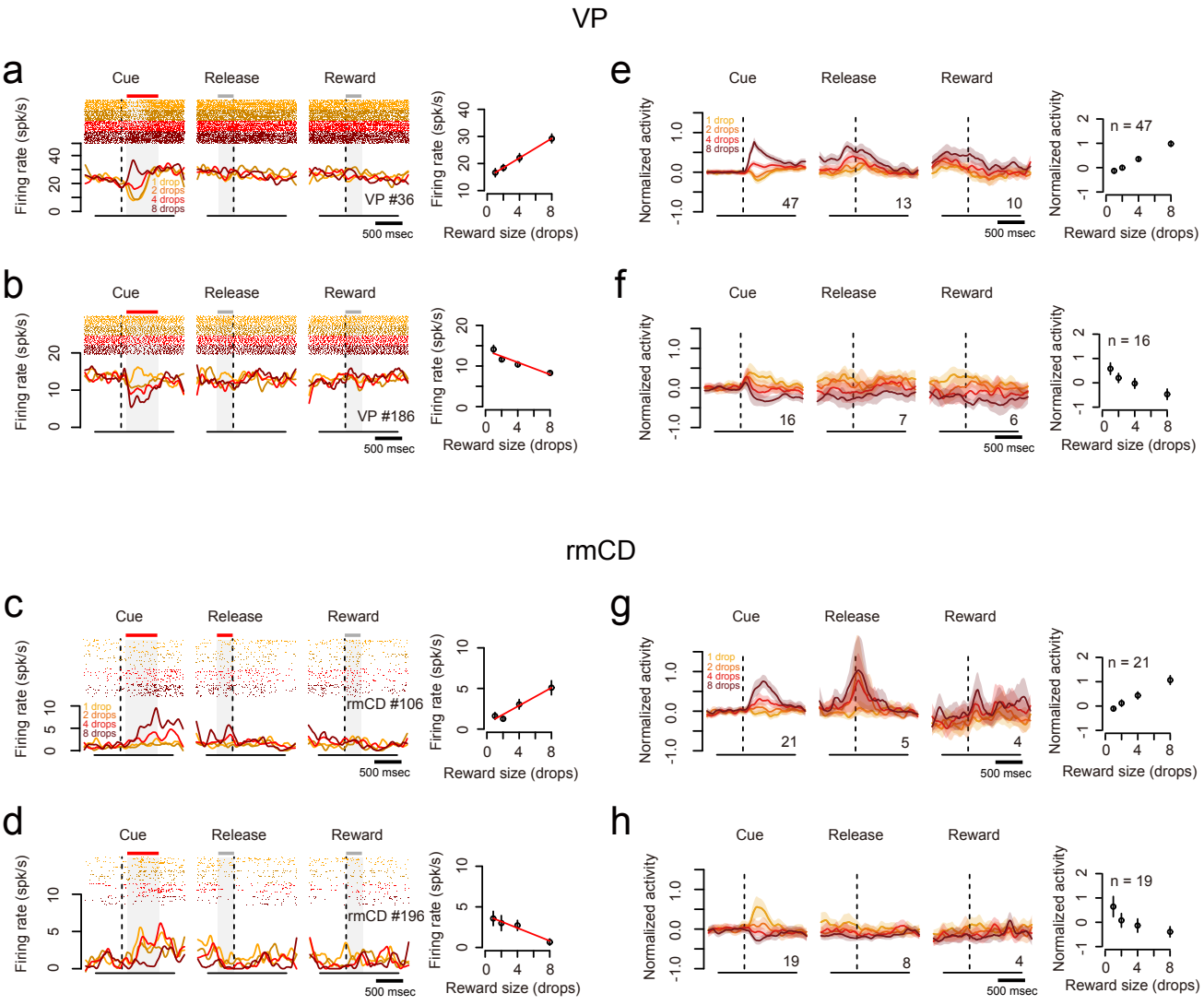

Figure 3

Fujimoto et al.

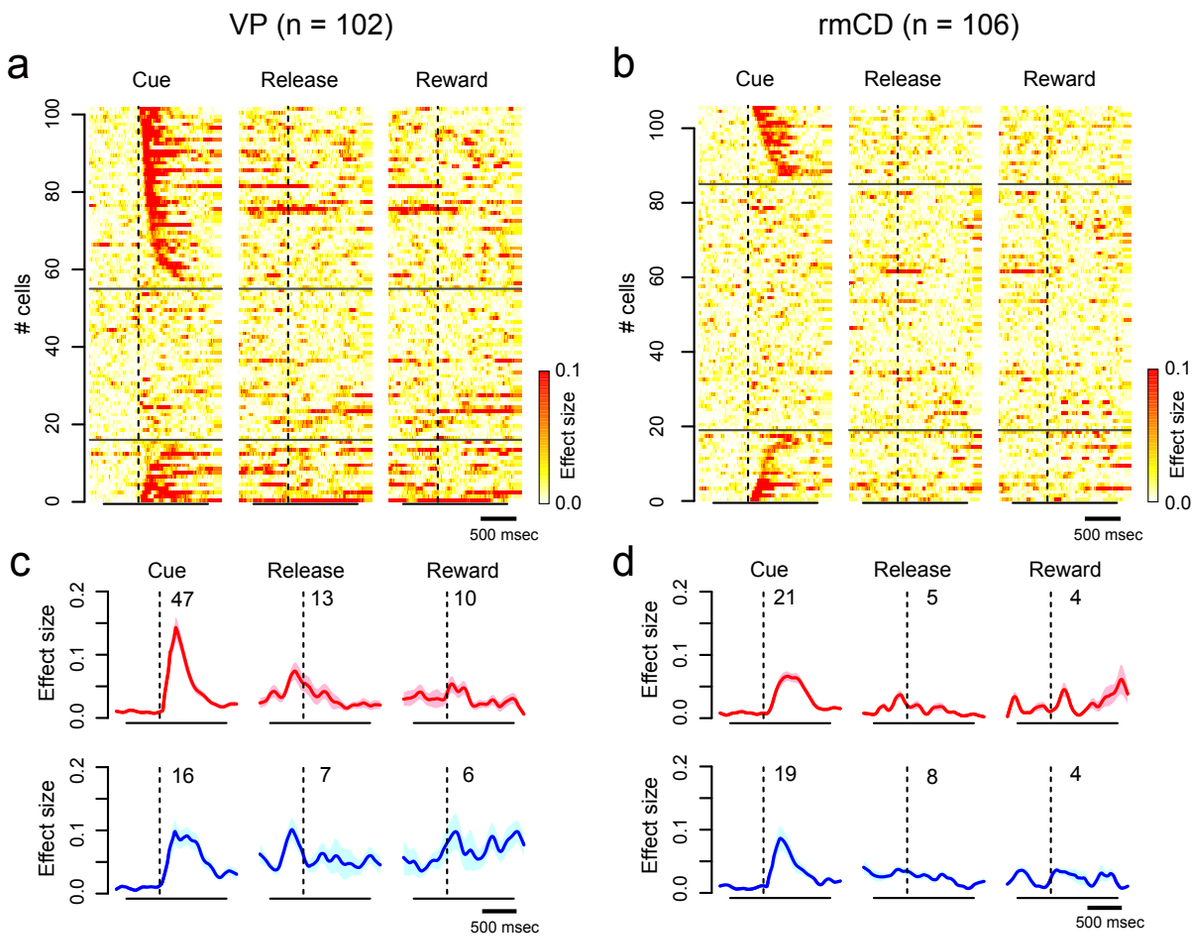

Figure 4

Fujimoto et al.

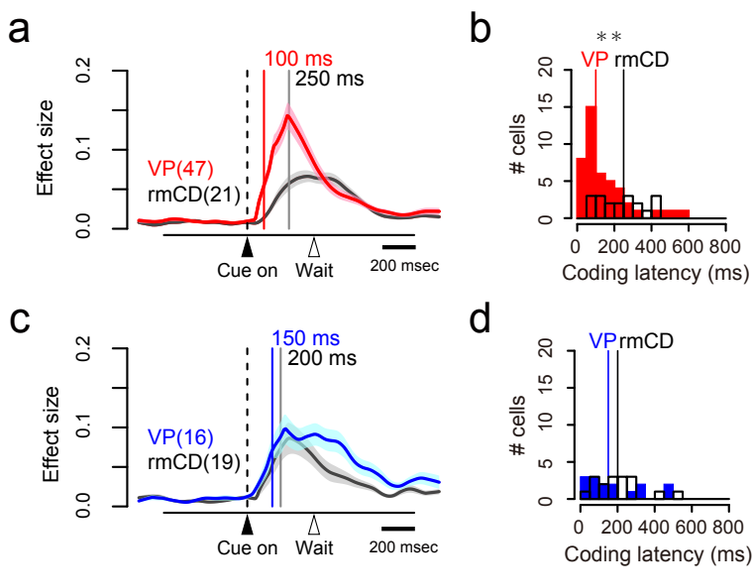

Figure 5

Fujimoto et al.

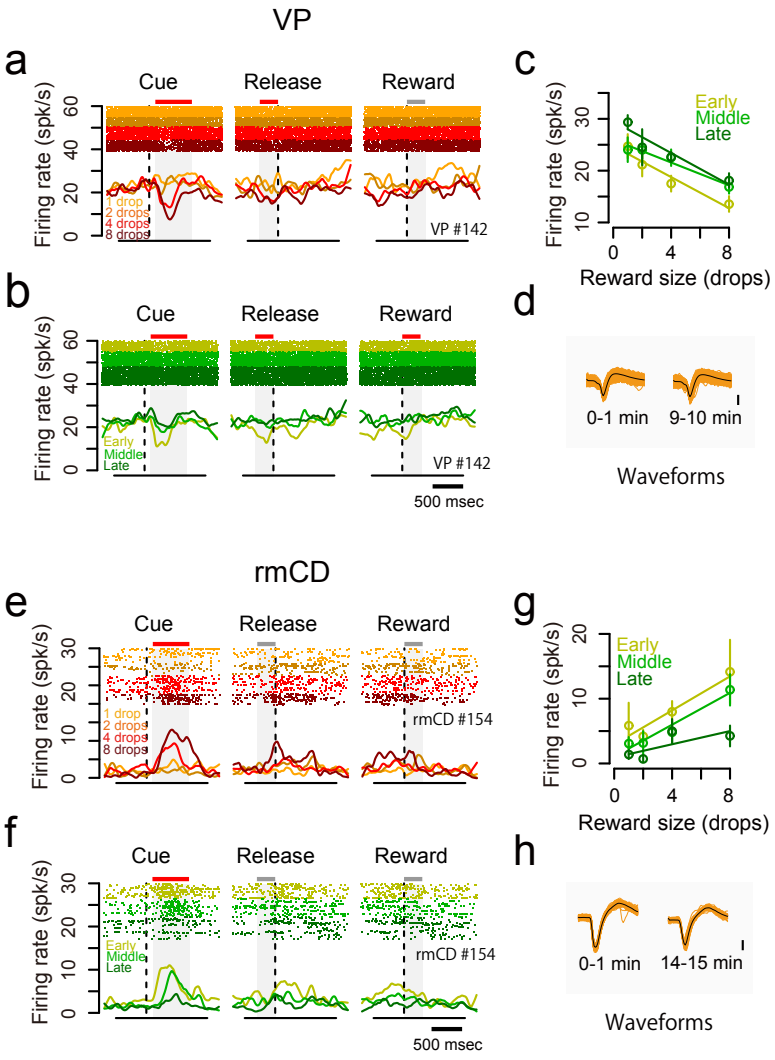

Figure 6

Fujimoto et al.

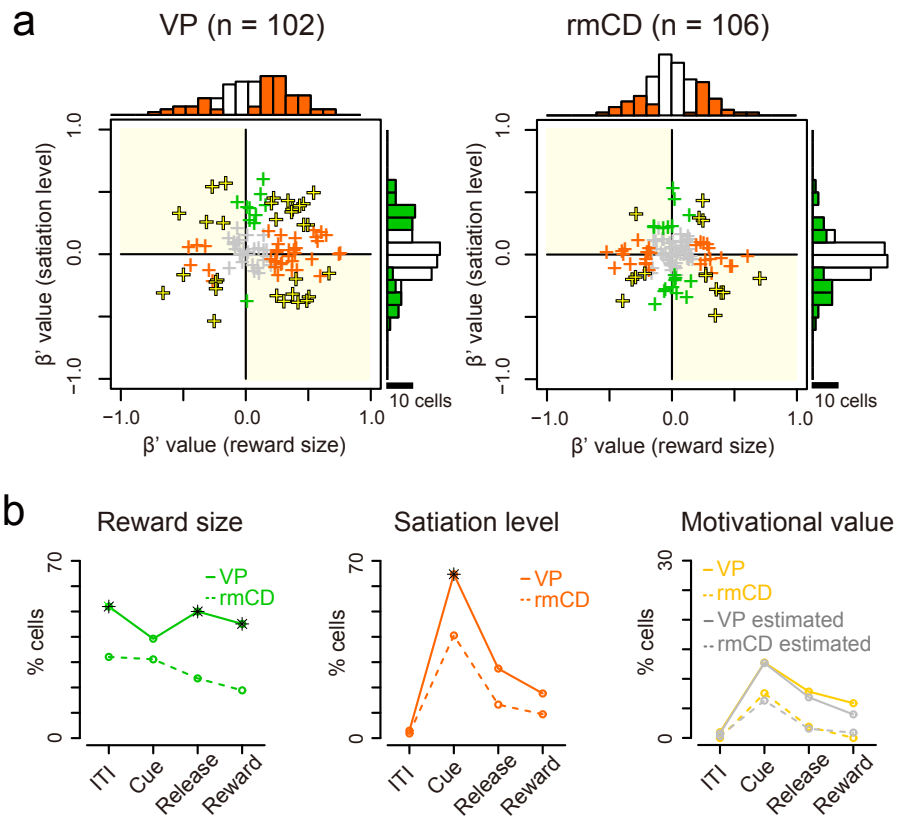

Figure 7

Fujimoto et al.

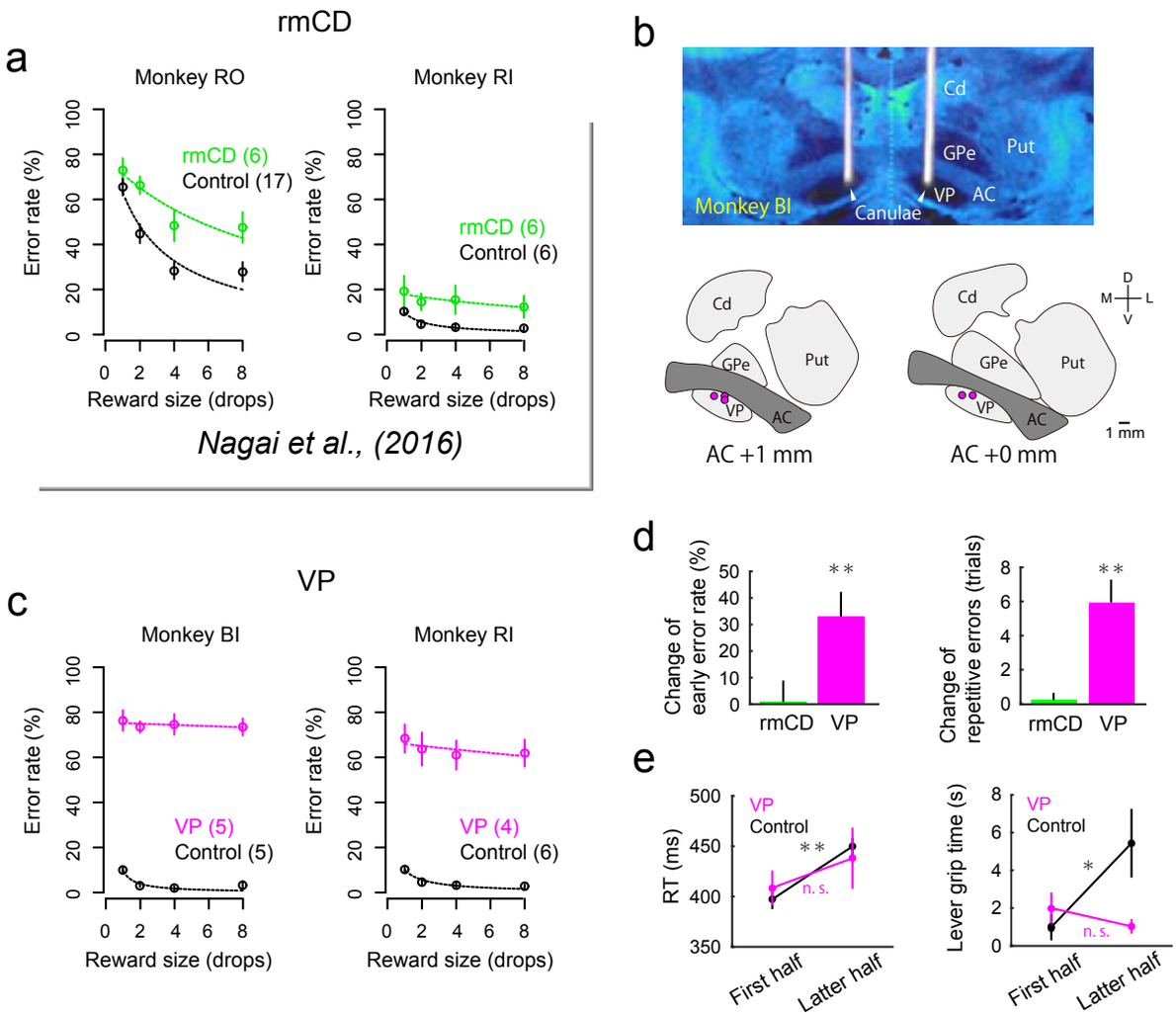

Figure 8

Fujimoto et al.