



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

The dorsomedial striatum encodes net expected return, critical for energizing performance vigor

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Wang, Alice Y., Keiji Miura, and Naoshige Uchida. 2013. "The dorsomedial striatum encodes net expected return, critical for energizing performance vigor." <i>Nature neuroscience</i> 16 (5): 10.1038/nn.3377. doi:10.1038/nn.3377. http://dx.doi.org/10.1038/nn.3377 .
Published Version	doi:10.1038/nn.3377
Accessed	April 17, 2018 4:36:29 PM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:11878977
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)



Published in final edited form as:

Nat Neurosci. 2013 May ; 16(5): . doi:10.1038/nn.3377.

The dorsomedial striatum encodes net expected return, critical for energizing performance vigor

Alice Y. Wang¹, Keiji Miura^{1,2,3}, and Naoshige Uchida¹

¹Center for Brain Science, Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA

²PRESTO, JST, Saitama, Japan

³Graduate School of Information Sciences, Tohoku University, Sendai, Japan

Abstract

Decision making requires an actor to not only steer behavior towards specific goals, but also determine the optimal vigor of performance. Current research and models have largely focused on the former problem of how actions are directed, while overlooking the latter problem of how they are energized. Here, we designed a self-paced decision-making paradigm that showed that rats' performance vigor globally fluctuates with the net value of their options, suggesting that they maintain long-term estimates of the value of their current state. Lesions of the dorsomedial (DMS), and to a lesser degree, in the ventral striatum (VS) impaired such state-dependent modulation of vigor, rendering vigor to depend more exclusively on the outcomes of immediately preceding trials. The lesions, however, spared choice biases. Neuronal recordings showed that the DMS is enriched with net-value-coding neurons. In sum, the DMS encodes one's net expected return, which drives the general motivation to perform.

Given stock options A and B, which one do you choose? The field of decision making and neuroeconomics have made significant progress on this problem of 'action selection'^{1,2}. Optimal behavioral selection, however, not only depends on the ability to choose *which* action to perform but also the appropriate vigor to perform. For example, it may be wise to flexibly adjust one's motivation to invest according to the overall state of the stock market, such as investing with lower frequency during an economic crisis.

The importance of properly regulating response vigor becomes apparent when one considers costs associated with performing an action. Rapid responding may increase the rate of obtaining rewards but may also increase energetic costs. Conversely, slow responding may be energetically efficient, yet it delays all future rewards. Importantly, the cost of delaying future rewards critically depends on the net expected future reward given the current state of the animal. It is thus proposed that selection of response vigor *should* depend on the average or net expected reward (or "state value") while action selection depends on values specific to individual options (and the relative value between them)³⁻⁵. This idea echoes two aspects of motivation proposed in classic animal psychology: the motivation to steer towards making a specific action (the "directing" effect or action-specific motivation) and the

Correspondence should be addressed to N.U. (uchida@mcb.harvard.edu)..

Supplementary information is available in the online version of the paper.

AUTHOR CONTRIBUTIONS A.W. wrote the manuscript, performed the experiments, performed the analyses, and was involved in the experimental design. K.M. performed the simulations. N.U. was involved in preparing the manuscript and experimental design.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

motivation to generally “arouse” or speed-up all pre-potent actions in a non-specific manner (the “energizing” effect or action-general motivation)^{3,6,7}. It should be noted that the directing effect may also speed-up actions towards particular goals but the energizing effect acts diffusely on a wider set of actions. Experimentally, many classic studies in animal psychology have shown that response vigor is modulated by the rate of reward, providing some limited support for the energizing effects of average reward rate⁸. More recently, it has been shown that manipulations of the size or probability of rewards affect choice direction and latencies in various choice tasks, which highlights motivation's directing effects^{9–11}. However, whether response vigor is indeed regulated by average or net expected reward, that is, whether motivation energizes behavior in a global manner, remains controversial.

It is believed that the basal ganglia play important roles in action selection^{1,2,12,13}. Some studies on patients with Parkinson's Disease and lesion studies using animal models, however, have suggested that the basal ganglia also play a prominent role in the regulation of response vigor¹⁴. Mounting evidence suggests that specific areas of the striatum encode specific types of values and regulate distinct aspects of value-dependent behavior^{15–17}. Historically, the striatum, particularly the VS has been linked to motivation^{18,19}, although other studies implicate the role of dorsal striatum in motivation^{20–22}. However, previous studies have not separated the directing versus energizing aspects of motivation, therefore whether these processes can be mapped onto specific parts of the striatum remains unknown.

To address these questions, we designed a task that allows us to study both the directing and energizing aspects of behavioral regulation. We first examined whether response vigor is indeed modulated by net expected future rewards and how this process is separable from the directing effects that are specific to individual actions. Second, using lesions, we examined which part of the striatum is involved in the regulation of net value-dependent response vigor. Finally, we recorded the activity of single neurons in DMS and VS. The results demonstrate a critical role of the DMS in net-value dependent regulation of response vigor.

RESULTS

Self-paced decision task

We designed a self-paced, two-choice behavioral paradigm, where rats self-initiate a trial by poking its snout into a central odor port. After poking into the odor port (trial initiation), a randomly-selected odor cue was presented after a variable (300–500ms) delay (Fig. 1a,b)²³. We used pure odors as well as binary mixtures of odors of various ratios, where the dominant component in the mixture determined which reward port delivers water. To examine the effects of the animal's option values on behavior, we systematically varied the amount of water delivered at the left and right ports in blocks of 40–60 trials (Fig. 1c)^{16,24}.

Consistent with previous studies^{16,24}, the animal's choices were biased by the relative value of the options (Fig. 1d). For example, animals chose the right port more frequently in blocks where the right port delivered relatively more water, as shown in the shifts in psychometric curves (Fig. 1e). Importantly, this bias towards the more valuable side was invariant to changes in the net value of the options (Fig. 1e, red versus pink, two-sided t-test, $t_{28} = 0.5$, $P > 0.05$; blue versus cyan, two-sided t-test, $t_{28} = 0.8$, $P > 0.05$, $n = 15$ rats). Further, choice biases existed primarily for mixed-but not pure-odor trials, indicating that rats did not prematurely make their selection until after odor sampling (Fig. 1d). Thus, action selection depended on the relative value of the animals' options.

Latency to trial initiation depends on the net value

We next examined the rats' response vigor by measuring the latency to initiate a trial after completing a previous trial. Overall, animals exhibited shorter latency in blocks of high net value compared with blocks of low net value (Fig. 1h; t-test, $n = 15$ rats, $t_{58} = 5.0$, $P < 0.05$, Supplementary Figs. 1 and 2). This change in initiation time consisted primarily of the slowing down of movement from the reward port to odor port, brief pauses near the port (as reflected by the increase in trials with 5–10s initiation time), as well as a relatively minor contribution of disengagement from the task (grooming or resting outside the task, as reflected by the increase in trials with long (>10s) initiation time) (Supplementary Fig. 1c). On average, alternations in initiation time occurred within 5 trials after a change in block value (time constant: 4.9 ± 1.3 trials, $n = 15$ rats) (Fig. 1g). These results demonstrated that trial initiation time depended on the net value of the animal's choices (Fig. 1i).

Past models predict that in states of high overall return, motivation should generally increase, while in states of low net value, motivation should generally decrease^{3,4}. However, the above results do not demonstrate that response vigor is block- or state-dependent. Under our task design, animals may simply initiate slower immediately after receiving small rewards and faster immediately after receiving large rewards. To examine whether the rats' response vigor is globally energized by the block, we next examined initiation time after left and right reward trials separately.

If the motivation to initiate trials is driven exclusively by the immediately preceding trial's reward size, the initiation time following left reward trials should be the *same* as far as the left reward size is identical (red and cyan blocks). However, we found that animals were significantly faster after left reward in *red* blocks (Fig. 2a, two sided t-test, $t_{12} = 2.4$, $P < 0.01$, $n = 7$ sessions in one rat). Moreover, latency after left trials in cyan blocks was no faster than in pink blocks, where left reward size is smallest (Fig. 2a, two sided t-test, $t_{12} = 0.7$, $P > 0.05$, $n = 7$ sessions in one rat). We saw a similar pattern for initiation time following all right reward trials (Fig. 2b) and similar patterns were observed in other rats. These results suggest that the motivation to initiate trials cannot be explained simply by the animal responding to the immediate previous trial's reward size. Rather, the overall value of the block determined the animal's motivation to perform.

To more succinctly quantify changes in initiation time across blocks (i.e., how initiation time varies with left and right value), we represented trial initiation time with a single vector projected onto "value space" with a polar angle (θ) and amplitude (r) (Fig. 2c). We regressed initiation time with value of left and right choices of the block (Q_L and Q_R , respectively):

$$\text{Trial initiation time (s)} = \beta_0 + \beta_L \cdot Q_L + \beta_R \cdot Q_R \quad (1)$$

These regression coefficients were used to project a single vector onto value space (Fig. 2c). The vector conveys two variables: θ (in polar coordinates) reveals how trial initiation time varies across blocks (i.e. the relative contribution of Q_L and Q_R) and r is the strength of the modulation (Fig. 2c). Essentially, the vector illustrates the strength and the relative contribution of left and right value on trial initiation time. For example, a negative correlation with net value (the larger the value of left *and* right reward, the shorter the initiation time) will be projected as a vector directed towards the lower-left quadrant of value space, or $\theta = 225^\circ \pm 22.5^\circ$ (Fig. 2d, left). If initiation time negatively correlates with left value only, the vector would be horizontal, or $\theta = 180^\circ \pm 22.5^\circ$ (Fig. 2d, middle). Further, if initiation time negative correlates with right value only, the vector would be vertical, or $\theta = 270^\circ \pm 22.5^\circ$ (Fig. 2d, right). For more stringent classification, we used a 95% confidence interval on θ calculated by bootstrapping.

Using the same example rat shown in Fig. 2a, we obtained a vector of $\theta = 222^\circ$ (95% confidence interval: lower bound = 213° , upper bound = 231°) and $r = 3.3$ (Fig. 2e). Since θ falls within 22.5° of 225° and the confidence interval did not cross 180° or 270° , the rat's trial initiation time is considered net value-dependent. This analysis allowed us to summarize initiation times following different trial outcomes, such as correct, error, left reward, or right reward trials, separately ($n = 1019, 244, 512, \text{ and } 507$ trials, respectively; Fig. 2e and f).

Using this analysis, we found that on average, trial initiation time was net value-dependent, regardless of previous trial type (Fig. 2g–h, $n = 15$ rats, correct trials: $\theta = 223^\circ \pm 2.3$, $r = 2.3 \pm 0.3$; error trials: $\theta = 222^\circ \pm 21$, $r = 2.4 \pm 0.4$; left trials $\theta = 212^\circ \pm 6.2$, $r = 2.4 \pm 0.3$; right trials $\theta = 233^\circ \pm 4.9$, $r = 2.5 \pm 0.3$). Thus, response vigor depended predominantly on the net value of both options, more so than on the immediate previous trial's reward size. However, we observed that θ_R (vector angle representing initiation time immediately after right reward) tended to be slightly larger than θ_L (vector angle representing initiation time after left reward) (Fig. 3a,b, paired t-test, $t_{14} = 2.9$, $P < 0.05$). This suggested that although initiation time primarily depends on the net value of the block (as most vectors were near 225°), the immediate previous trial may still have an influence.

To quantify how the outcomes of multiple previous trials affected initiation time, we regressed initiation time by previous reward outcomes, where T is the current trial (e.g., Q_{T-1} is the reward outcome of one trial before):

$$\text{Trial initiation time (s)} = \beta_0 + \beta_{T-1} \cdot Q_{T-1} + \beta_{T-2} \cdot Q_{T-2} + \dots + \beta_{T-5} \cdot Q_{T-5} \quad (2)$$

We obtained the regression coefficients and fitted them to an exponential to obtain a decay constant, τ (Fig. 3c). The result showed that initiation time depends on integration over multiple previous trials ($\tau = 4.6 \pm 1.5$ trials, $n = 15$ rats). The difference between the angles, $\theta_R - \theta_L$, was very small (Fig. 3d, median: 21.3° degrees, mean: 19.6°), and was not significantly different from the $\theta_R - \theta_L$ for trial shuffled controls ($P > 0.05$, bootstrap test), suggesting that the effect of the immediate previous trial outcome was small. The analysis based on the time constant (τ) complements the vector analysis in that although the effects of various trial types (e.g. left, right, correct, error) are masked in the process of computing τ , we could obtain the number of trials back in time on which initiation time depends.

In all, these results demonstrate that the overall value of the block energized the initiation of trials: animals exploit the task during states of high value, since the opportunity cost of wasted time is high, as predicted by normative model of motivation⁴.

DMS is critical for net-value dependence of response vigor

The above behavioral paradigm allows us to dissociate the effect of values on choice and response vigor. We next performed selective lesions of the DMS or VS to test whether either region is involved in choice bias or response vigor. After behavioral training, we injected ibotenic acid, an excitotoxic agent, into the anterior DMS or VS bilaterally, which caused lesions in a relatively large portion of each area (~ 2 mm in diameter) (Fig. 4 and Supplementary Fig. 4). In control animals, we injected saline (“sham-lesion”). The behavior of VS, DMS, and sham-lesioned animals were then examined ($n = 5$ rats per condition; $3,365.6 \pm 115$ trials in 7.86 ± 0.09 sessions per animal). None of the animals showed impaired choice biases (Fig. 4b). VS-lesioned animals showed a trend for larger choice biases (Fig. 4b) but the shifts were not statistically significant (Fig. 4c, unpaired two-sided t-test, $n = 5$ rats per lesion condition; sham versus VS-lesion for cyan and blue blocks, $t_{18} = 0.6$, $P > 0.05$; sham versus VS-lesion for pink and red blocks, $t_{18} = 1.8$, $P > 0.05$). For mix-odor trials only, VS-lesioned animals had significantly more errors towards the relatively

better choice (Fig. 4d, sham: 37.9%, DMS: 38.6%, VS: 45.1% error trials towards larger choice for mixed-odor trials; unpaired two-sided t-test between sham and VS-lesioned condition, $n = 5$ rats per condition, $t_8 = 3.1$, $P < 0.05$). This higher error rate for difficult trials was not due to inadequate odor sampling duration (Fig. 4e, unpaired two-sided t-test, all $P > 0.05$; sham vs. DMS pure odors: $t_8 = 0.2$; mixed odors: $t_8 = 0.6$; sham vs. VS pure odors: $t_8 = 0.4$, mixed odors: $t_8 = 0.8$).

Next, we examined the effect of lesions on trial initiation time. Although the mean initiation time was normal (Fig. 5a, left panels, unpaired two-sided t-test comparing high and low net value blocks, $n = 5$ rats per condition, all $P < 0.05$; sham, $t_{18} = 2.6$; DMS, $t_{18} = 3.0$; VS, $t_{18} = 4.3$), we observed the effect of lesions when trial initiation time was analyzed separately for different trial types, using the vector representation method (Fig. 5b–c and Supplementary Fig. 5b). Trial initiation vectors for correct, error, and all trials were dependent on net-value (i.e., within 22.5 degrees of 225° in all example rats, $n = 3413$, 3041, 3776 total trials for sham, DMS, and VS example). In contrast, vectors obtained for left and right reward trials separately revealed striking differences across lesion conditions (Fig. 5c). In the sham animals, trial initiation vectors for both left and right trials were within $\pm 22.5^\circ$ of 225°, indicating net value-dependent regulation of response vigor (Fig. 5c, sham example: left trials $\theta_L = 216^\circ$; right trials $\theta_R = 225^\circ$; Supplementary Fig. 5b, population: left trials $\theta_L = 219^\circ \pm 5.8$, right trials $\theta_R = 236^\circ \pm 7$). For the DMS-lesioned rat, however, initiation times after left trials depended on left value alone while initiation after right trials depended on right value alone (Fig. 5c, DMS example: $\theta_L = 190^\circ$, $\theta_R = 289^\circ$; Supplementary Fig. 5b, population: left trials $\theta_L = 189^\circ \pm 4.5$, right trials $\theta_R = 275^\circ \pm 6.8$). The DMS-lesioned animal's trial initiation vectors for left reward and right reward trials were nearly orthogonal to one another (Fig. 5c, DMS example: $\theta_R - \theta_L = 99^\circ$; Fig. 5e, DMS: $\theta_R - \theta_L = 87^\circ \pm 8.5$) unlike sham-lesioned animals (Fig. 5e, sham: $\theta_R - \theta_L = 20.1^\circ \pm 5.3$; DMS vs. sham condition, $P < 0.05$, t-test; $t_8 = 6.1$), indicating that its motivation depend primarily on the value of the immediately preceding trial. A less dramatic but statistically significant effect was seen in the VS-lesioned condition (Fig. 5c, VS example: left trials $\theta_L = 207^\circ$, right trials $\theta_R = 248^\circ$; Supplementary Fig. 5b, population: left trials $\theta = 203^\circ \pm 3.7$, right trials $\theta = 243^\circ \pm 2.9$). Vectors for left and right trials were more widely separated in VS animals (Fig. 5c, VS example: $\theta_R - \theta_L = 41^\circ$; Fig. 5e, VS: $\theta_R - \theta_L = 40.7^\circ \pm 5.4$) than for the sham animal (VS vs. sham condition: t-test, $t_8 = 2.8$, $P < 0.05$), yet not orthogonal as in the DMS condition (VS vs. DMS: t-test, $t_8 = 5.1$, $P < 0.05$). Lastly, we examined the effects of multiple previous trial outcomes on initiation time and found that the influence of the immediately previous trial was significantly stronger for the DMS-lesion condition compared to the sham condition (Fig. 5f, $\tau = 1.2 \pm 0.4$ trials for DMS; DMS vs. sham: $t_8 = 2.7$, $P < 0.05$). A similar but nonsignificant tendency was seen in the VS (Fig. 5f, $\tau = 2.6 \pm 0.9$ trials for VS; VS vs. sham, t-test, $t_8 = 1.7$, $P > 0.05$). In total, these results demonstrate that the DMS had a larger role than the VS for integrating the total value of one's options, which was critical for promoting net value-dependent modulation of response vigor.

DMS is enriched with neural representations of net value

To examine the neural mechanisms that may underlie the DMS' importance in net-value dependent modulation of trial initiation, we monitored the neural activity in the DMS and VS (Fig. 6a,b) of another set of rats as they performed the behavioral paradigm (Supplementary Fig. 6). We recorded 522 neurons, with 364 from DMS and 158 from VS while rats performed the task ($n = 4$ rats; 91 ± 37.4 and 39.5 ± 56.9 neurons in DMS and VS per animal; mean \pm s.d.; total neurons recorded from each rat: 96, 154, 209 and 63; all from the left hemisphere).

First, we examined each neuron's time of maximum firing by pooling all trial types, and found that as a population, the inter-trial epoch was a time of high neural activity (30%, or 152/522 cells had maximal firing during this inter-trial interval)(Fig. 6a, Note that PETHs show activity aligned to the median timing of the epochs. Since inter-trial intervals were especially variable, the peak timing of neurons appears especially smeared during this period). Based on this result and the lesion results on trial initiation, the following neural analysis focuses on the epoch immediately preceding trial initiation (0–300ms before odor-poke-in). During this pre-initiation epoch, rats have just prepared to start a new trial but have not yet made their decision. Moreover, this period is less contaminated by movement activities that are unrelated to the task such as grooming or resting, since animals are prepared to perform the task. Finally, due to the blocked task structure, animals already know the expected values of their options during this epoch (except for first 5–10 trials of each block), enabling us to examine value-related activity.

To quantify how the neural responses during the pre-initiation epoch are modulated by block-wise changes in value, we projected each neuron onto the same value space as our behavioral analysis (Fig. 6c–h). We regressed firing rate with left and right value (Q_L and Q_R , respectively).

$$\text{Firing rate (spikes/s)} = \beta_0 + \beta_L \cdot Q_L + \beta_R \cdot Q_R \quad (3)$$

We systematically mapped neural responses to their respective decision making processes: neurons encoding left value alone point towards 0° or 180° ; those encoding right value are 90° or 270° . Such neurons encode values in a 'menu-invariant' manner ("absolute value"). Next, neurons encoding relative-values ($Q_R - Q_L$) are modulated towards 135° or 315° , and neurons encoding net-values ($Q_L + Q_R$) are modulated along 45° or 225° . Value-coding neurons are defined as those that showed significant modulation in any direction in this value space ($P < 0.01$, F-test).

We then examined the population of neurons whose activity was significantly modulated by value during pre-trial initiation in the DMS and VS (Fig. 7a and Supplementary Fig.7). The activity of 31% (113/364) of DMS neurons and 22.8% (36/158) of VS neurons were significantly modulated by value ($P < 0.01$, F-test in regression analysis). Although the DMS contained a higher proportion than the VS, the difference was not significant (χ^2 test, $\chi^2_1 = 3.41$, $P = 0.055$). However, the distribution of the types of value-coding neurons in the two regions differed. In the VS, the proportion of absolute and net value coding neurons were significantly represented above chance ($P < 0.01$, binomial test), but relative value coding neurons were not (Fig. 7b, bottom; $P > 0.0125$, Bonferroni correction for proportion of relative value coding neurons). Furthermore, the distribution of value-coding categories (left-absolute, right-absolute, relative, and net value) was not significantly different from uniform (χ^2 goodness of fit test against uniform, $\chi^2_3 = 2.0$, $P = 0.57$). In the DMS, all categories of neuron types were represented above chance level ($P < 0.0125$, binomial test). Moreover, the distribution was significantly different from uniform (χ^2 goodness of fit test against uniform, $\chi^2_3 = 12.0$, $P = 0.0077$). The non-uniformity was due to the predominance of a single category: net value coding neurons (Fig. 7c). Net value coding neurons formed the only category that significantly deviated from what is expected from a uniform distribution ($P = 0.0028$, binomial test). In total, net value coding neurons were most dominantly represented in the DMS, although all coding types were present.

An advantage of the polar coordinate method is that in addition to examining the number of neurons per category (classified by θ), we can also take into account the strength of representation (amplitude, r , of the vectors). Thus, we computed resultant vectors for each category of neurons with their 95% confidence boundaries (bootstrap) and projected them

onto the polar coordinate (Fig. 7d). This analysis also supported that the DMS predominantly represents net value.

Finally, we examined whether the activity of net value coding neurons were significantly affected by the immediately preceding trial's outcome. The vector analysis indicated that the majority of neurons (>80%) showed net-value dependent firing modulation regardless of previous trial's outcome (Fig. 8a,b). As a population, the angle between vectors for firing rate immediately after left reward and after right reward trials were not statistically significant (Fig. 8c, paired t-test, θ_L vs. θ_R , $t_{53} = 0.9$, $P > 0.05$). Moreover, we regressed firing rate with net value and previous trial outcome and found that 94% (51/54) of the neurons were significant for net value while (7/54) 13% were significant for both net value and previous trial outcome while only 3.7% (2/54) were significant for previous trial outcome alone.

DISCUSSION

In the present study, we teased apart two orthogonal components of decision making: how to choose between alternative actions and how to choose the vigor applied to actions. First, we demonstrated that response vigor fluctuates globally with the net value of the animal's options (or the average expected reward rate). Second, lesions in the DMS strongly diminished the global effects of net value on motivation, rendering animals' response vigor to depend on immediately preceding trial outcomes rather than on the net value. A weaker effect was observed in the VS. In contrast, action selection depended on the relative value of available options and was not affected by lesions of DMS nor VS. Finally, the DMS was more enriched with net value coding neurons than the VS. Together, these results demonstrated a critical role of the DMS in net-value dependent regulation of response vigor.

In decades past, measurement of response vigor in behaviors such as lever-press tasks and key-pecking tasks have been central^{6,7,25}, while the study of action selection was not well measured. In recent years, the opposite problem has emerged, where the primary focus is on action selection, often neglecting inter-trial intervals. Although there have been a few studies that examined how incentive values of goals modulate reaction times of actions directed towards specific goals, such questions differ in that they do not address how long-term estimates of the animal's state can have global effects on response vigor^{9-11,26}. The present study is one of the first to demonstrate that the total value of the animal's options (i.e. the value of the animal's state) globally influences vigor. Such global considerations for choosing vigor have been previously hypothesized to be an optimal strategy for appropriately exploiting a task only when it is worthwhile, and slowing down or relaxing when it is not. The present study paves a way towards dissociating and examining both aspects of decision making in a unified task.

The role of dorsomedial striatum in energizing motivation

The VS has long been linked to motivation^{18,19}, because of its role in intracranial self-stimulation²⁷, drugs addiction²⁸, effort-related decision making²⁹, and its anatomical connections with limbic structures³⁰. In contrast, dorsal striatum is less known for motivation but studies using dopamine-deficient mice have shown that restoration of dopamine into the dorsal striatum can sufficiently rescue animals' motivation to engage in reward-oriented behavior²⁰. Human studies have also demonstrated the dorsal striatum's role in motivation^{9,31,32}. Further, DMS lesions impair reaction time and initiation latency in rats^{22,33,34}, electrical stimulation of primate caudate nucleus influences reaction time towards immediate goals¹⁰ and several physiology and lesion studies implicated the DMS in flexible, reward-oriented behavior³²⁻³⁶. Despite these results, none of the studies explicitly separated the directing and energizing effects of motivation. Our results suggest that the

DMS may normally be critical for integrating the values of the animal's options, or computing net value, and regulating response vigor. In support of this, our electrophysiological results show that net-value representation is enriched in the DMS.

Our results do not exclude the possibility that the VS was involved in the regulation of response vigor in other behavioral contexts. It should be noted that in our task, the rate of reward depends on animals' actions, that is, the task is instrumental (rather than Pavlovian). It is possible that there are distinct control mechanisms for response vigor in Pavlovian vs. instrumental contexts.

Recent studies in rats and humans have indicated that DMS is involved in goal-directed behavior (or model-based learning)³⁷. A hallmark of goal-directed behavior, but not of habitual behavior, is its sensitivity to the changes of the outcome values associated with specific actions (action-outcome associations). For example, if the outcome value of one of two potential actions is reduced, the performance of the devalued, but not the non-devalued action decreases^{38,39}. It has been shown that lesions of DMS render animals insensitive to such devaluations, suggesting that DMS supports behavior based on associations between specific actions and outcomes. In contrast, our result showed that DMS lesions impaired the response vigor's dependency on the net value (values general to potential actions) but spared choice biases that depend on the values associated with specific actions. We note that previous studies highlighted the importance of the more posterior region of the DMS than that studied in the present study³⁸. It is, therefore, possible that the anterior and posterior DMS underlie two distinct aspects of goal-directed behavior: anterior DMS in behavior dependent on action-general values (net values) and posterior DMS in behavior dependent on action-specific values.

The present study focused on the value of options within the task. However, one can argue that rats are choosing between performing the task we presented and other activities outside the task (such as grooming or resting). It is thus possible that the change in trial initiation time is due to the changes in the value of performing the task relative to other activities. Alternatively, the normative theory of motivation would predict that the 'vigor' of grooming or other activities outside the task is also slowed down during the period of low-net value, similar to trial initiation time in the task⁴. Measurement of the vigor of outside activities (e.g. testing whether rats groom rapidly during high net value) will allow us to tease apart these possibilities.

A popular model of decision making describes a hierarchical architecture, where values of individual actions (absolute values) are represented in the DMS, and these representations are read out and used by downstream areas to compute the relative-values needed for action selection¹. Although this hypothesis comes from the finding that neurons in DMS (including the anterior- to mid-caudate) encode values of specific actions (absolute value)^{15,16}, our results show that net value representation in the DMS is more enriched than previously thought. One difference between the present and previous studies is task epochs used for the analyses: we focused on the pre-trial initiation period while previous studies focused on the period just preceding action selection. However, this difference alone may not account for the discrepancy. First, our analysis using pre-action selection epochs did not support the dominance of absolute value representation (Supplementary Fig.8c, right). Second, and more importantly, the analysis methods of past studies had biases in the classification of their neuronal responses (Supplementary Figs. 8 and 9). Our analysis using polar representations is not prone to these biases.

Diversity of value coding

Our data show that value coding neurons in the striatum are distributed rather continuously in the polar coordinate (Fig. 7a). Rather than simply forming distinct categories of value coding types, the population seems to encode diverse linear combinations of value. Future studies will be needed to understand how this diversity is generated. One possibility is that relative and net-coding activities are indeed secondarily derived from absolute value representations. It is also possible that net- and relative-value representations in the DMS do not depend on reading out absolute-value representations, and are explicitly or directly represented. For example, tonic dopamine levels may convey net-value information to DMS neurons⁴. It will be crucial to examine how cortical afferents, dopaminergic, and other neurotransmitter systems modulate this diversity of value-modulated responses.

Choosing the general pace of performance in conjunction with what specific action to take is vital for behavioral regulation. By providing a theoretical framework, a behavioral paradigm, and analytical tools, our study promotes a more inclusive understanding of decision making. Applying these approaches to future experiments in different brain regions will further our understanding of how the brain regulates value representation and goal-oriented behavior.

ONLINE METHODS

All procedures involving animals were carried out in accordance with NIH standards and approved by the Harvard University Institutional Animal Care and Use Committee (IACUC). All values were represented by the mean \pm standard error (SE) unless otherwise noted.

Behavior

Fifteen male Long-Evans hooded rats (250–300 grams) were trained to perform an odor discrimination task for water reward²³. Animals were pair-housed under a 12 hour light/dark cycle during the training period (and were individually housed later for electrophysiology and lesion experiments). All experiments were performed during the animal's dark cycle. Rats self-initiated each trial by introducing their snout into a central port, which triggered odor delivery. Valid odor pokes were restricted to trials where animals delivered a nose poke that lasted at least 20 ms and delivered a single nose poke into the odor port. Multiple successive pokes such as two pokes in a row abort the trial and triggers a four second inter-trial interval. Valid odor pokes also must occur outside the inter-trial interval. After a variable delay, drawn from a uniform random distribution of 0.3–0.5 s, a binary mixture of two pure odorants, caproic acid and 1-hexanol, was delivered at one of 4 concentration ratios (100/0, 60/40, 40/6, 0/100) in pseudorandom order within a session. After a variable odor sampling time, rats responded by withdrawing from the central port, which terminated the delivery of odor, and moved to the left or right water port. Choices were rewarded according to the dominant component of the mixture, that is, at the left port for mixtures A/B < 50/50 and at the right port for A/B > 50/50. The next 'trial start' time commences 4 seconds after closure of the water valve. The animal's trial initiation time is defined as the latency it takes for animals to poke back into the odor port after this Trial Start time. Blocks were randomly interleaved within a session, which contained 5 to 9 blocks.

Odor pokes during the inter-trial interval

The task had a forced 4s-time-out between trials, and the analyses focused on valid odor pokes only. However, it is possible that animals poke during this inter-trial interval. However, very few odor pokes occurred before the end of the trial (<20%, odor pokes before

time 0 seconds; Supplementary Fig. 1), and our conclusions are not significantly affected by the inter-trial interval.

Movement time

Under vigor theory, one may predict that all components of actions will speed up during times of high net value⁴. We analyzed movement time (time from odor port exit to reward port) across blocks in individual animals and the population. Although some animals showed faster movement time in high net value blocks (Supplementary Fig. 2a, t-test between high and low value blocks, $t_{28} = 2.8$, $P < 0.05$, $n = 7$ sessions per block type for one rat) this effect was very small and not consistent across the population (Supplementary Fig. 2b, t-test between high and low value blocks, $t_{58} = 0.5$, $P > 0.05$, $n = 15$ rats per block type). In the example rat shown, movement time was fastest at 0.28 seconds in high value blocks and slowest at 0.31 seconds in low value blocks (Supplementary Fig. 2a). Thus, the difference in time between high and low blocks is on the order of tens of milliseconds, possibly because the animal is already moving near maximum speed in high value blocks. Nevertheless, there is a trend for the population to speed up movement time in high value blocks.

Comparing τ to the difference between θ_R and θ_L

A small difference in θ_R and θ_L corresponds to less dependence of trial initiation time on immediately preceding trials and should correspond to larger τ . Conversely, a large difference between θ_R and θ_L ($\sim 90^\circ$) should correspond to smaller τ . To understand the quantitative relationship between τ and $\theta_R - \theta_L$, we performed the following simulation. Using a given value of τ , we predicted trial initiation times based on rat's actual reward history. We then obtained θ_R and θ_L , using the predicted trial initiation times for right reward and left reward trials (Supplementary Fig. 3). This simulation predicted that there is indeed a negative relationship between τ and $\theta_R - \theta_L$, and that $\theta_R - \theta_L$ and τ obtained from the data fall within this prediction (Supplementary Fig. 3, black cross: τ and $\theta_R - \theta_L$, mean and standard error). This result demonstrates that our observations using trial initiation vectors and time constants (regression with reward history) are quantitatively consistent.

Lesions

After 6 weeks of training on the reward-manipulation task, rats were randomly selected to be in one of three conditions: DMS-lesion, VS-lesion, and sham-lesion. The experimenter was not blind to group allocation. Rats were individually housed following surgery. Rats were lesioned with ibotenic acid (250nl, 10mg/ml) in either the DMS (AP 1.68, ML 2.0 and DV 4.5) or VS (AP 1.68, ML 1.5 and DV 7.4). Sham lesioned animals were bilaterally infused with saline of the same volume. We used 5 animals per category, as it was the minimum needed to obtain statistical significance for our analyses. After one week of recovery, animals were water deprived. Behavioral performance for the first 7–8 days after recovery was used for the behavioral analysis. All animals ($n=15$) were perfused and stained with Nissl as described before⁴⁰. Some sections were used for neuron-specific labeling, where NeuN was used as primary antibody (1:200) and Alexa568 (1:500) used as the secondary antibody⁴¹ (Supplementary Fig. 4).

Electrophysiology

Electrophysiological experiments were performed as described before⁴⁰. Briefly, rats were implanted with custom-made microdrives in the left, anteromedial striatum (1.7 mm anterior to bregma and 2.1 mm lateral to the midline). Extracellular recordings were obtained with twelve independently movable tetrodes using the Cheetah system (Neuralynx) and single units were isolated by manually clustering spike features with MClust (A. D. Redish). Cells

were recorded at various depths between 3.5 and 9 mm ventral to the surface of the skull. The boundary between dorsal and ventral striatum was 6 mm deep. The depth of each cell was reconstructed by calculating the number of turns made on each tetrode screw (each turn = 0.32mm) and further confirmed using the final length of each tetrode (through histological examination and measuring the length of the tetrodes after removal of the drive). We recorded 522 neurons, with 364 from DMS and 158 from VS. The sample size is comparable to similar studies in the field and was sufficient for the statistical analyses used in the present study.

Data analysis

Behavior

The polar coordinate analysis of trial initiation time: We developed a method to represent, using a single vector, how behavior and neural activity were modulated by changes across different reward blocks. The first and last 30 trials per session were excluded from the analysis for examining trial initiation, in order to examine steady-state behavior that is independent of satiety. The first 10 trials after each block transition were eliminated to exclude the effects of learning. To obtain a behavioral vector, we regressed each animal's trial initiation time by the reward amounts of the left and right water ports, which varied across blocks. The values of the left and right choices, Q_L and Q_R , were defined by reward amounts (water valve duration; we confirmed the linear relationship between the valve durations and the delivered reward amounts).

$$Trial_Initiation = \beta_0 + \beta_L \cdot Q_L + \beta_R \cdot Q_R$$

We used F-test to see whether behavior significantly modulated by blocks. The F-test tests whether a proposed regression model as a whole fit the data significantly better than a simpler model ($Trial\ initiation = \beta_0$)⁴².

The amplitude (r) of the behavior vector was calculated by taking the square root of the sum of the square of the coefficients:

$$r = \sqrt{(\beta_L)^2 + (\beta_R)^2}$$

the polar angle (θ) was calculated by taking the four-quadrant arc-tangent of the coefficients:

$$\theta = \tan^{-1} \left(\frac{\beta_R}{\beta_L} \right)$$

To determine how trial initiation time was modulated by reward size, we divided the polar plot into 8 segments of 45°. θ values of the behavior vectors falling between -22.5° and $+22.5^\circ$ represents a positive correlation between trial initiation and size of left reward; behavior vectors whose θ fall between 22.5° and 67.5° represents a positive correlation between trial initiation and net-value, θ between 67.5° and 112.5° represents a positive correlation with right-reward, θ between 112.5° and 157.5° represents positive correlation with right > left, θ between 157.5° and 202.5° represents negative correlation with left reward, etc.

To determine whether the two vectors for left and right trials have significantly different angles, we shuffled left and right trials so that we effectively ignore choice direction. We

then obtain trial initiation vectors for the shuffled trials and determined whether their angle difference is significantly different from that derived from the original, un-shuffled data.

Neuronal data analysis

Classification of striatal responses—Our analysis focused on the pre-trial initiation period, the 0–300 ms window before odor poke in. The first 10 trials of every block were not used to eliminate the effect of learning. Next, we regressed each neuron's firing rate by the reward amounts, which varied across blocks.

$$FiringRate = \beta_0 + \beta_L \cdot Q_L + \beta_R \cdot Q_R$$

We used F-test to select for neurons ($P < 0.01$). The F-test tests whether a proposed regression model as a whole fit the data significantly better than a simpler model ($Firing\ rate = \beta_0$). Because this fitting is invariant to the choice of axes, F-test is also invariant to the choice of independent variables.

The amplitude (r) of each neuron was calculated by taking the square root of the sum of the coefficients:

$$r = \sqrt{(\beta_L)^2 + (\beta_R)^2}$$

the polar angle (θ) was calculated by taking the four-quadrant arc-tangent of the coefficients:

$$\theta = \tan^{-1} \left(\frac{\beta_R}{\beta_L} \right)$$

Similar to the behavioral vector classification, we divided the polar plot into 8 segments of 45° to classify the neural responses. Neurons whose θ values fall between -22.5° and $+22.5^\circ$ were classified as left-positive value-coding, neurons whose θ fall between 22.5° and 67.5° were positive-state-value-coding, neurons whose θ fall between 67.5° and 112.5° were right-positive value-coding, neurons whose θ fall between 112.5° and 157.5° were right-preferring relative-value-coding, etc.

More stringent classification of neural responses—Further, we used a more stringent method to classify the responses of our population. By bootstrapping, we obtained a 95% confidence interval on each angle. Here, only neurons whose confidence intervals crossed exactly one classification boundary can be included (each boundary being 0, 45, 90, 135, 180, 225, 270, and 315). With the more stringent classification method, 80/522 neurons were classified as value-coding. Net-value coding neurons were in the majority, comprising 33.8% of the 80 value-coding neurons.

Peri-event time histograms (PETHs)—PETHs in Fig. 6 plot the firing rates of each block condition between odor port entry and water port entry. As the timing between each task event is variable, for visualization purposes, the neural activities were aligned to all events, and 'time-warped,' so that firing rate (spikes/s) is preserved, but the time windows between epochs are kept constant. The timing of each epoch (Odor poke in, odor valve on, etc.) was averaged across all trials. Firing rates between epochs were obtained for every trial and then averaged for the PETH display. Time-warped firing rates were used only for displaying PETHs and not for the analyses below.

Alternative analysis window—We selected all neurons that had their peak firing activity between water-poke-out and odor-poke-in and obtained 181 neurons in total (127/364 in DMS and 54/158 in VS). We performed the above regression analysis and vector projection using a 250ms window around the time of the each neuron's time of peak activity. With this analysis, the dominant neural response in the striatum was net-value coding (Supplementary Fig. 7b, left). DMS was net-value coding in both the positive and negative direction (Supplementary Fig. 7b, middle), while in the VS, was primarily negative-net value coding (Supplementary Fig. 7b, right).

Comparing our proposed vector analysis with previous studies' analyses—

Our finding that the DMS encodes net-value seemingly contradicts with the current framework that the caudate (or DMS) primarily encodes the absolute values of choices, or 'action values'^{15,16}. Previous studies used different animal models (primates vs. rodents), task structures (varying reward value by temporal discounting or reward probability vs. varying reward size), and analysis epochs (after stimulus onset instead of during before trial initiation). However, the most critical differences may be in our analysis methods. In studies that suggested that DMS primarily encodes absolute action values, firing rates were regressed by the independent variables, Q_L and Q_R , representing the 'action value' of left and right, respectively. After running multiple regression (or Mann-Whitney U-test), neurons were classified to be absolute-value coding if the coefficients for *either* Q_L or Q_R , but *not both*, were significantly different from zero^{15–17}. Yet, neurons were relative- (or net-) value coding if coefficients for *both* Q_L and Q_R were significantly different from zero (Supplementary Fig. 8a conventional method type 1). Conversely, other studies that examined the role of VS and other areas in net-value coding applied the opposite regression: $Q_R - Q_L$ and $Q_L + Q_R$ as independent variables, where neurons were considered relative- or net-value coding if *either* $Q_R - Q_L$ or $Q_L + Q_R$ were significant, respectively (conventional method type 2)^{17,43,44}.

We first performed a simulation (Supplementary Fig. 8) to test whether the two different methods can correctly capture a neural population that is uniformly involved in all aspects of value coding, i.e. uniformly distributed in the polar coordinate representation (an 'extreme' case). To match numbers obtained in past striatal studies, we assumed about 36% of the population was statistically significant at the 5% level (for more general cases, Supplementary Fig. 9). Applying conventional method 1, 32% of the population are classified as absolute-value coding, but only 4% are relative- or net-value coding. Therefore, the vast majority (89%) of value coding neurons was classified as absolute-value coding, or participate in valuation. Conversely, by applying conventional method 2 (Supplementary Fig. 9), 89% of significant neurons were relative- or net-value coding, while the rest were absolute-value coding. Therefore, neither choice of axes for the regression was fair at classification for a uniform distribution. We visualize these systematic biases in classification in Supplementary Fig. 8a–b (left panels), by plotting a randomly generated and uniformly distributed neural population distributed in two-dimensional space onto each methods' regression plane (neurons falling into different sectors of the critical regions are color-coded and classified accordingly). By replacing the Cartesian coordinate system with a polar coordinate system, so that each neuron's activity is represented by a polar angle (θ) and amplitude (r), equal numbers of neurons were classified into the four categories (Supplementary Fig. 8c). Therefore, we believe that analyzing neural responses with our polar method is fairer at classifying changes related to the absolute, relative, and net values of the animal's options.

We obtained similar biases when we applied past regression methods onto our own data set. Similarly, we found that with conventional method type 1 (regression with Q_L and Q_R), 74.6% of all significant striatal neurons were classified as absolute-value coding

(Supplementary Fig. 8a, middle right panel). On the contrary, using conventional method type 2, 79.4% of significant neurons were relative or state coding (Supplementary Fig. 8b, middle right panel). Lastly, we examined another time epoch (pre-odor valve on) for the analysis and obtained similar biases with conventional method 1 and 2 (Supplementary Fig. 8, right-most panels).

Therefore, discrepancies between our and previous studies may stem from differences in regression methods.

Simulations: Results of categorization by both conventional methods are sensitive to the strengths of neural responses—In conventional regression methods, the relative proportions of different value coding categories (absolute-value and relative/net coding neurons) are sensitive to the total proportion of statistically significant neurons. To illustrate this, we simulate neural populations that are uniformly distributed in two dimensional space with various standard deviations (SD). When there are no signals in the population, (Supplementary Fig. 9a, a normal distribution with $SD=1$), only chance levels of neurons are selected with a criterion of 5%. In this case, most neurons are significant for a single independent variable (S: single-positive) and very few are significant for *both* independent variables (D: double-positive). On the other hand, in distributions with high SD and thus larger proportions of significant neurons, (Supplementary Fig. 9d, $SD=6.0$), most of these significant neurons are significant for *both* independent variables (D: double-positive) while a smaller fraction are significant for only one independent variable (S: single-positive). Thus, the proportions of neurons classified as single-positive or double positive are sensitive to the SDs of the distributions, or the amount of signal in a population. (Note that with conventional method type 1, single-positive types are absolute-value coding while double-positive types are relative/state coding; the opposite holds for conventional method type 2).

By defining p as the probability of significance for each independent variable, then the probabilities of single-positive, double-positive and total-positive (S+D) types are given as follows (supplementary Fig. 9e):

$$\begin{aligned} \text{Single positive: } & 2p - 2p^2 \\ \text{Double positive: } & p^2 \\ \text{Total significant: } & 2p - p^2 \end{aligned}$$

Therefore, the relative frequency of single-positive points among total positive points is given (supplementary Fig. 9f):

$$\text{Single positive/Total: } \frac{2 - 2p}{2 - p}$$

Therefore when $p = 0.20$ (20% of the neurons are significantly modulated by either variable), the distribution is about 1.53 SD, so 88.9% of significant neurons are classified as single-positive while only 11.1% are classified as double-positive. (In this case, about 36% of neurons in total are deemed 'significant', which is close to the percentage obtained in the present and previous studies).

Contrary to the conventional methods, our method (the polar method) is invariant to overall proportions of statistically significant neurons.

Our proposed method is invariant to the choice of axes for regression—We applied Q_L and Q_R as independent variables for multiple regression. With our proposed method of categorization, 32, 32, 30, and 55 neurons out of 522 in total were left absolute, right absolute, relative, and net value coding, respectively. When we applied Q_L+Q_R and Q_R-Q_L as independent variables, we achieved exactly the same results (to categorize the neurons, all categories are simply shifted +45 degrees). This confirms that our method is invariant to the choice of independent variables (axes).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank J. Assad, K. Blum, O. Hikosaka, D. Lee, M. Livingstone, J. Maunsell, and M. Meister for their comments on the manuscript. We also thank R. Born and the members of the Uchida lab for discussions. We thank Scott Kim for support in behavioral experiments. This work was supported by a fellowship from National Science Foundation (A.Y.W), PRESTO (K.M.), a Smith Family New Investigator Award, the Alfred Sloan Foundation, the Milton Fund and the Startup Fund from Harvard University (N.U.).

REFERENCE

1. Kable JW, Glimcher PW. The neurobiology of decision: consensus and controversy. *Neuron*. 2009; 63:733–45. [PubMed: 19778504]
2. Lee D, Seo H, Jung MW. Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci*. 2012; 35:287–308. [PubMed: 22462543]
3. Niv Y, Joel D, Dayan P. A normative perspective on motivation. *Trends Cogn Sci*. 2006; 10:375–81. [PubMed: 16843041]
4. Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. 2007; 191:507–20. [PubMed: 17031711]
5. Guitart-Masip M, Beierholm UR, Dolan R, Duzel E, Dayan P. Vigor in the face of fluctuating rates of reward: an experimental examination. *J Cogn Neurosci*. 2011; 23:3933–8. [PubMed: 21736459]
6. Hebb DO. Drives and the C.N.S. (conceptual nervous system). *Psychol Rev*. 1955; 62:243–54. [PubMed: 14395368]
7. Hull, CL. *Principles of Behavior: An Introduction to Behavior Theory*. D. Appleton-Century Company, Inc.; Oxford, England: 1943.
8. Ferster, CB.; Skinner, BF. *Schedules of reinforcement*. D. Appleton-Century Company, Inc.; Oxford, England: 1957.
9. Nakamura K, Hikosaka O. Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *J Neurosci*. 2006; 26:5360–9. [PubMed: 16707788]
10. Ding L, Gold JJ. Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. *Neuron*. 2012; 75:865–74. [PubMed: 22958826]
11. Roesch MR, Olson CR. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science*. 2004; 304:307–10. [PubMed: 15073380]
12. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol Cybern*. 2001; 84:401–10. [PubMed: 11417052]
13. Hikosaka O, Nakamura K, Nakahara H. Basal ganglia orient eyes to reward. *J Neurophysiol*. 2006; 95:567–84. [PubMed: 16424448]
14. Turner RS, Desmurget M. Basal ganglia contributions to motor control: a vigorous tutor. *Curr Opin Neurobiol*. 2010; 20:704–16. [PubMed: 20850966]
15. Lau B, Glimcher PW. Value representations in the primate striatum during matching behavior. *Neuron*. 2008; 58:451–63. [PubMed: 18466754]
16. Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science*. 2005; 310:1337–40. [PubMed: 16311337]

17. Cai X, Kim S, Lee D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron*. 2011; 69:170–82. [PubMed: 21220107]
18. Kelley AE. Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. *Neurosci Biobehav Rev*. 2004; 27:765–76. [PubMed: 15019426]
19. Mogenson GJ, Jones DL, Yim CY. From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol*. 1980; 14:69–97. [PubMed: 6999537]
20. Palmiter RD. Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice. *Ann N Y Acad Sci*. 2008; 1129:35–46. [PubMed: 18591467]
21. Amalric M, Koob GF. Depletion of dopamine in the caudate nucleus but not in nucleus accumbens impairs reaction-time performance in rats. *J Neurosci*. 1987; 7:2129–34. [PubMed: 3112323]
22. Hauber W, Schmidt WJ. Differential effects of lesions of the dorsomedial and dorsolateral caudate-putamen on reaction time performance in rats. *Behav Brain Res*. 1994; 60:211–5. [PubMed: 8003250]
23. Uchida N, Mainen ZF. Speed and accuracy of olfactory discrimination in the rat. *Nat Neurosci*. 2003; 6:1224–9. [PubMed: 14566341]
24. Rorie AE, Gao J, McClelland JL, Newsome WT. Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS ONE*. 2010; 5:e9308. [PubMed: 20174574]
25. Skinner, BF. *The behavior of organisms: an experimental analysis*. D. Appleton-Century Company, Inc.; Oxford, England: 1938.
26. Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci*. 2012; 15:1281–9. [PubMed: 22902719]
27. Mogenson GJ, Takigawa M, Robertson A, Wu M. Self-stimulation of the nucleus accumbens and ventral tegmental area of Tsai attenuated by microinjections of spiroperidol into the nucleus accumbens. *Brain Res*. 1979; 171:247–59. [PubMed: 572734]
28. Imperato A, Di Chiara G. Preferential stimulation of dopamine release in the nucleus accumbens of freely moving rats by ethanol. *J Pharmacol Exp Ther*. 1986; 239:219–28. [PubMed: 3761194]
29. Salamone JD, Correa M, Farrar AM, Nunes EJ, Pardo M. Dopamine, behavioral economics, and effort. *Front Behav Neurosci*. 2009; 3:13. [PubMed: 19826615]
30. Newman R, Winans SS. An experimental study of the ventral striatum of the golden hamster. I. Neuronal connections of the nucleus accumbens. *J Comp Neurol*. 1980; 191:167–92. [PubMed: 7410590]
31. Delgado MR, Stenger VA, Fiez JA. Motivation-dependent responses in the human caudate nucleus. *Cereb Cortex*. 2004; 14:1022–30. [PubMed: 15115748]
32. Volkow ND, et al. "Nonhedonic" food motivation in humans involves dopamine in the dorsal striatum and methylphenidate amplifies this effect. *Synapse*. 2002; 44:175–80. [PubMed: 11954049]
33. Brown VJ, Robbins TW. Elementary processes of response selection mediated by distinct regions of the striatum. *J Neurosci*. 1989; 9:3760–5. [PubMed: 2585053]
34. Bailey KR, Mair RG. The role of striatum in initiation and execution of learned action sequences in rats. *J Neurosci*. 2006; 26:1016–25. [PubMed: 16421321]
35. Thorn CA, Atallah H, Howe M, Graybiel AM. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*. 2010; 66:781–95. [PubMed: 20547134]
36. Kimchi EY, Laubach M. Dynamic encoding of action selection by the medial striatum. *J Neurosci*. 2009; 29:3148–59. [PubMed: 19279252]
37. Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010; 35:48–69. [PubMed: 19776734]
38. Yin HH, Knowlton BJ, Balleine BW. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci*. 2005; 22:505–12. [PubMed: 16045503]

39. Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci.* 2005; 22:513–23. [PubMed: 16045504]
40. Cury KM, Uchida N. Robust odor coding via inhalation-coupled transient activity in the mammalian olfactory bulb. *Neuron.* 2010; 68:570–85. [PubMed: 21040855]
41. Jongen-Relo AL, Feldon J. Specific neuronal protein: a new tool for histological evaluation of excitotoxic lesions. *Physiol Behav.* 2002; 76:449–56. [PubMed: 12126979]
42. Draper, NR.; Smith, H. *Applied Regression Analysis.* John Wiley & Sons, Inc.; New York: 1966.
43. Ito M, Doya K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J Neurosci.* 2009; 29:9861–74. [PubMed: 19657038]
44. Seo H, Lee D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci.* 2007; 27:8366–77. [PubMed: 17670983]

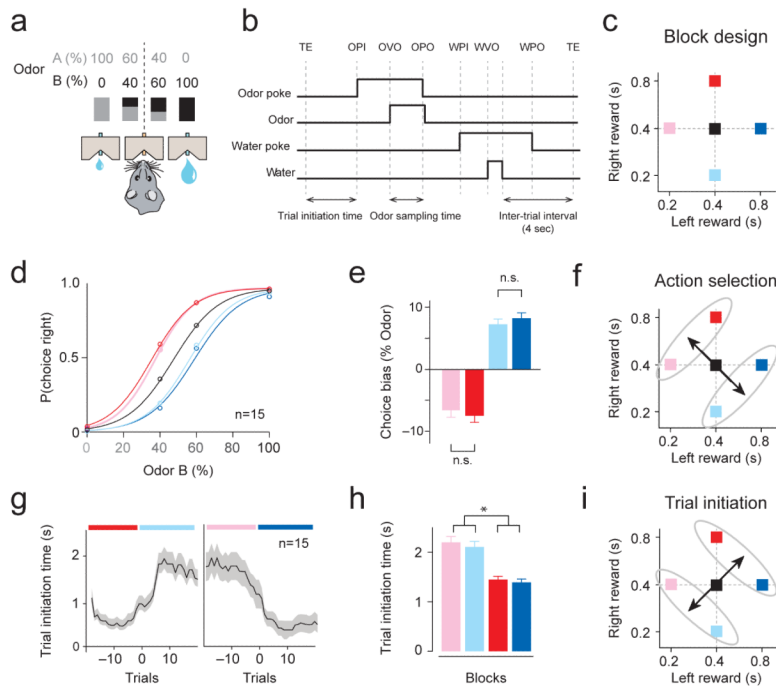


Figure 1. A self-paced decision making task design: dissociating the effects of relative and net value on choice and performance vigor

(a) Odor mixture categorization task. Rats were trained to respond to the left or right reward port depending on the dominant component of a binary mixture of odors. Concurrently, the sizes of reward at the left and right reward ports were manipulated in a block-wise manner.

(b) Task epochs. Rats initiated a trial by making a nose poke into the central odor port. TE, trial end; OPI, odor poke in (entry); OVO, odor valve on (opening); OPO, odor poke out (exit); WPI, water poke in; WVO, water valve on (opening); WPO, water poke out (exit).

(c) Design of block-wise reward manipulations. Reward values of the left and right ports varied across blocks. Each square represents a block of trials. The size of reward was manipulated by changing the duration of water valve opening.

(d) Psychometric functions in different blocks. The average performance over 15 rats is shown. Colors refer to the blocks in **Fig. 1c**.

(e) The magnitude of shift in psychometric curves in bias blocks compared to no-bias blocks. Mean \pm s.e.m. ($n = 15$ rats).

(f) Relative value of options guides action selection.

(g) Time course of alternations in trial initiation time across block changes from the population ($n=15$, mean \pm s.e.m., smoothed over 5 trials). Gradual shifts preceding a block switch (Trial = 0) are due to smoothing of the data. Colors refer to the blocks in the **Fig. 1c**.

(h) Trial initiation time (mean \pm s.e.m. of the median trial initiation time per block).

* $P < 0.05$ (t-test).

(i) Net value of options guides trial initiation time.

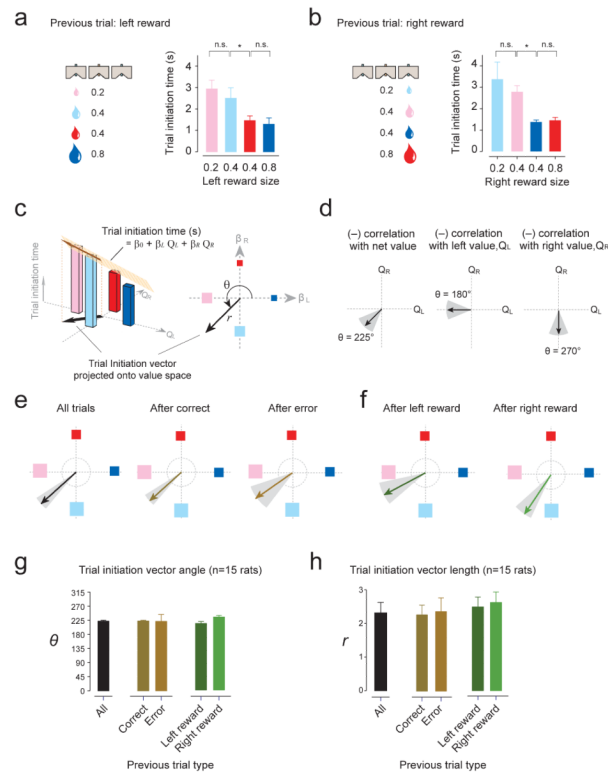


Figure 2. The net value of the animal's current block guides performance vigor

(a–b) Trial initiation time across blocks immediately after left (a) or right reward (b) from an example animal. Colors refer to the blocks in the Fig. 1c.

(c) Trial initiation vector analysis. Trial initiation times (bar graphs) were regressed by left and right reward values (regression equation in black). The vector (brown arrow) represents the relative contribution of left and right value (Q_L and Q_R) on trial initiation time and was projected onto the 'value space' (the black arrow in both panels in 2c represents the same projected vector). The vector (black arrows in either panel of 2c) represents the direction (θ) and magnitude (r) of modulation in value space based on the regression coefficients (right). Essentially, the vector illustrates the strength and the relative contribution of left and right value on trial initiation time.

(d) Categorization of trial initiation vectors based on the angle of the vector in value space. Trial initiation time can be correlated, for example, with net value (left) or the value of one of the options (middle and right).

(e–f) Example vectors of a single rat. Vectors were calculated separately for trial initiation times immediately after correct, error, or pooled trials (e), or after left or right reward trials (f). Grey shading indicates the 95% confidence interval of the vector angle obtained by bootstrap method.

(g) Population summary of the angles (θ) of vectors (mean \pm s.e.m., $n = 15$ rats).

(h) Population summary of the amplitudes (r) of vectors (mean \pm s.e.m., $n = 15$ rats).

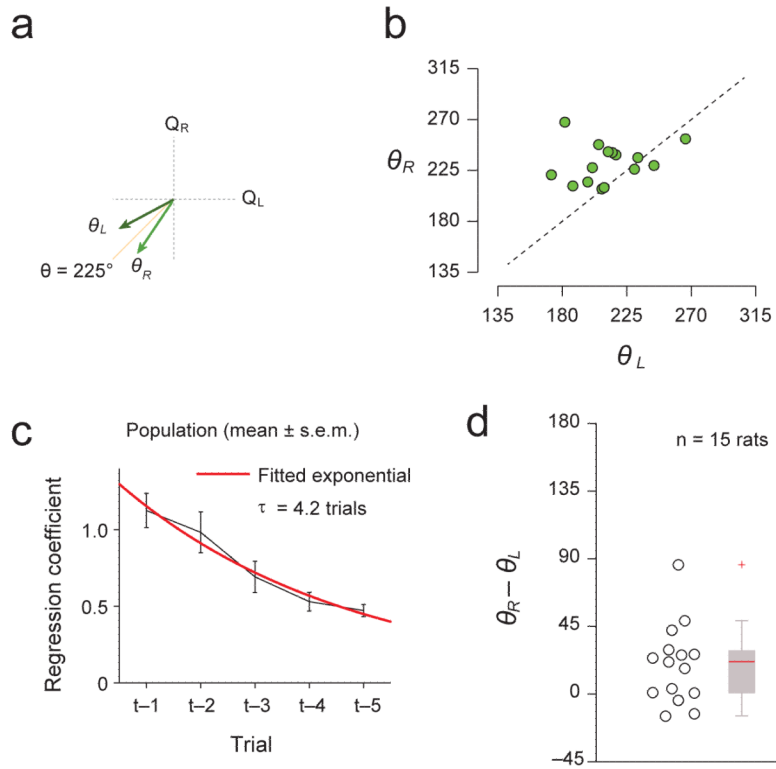


Figure 3. Performance vigor depends on the integration of outcomes of multiple previous trials
(a) Schematic showing how $\theta_R - \theta_L$ is computed. θ_L and θ_R : the angle of trial initiation vectors after left or right reward trials, respectively.
(b) Scatter plot of θ_R versus θ_L .
(c) Effect of multiple previous trial outcomes on trial initiation time for the population. Black: regression coefficients (mean \pm S.D., n = 15 rats), Red: fitted exponential curve (time constant: $\tau = 4.24$ trials).
(d) Difference in angle between vectors representing trial initiation time immediately after left reward and after right reward trials, ($\theta_R - \theta_L$). The lower dotted line represents the median angle of trial-shuffled controls, which is the expected angle when the two vectors did not differ but has finite sample size. The upper dotted line represents the expected difference in angle if initiation time depended only on previous reward size.

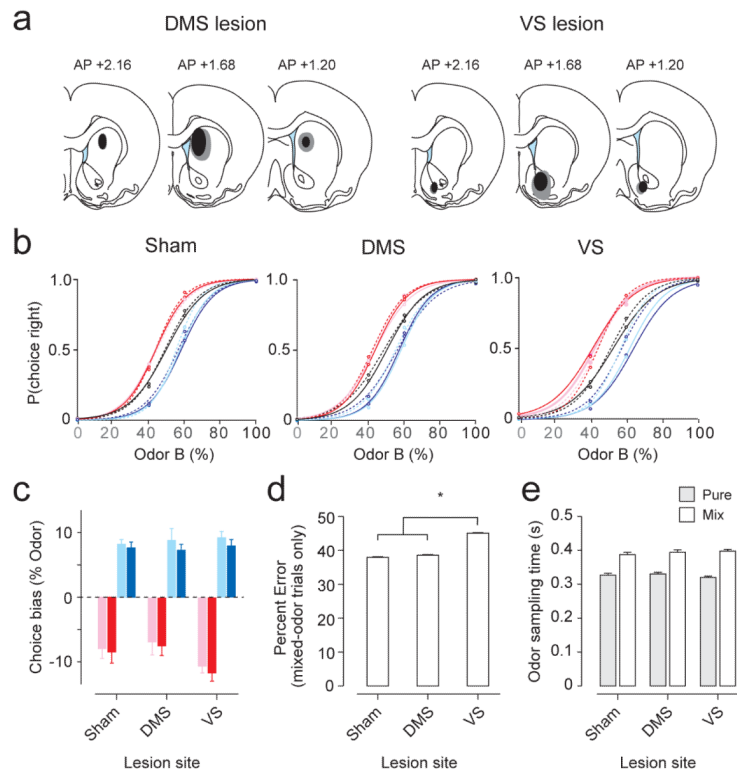


Figure 4. Lesions of DMS and VS have little effect on action selection

(a) Locations of selective lesions of the anterior DMS and VS. Gray area, maximum extent of lesion. Black area, minimum extent of lesion.

(b) Psychometric curves of different lesion conditions. Dotted lines indicate pre-lesion performance; solid lines indicate post-lesion performance (average performance over 15 rats).

(c) Quantification of the shifts in psychometric curves for all reward-bias blocks in all conditions. Colors represent reward-bias blocks shown in Fig. 1c. * $P < 0.05$ (t-test). We obtained the inflection points of psychometric curves for each block and animal. Leftward or rightward shifts in psychometric curves were calculated by subtracting the inflection points of reward-bias blocks (colored curves) from blocks with no reward bias (black curves, **Fig. 4b**).

(d) Performance accuracy for mixed-odor trials (mean \pm s.e.m., $n = 15$ rats).

(e) Odor sampling time for pure- and mix-odor trials (mean \pm s.e.m., $n = 15$ rats).

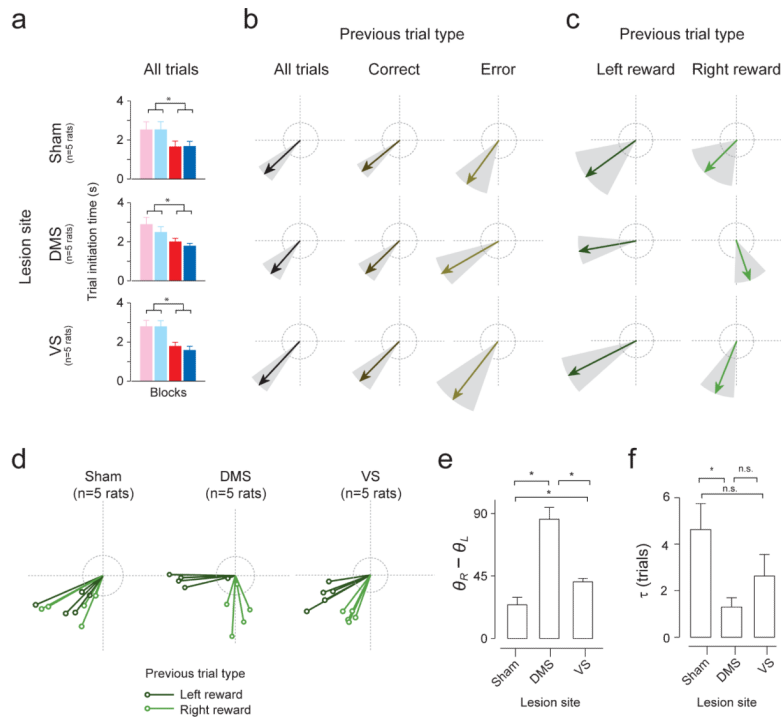


Figure 5. Lesions in DMS impaired net value-dependent modulation of vigor

(a) Trial initiation time across blocks (pooled for all trials) in different lesion conditions (mean \pm s.e.m., $n = 5$ rats for each condition).

(b–c) Trial initiation vectors for one rat separated by previous trial outcome (correct, error, left reward and right reward). Grey, 95% confidence interval. Dotted rings represent a scale of vector length ($r=1$).

(d) Trial initiation vectors for all animals, separated by previous left and right reward outcome. $n = 5$ rats for each condition.

(e) Difference in polar angle between left and right trial vectors of each animal ($*P < 0.05$, t -test, $n = 5$ rats for each condition).

(f) Effects of multiple previous trial outcomes on trial initiation time. Time constants are derived from exponential curves fitted to coefficients of multiple regression (mean \pm s.e.m., $n = 5$ rats for each condition).

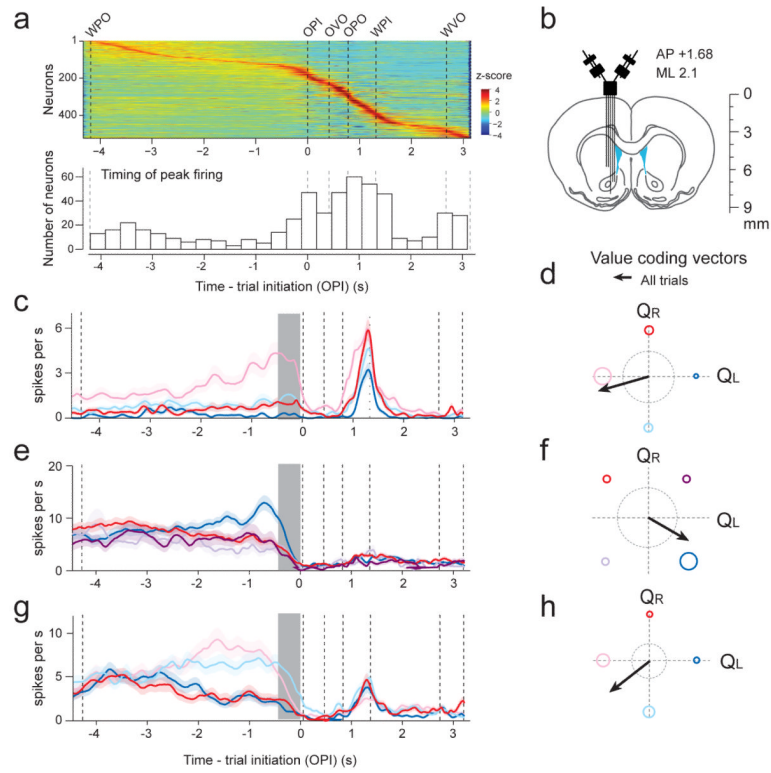


Figure 6. Value-coding neurons in the rat striatum

(a) Top, Peri-event time histograms (PETHs) of all neurons recorded. Neurons are sorted by time of peak firing (z-scored). Bottom, histogram of time of peak firing (n=522 neurons).

(b) Recording configuration.

(c–e) PETHs of single neurons. Colors refer to blocks (see **Fig. 6f–h** or Fig. 1c and Supplementary Fig.6).

(f–h) Vector representation of firing rate modulation during pre-trial initiation period (f, g, and h correspond to PETHs in c, d, and e, respectively). Dotted rings represent a scale of vector length ($r=4$).

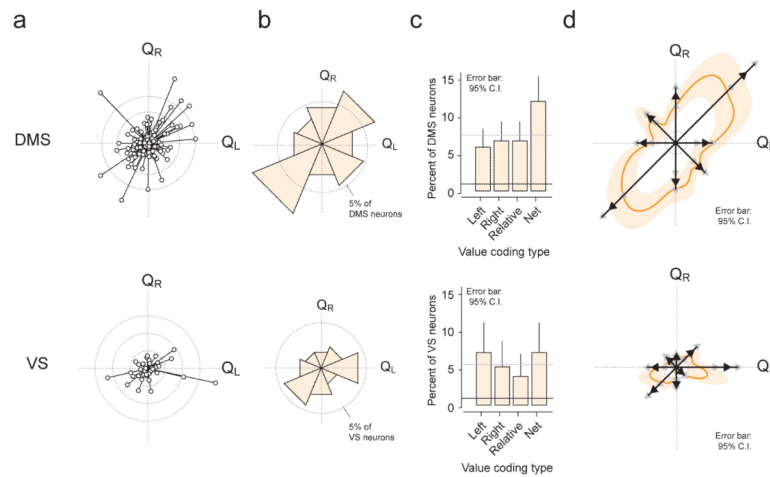


Figure 7. DMS, rather than VS, predominantly encodes net value

(a–c) Top, DMS value coding neurons; Bottom, VS value coding neurons

(a) Population data: vector representation of all value coding neurons ($P < 0.01$, $n = 149$ out of 522 neurons, F-test in regression, see text). Each circle represents one neuron. Inner dotted ring: $r=8$; middle ring, $r=16$; outer ring, $r=24$.

(b) Circular histogram depicting the proportion of neurons in each value coding category. Dotted ring, 5%.

(c) Distribution of neurons encoding value of one action (left and right), relative value, and net value. Mean \pm 95% confidence interval (bootstrap).

(d) Smoothed vector representations (thick orange line). Population vectors (Fig. 7a) were smoothed using a Von Mises distribution (κ , 20). Mean \pm 95% confidence interval (bootstrap). (thick orange line and light shading, respectively). Vectors: resultant vectors of each value coding type (black arrowheads). Grey arrow heads represent the boundaries of the 95% confidence interval.

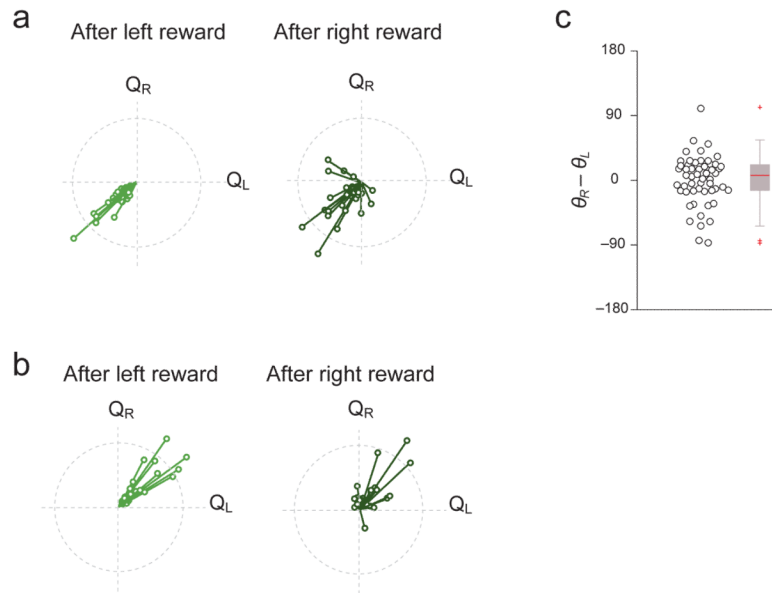


Figure 8. The outcome of immediately preceding trial does not significantly affect the activity of net value coding neurons

(a) Neurons having negative correlation with net value: vectors representing direction of value modulation are plotted separately for trials following left and right reward.

(b) Same as **Fig. 8a** for neurons having positive correlation with net value.

(c) Difference in angle between vectors for firing rate after left and right reward trials ($\theta_R - \theta_L$). Each circle represent one value-coding neuron ($n = 54$ neurons). In the box plot (right), the central mark indicates the median, and the edges of the box are 25th and 75th percentiles.