

Neural dynamics underlying self-control in the primate subthalamic nucleus

Benjamin Pasquereau^{1,2} and Robert S. Turner³

¹ Institut des Sciences Cognitives Marc Jeannerod, UMR 5229, Centre National de la Recherche Scientifique, 69675 Bron Cedex, France.

² Université Claude Bernard Lyon 1, 69100 Villeurbanne, France

³ Department of Neurobiology, Center for Neuroscience and The Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

Address correspondence to:

Benjamin Pasquereau, PhD
Institut des Sciences Cognitives Marc Jeannerod
CNRS UMR-5229
67 Boulevard Pinel
69675 BRON Cedex, France
Phone: +33 437 911 241
benjamin.pasquereau@isc.cnrs.fr

36 **ABSTRACT**

37 The subthalamic nucleus (STN) is hypothesized to play a central role in neural processes that regulate
38 self-control. Still uncertain, however, is how that brain structure participates in the dynamically
39 evolving estimation of value that underlies the ability to delay gratification and wait patiently for a
40 gain. To address that gap in knowledge, we studied the spiking activity of neurons in the STN of
41 monkeys during a task in which animals were required to remain motionless for varying periods of
42 time in order to obtain food reward. At the single-neuron and population levels, we found a cost-
43 benefit integration between the desirability of the expected reward and the imposed delay to reward
44 delivery, with STN signals that dynamically combined both attributes of the reward to form a single
45 integrated estimate of value. This neural encoding of subjective value evolved dynamically across the
46 waiting period that intervened after instruction cue. Moreover, this encoding was distributed
47 inhomogeneously along the antero-posterior axis of the STN such that the most dorso-posterior-
48 placed neurons represented the temporal discounted value most strongly. These findings highlight
49 the selective involvement of the dorso-posterior STN in the representation of temporally discounted
50 rewards. The combination of rewards and time delays into an integrated representation is essential
51 for self-control, the promotion of goal pursuit and the willingness to bear the costs of time delays.

52

53

54

55

56

57

58 INTRODUCTION

59 Imagine you are standing in a queue in front of a bakery. How long are you willing to wait for your
60 favorite pastry? Many of us lose patience after about five minutes, while others persevere and keep
61 waiting during longer delays. Our individual ability to delay gratification and maintain self-control
62 depends on an internal process that estimates continuously the trade-off between the desirability of
63 the benefit expected and the cost of waiting (Ainslie, 1975). All animals, including humans, prefer to
64 receive rewards sooner rather than later, a phenomenon known as temporal discounting (Frederick
65 et al., 2002; Loewenstein and Prelec, 1993; Mazur, 2001; Vanderveldt et al., 2016). Accordingly,
66 people with low discount rates tend to pursue their long-term goals patiently, whereas people with
67 high discount rates often abandon their goals impulsively and move on (Janakiraman et al., 2011). In
68 economic behavior, the net payoff for such a cost-benefit dilemma is typically evaluated by
69 integrating the magnitude of the future reward with a hyperbolic discounting function (Green and
70 Myerson, 2004; Kirby, 1997; Loewenstein et al., 1992). Most studies of the neural correlates of
71 temporal discounting have focused on the task instruction period, the point in a trial when the
72 subject is informed of the size of the reward to be delivered and of the delay in time until its delivery
73 (Berns et al., 2007). Far less is known about neuronal activity during the subsequent post-instruction
74 delay period, during which subjects may exhibit varying degrees of patience (e.g., self-control) and
75 anticipation of reward. It is quite possible that neuronal activity related to those factors also evolves
76 dynamically across this time period. For example, as the subjective value of a future reward is
77 updated with the passage of time, the motivation to achieve a delayed goal may vary gradually.
78 Indeed, functional imaging studies in humans suggest that neural activity related to temporal
79 discounting evolves dynamically during a post-instruction delay period in patterns that differ
80 distinctly between brain regions (Jimura et al., 2013; McGuire and Kable, 2015; Tanaka et al., 2020).
81 How such dynamically evolving encodings of temporally-discounted subjective value are instantiated
82 at the single-unit level remains poorly understood.

83

Because the subthalamic nucleus (STN) is thought to be crucial in inhibitory control by preventing impulsivity (Aron et al., 2016; Bonnevie and Zaghloul, 2019; Jahanshahi et al., 2015) and modulating the performance of reward-seeking actions (Baunez et al., 2007; Baunez and Robbins, 1997), we hypothesized that this structure could contribute to the maintenance of adaptive behaviors by dynamically computing the temporally discounted value. The STN occupies a unique position for translating motivational drives into behavioral perseverance, standing at the crossroads between the basal ganglia indirect pathway and many hyperdirect inputs from prefrontal areas involved in motivational, cognitive and motor functions (Haynes and Haber, 2013; Parent and Hazrati, 1995). Current functional models of the STN propose that increased activity in the STN extends the time to action initiation by elevating decision thresholds, preventing suboptimal early responses or decisions, especially in situations in which the motivational options are conflicting (Cavanagh et al., 2011; Frank, 2006; Mansfield et al., 2011). In support of these models, a series of lesion studies performed on rats provided causal evidence that STN restrains premature responding in instrumental tasks (Baunez and Robbins, 1997; Wiener et al., 2008) and controls the willingness to work for food (Baunez et al., 2005, 2002). Dysfunctions of STN circuits even produced perseverative actions with a reduced ability to switch between behaviors (Baker and Ragozzino, 2014; Baunez et al., 2007), making this brain region a good candidate for regulating self-control and delayed gratification. Until now, however, existing evidence is mixed on whether the STN is involved in temporal discounting (Aiello et al., 2019; Evens et al., 2015; Seinstra et al., 2016; Seymour et al., 2016; Uslaner and Robinson, 2006; Voon et al., 2017; Winstanley et al., 2005), and no previous study has investigated how STN neurons process value information across delays.

Aside from its role in motor control, clinical studies support the involvement of the STN in motivational functions. In particular, deep brain stimulation (DBS-STN), which is effective at alleviating motor symptoms in parkinsonian patients, may induce a variety of side effects related to altered motivation such as depression, excessive eating behavior and hypomania (Berney et al.,

2002; Castrioto et al., 2014; Jahanshahi et al., 2015; Voon et al., 2006). Electrophysiological recordings collected from the STN of these patients have shown low-frequency oscillations (<12Hz) and spiking activities related to various aspects of reward processing, with neural signals modulated by the magnitude of monetary reward and cost-benefit value attribution (Fumagalli et al., 2015; Justin Rossi et al., 2017; Zénon et al., 2016). In non-human animals, the ability of the STN to represent the subjective desirability of actions has also been evidenced by studies that show neurons firing as a function of the expected reward and the associated effort cost (Breyse et al., 2015; Espinosa-Parrilla et al., 2013; Nougaret et al., 2022). Although substantial effort has been directed to elucidate the role of the STN in valuation-related processes at the time of decision-making (Cavanagh et al., 2011; Coulthard et al., 2012; Frank et al., 2007) and in movement incentive (Nougaret et al., 2022; Tan et al., 2015; Zénon et al., 2016), much less attention has been paid to STN involvement in the computation of temporally discounted value during a waiting period, when behavioral inhibition must be sustained patiently over time. In addition, it is still unclear how these roles for STN in cost-benefit valuation and motivational processing relate to the known organization of this nucleus into anatomically and functionally distinct territories (Alexander et al., 1990; Nambu et al., 2002; Parent and Hazrati, 1995).

To determine if the STN conveys signals consistent with its predicted role in pursuing delayed gratification, we trained two monkeys to perform a delayed reward task in which animals were required to remain motionless during post-instruction delay periods of varying durations in order to obtain food reward. We hypothesized that STN neurons exhibit a dynamic encoding of temporally discounted value over the time course of the delay period consistent with a continuously evolving value attribution essential for self-control. Here we tested this hypothesis by studying spiking activity in the STN while monkeys performed the task. At the single-neuron and population levels, our results support a role for the STN in temporal discounting, and indicate that neural signals underlying the valuation of reward size and delay are integrated dynamically into subjective value along an antero-

posterior axis in this nucleus. Such dynamic value integration through the STN may regulate the expression of persistent behaviors for which a continuously evolving cost-benefit estimation is required to monitor and sustain goal achievement.

RESULTS

Monkeys' behavior reflects reward size and delay in an integrated manner

Two monkeys (H and C) were trained to perform a delayed reward task in which they were required to align a cursor on a visual target and to maintain this arm posture for varying periods of time before delivery of food rewards (Fig. 1A). At the beginning of each trial, an instruction cue appeared transiently signaling one of six possible reward contingencies. Cue colors indicated the size of reward (1, 2, or 3 drops of food) and symbols indicated the delay-to-reward [short delay (3.5-5.6s) or long delay (5.2-7.3s)]. Animals were given the option to reject a proposed trial by moving the cursor outside of the target (e.g., if they did not think it was worth waiting for the expected quantity of reward). In this task, the rejection rate (i.e. the proportion of trials with a failure to keep the cursor in the target) reflects the monkey's motivation to stay engaged in the task and to successfully complete the trial according to its prediction about the forthcoming reward. The six instruction cues effectively communicated six different levels of motivation or subjective value as evidenced by consistent effects on the animals' task performance (Fig. 1B-C). Rejection rates were affected by both reward size (two-way ANOVAs; Monkey H: $F_{(2,666)}=10.47$ $P<0.001$; Monkey C: $F_{(2,708)}=5.36$ $P=0.0049$) and delay to reward (H: $F_{(1,666)}=22.62$ $P<0.001$; C: $F_{(1,708)}=8.03$ $P=0.0047$). Although the total proportion of rejected trials differed across monkeys (two-sample t-test; $t_{(229)}=3.96$ $P<0.001$), a similar behavioral pattern was observed in both animals during the task. The proportion of rejected trials was higher for smaller rewards and longer delays, while both animals waited more patiently to obtain larger rewards.

Interactive effects between reward size and delay (H: $F_{(2,666)}=19.31$ $P<0.001$; C: $F_{(2,708)}=10.31$ $P<0.001$) revealed an integration of both task parameters to estimate the overall desirability or subjective value of each cost-benefit condition. To characterize how subjective value declined with delay, we fitted the averaged rejection rates to a hyperbolic discounting model (as expressed by Eq. 2). To be more specific, we inferred the temporal discount factor (k) that maximized the inverse relation between each monkey's behavior and the subjective value calculated from a hyperbolic function. Consistent with other monkey studies (Hori et al., 2021; Minamimoto et al., 2009), the animals' task performance was well approximated by an inverse relation with hyperbolic delay discounting (H: $R^2=0.98$; C: $R^2=0.86$; Figs 1D-E). The resulting discount rates calculated for the two animals were relatively similar in value (Figs 1F-G). In comparison, however, Monkey C was a bit more impatient with a steeper delay discounting ($k=1.62\text{ s}^{-1}$), while the subjective value estimated by monkey H was slightly less impacted by the cost of waiting ($k=1.28\text{ s}^{-1}$).

Controls

We recorded EMGs from different muscles (trapezius, deltoid, pectoralis, triceps, biceps) while monkey H performed the behavioral task. During the post-instruction waiting interval, when the animal remained static, the maintenance of the arm posture resulted in a slight increase in the tonic activity of shoulder muscles (Figs 1H-I). As evidenced by a series of two-way ANOVAs (reward x delay, $P<0.05/173$ -time bins), muscle patterns were not altered by reward contingencies. This suggests that monkeys controlled their posture with a constant motor output across trial conditions, independent of reward size and delay. Alternatively, as monkeys were not required to control their gaze while performing the task, we found that their eye positions varied according to the type of trial (Fig.1-figure supplement 1). Eye position was affected by both the expected reward size and delay after the presentation of instruction cues (two-way ANOVAs, $P<0.05/173$ -time bins). Reward-by-delay interactions detected in eye position after instruction offset reinforce the view that cost-benefit parameters were integrated into a common valuation by monkeys.

186

187 To confirm the ability of our animals to recognize and evaluate appropriately the different instruction
188 cues, the animals also performed a variant of the task that required decision making. In this variant,
189 the monkey was allowed to choose freely between two alternate reward size or delay conditions.
190 We observed appropriately strong preferences for the cues that predicted large rewards and short
191 delays. Monkeys selected the more advantageous option in terms of reward when the delays were
192 equal (H: 97%, $t_{(14)}=26.99$ $P<0.001$; C: 99%, $t_{(16)}=30.55$ $P<0.001$) and the more advantageous delay
193 option when reward sizes were held constant (H: 99%, $t_{(11)}=29.71$ $P<0.001$; C: 95%, $t_{(10)}=24.82$
194 $P<0.001$).

195

196 **Neuronal activity of STN reflects reward size and delay**

197 While the monkeys performed the delayed reward task, we recorded single-unit activity from 231
198 neurons in the right STN (112 from monkey H; 119 from monkey C). Similar to our previous study
199 (Pasquereau and Turner, 2017), STN neurons were identified based on location and standard
200 electrophysiological criteria (Figs 2A-C). Most STN neurons exhibited changes in firing rate at one or
201 more times in the task. Approximately 42% of neurons demonstrated a peak in activity in the first
202 second following presentation of the instruction cues, while 32% of neurons exhibited highest
203 discharge rates later during the waiting period (Fig. 2D). Despite the fact that phasic changes evoked
204 by instruction cues dominated the population-averaged activity (Fig. 2E), we found that the
205 variability of neuronal activities across the six reward/delay conditions was maintained at an
206 elevated constant level across several seconds of the trial, as evidenced by the Fano Factor shown in
207 Figure 2F. This suggests that task-relevant information was processed by STN neurons not only
208 immediately after the presentation of the instruction cue, but also later in the course of the post-
209 instruction delay period, when the animal was maintaining a stable arm position in anticipation of
210 reward delivery.

211

To determine whether and when STN neurons were involved in the evaluation of different task conditions, we tested the neural activities for effects of reward size and delay using two-way ANOVAs combined with a sliding window procedure ($P < 0.05/174$ -time bins). Because the variability of neuronal activities across task conditions was sustained over time, we analyzed the spiking activity of each neuron in a time-resolved way across a continuous 3.5 second period following the presentation of the instruction cue. Of the 231 neurons recorded, 112 (48%; 21 from monkey H and 91 from monkey C) and 91 (39%; 18 from monkey H and 73 from monkey C) were modulated by reward size and delay, respectively (Figs. 2G-J). Interestingly, the two types of encoding occurred preferentially during different periods of the trial. Immediately following cue presentation, neurons were strongly influenced by the reward size signaled by the instruction while, later in the trial, as animals endured the waiting period, encoding of delay became more common. Among the 79 neurons (34%; 14 from monkey H and 65 from monkey C) sensitive to both parameters at some point over the course of the trial, 50 (22%; 8 from monkey H and 42 from monkey C) were influenced simultaneously by reward size and delay within the same time bins, thereby reflecting a direct integration of cost-benefit conditions by individual neurons (Figs. 2K-L). Reward-by-delay interactions were scattered in a roughly uniform distribution across the course of the post-instruction period. Overall, these results suggest that the way STN neurons represented task conditions evolved dynamically across the course of a trial.

Dynamic encodings of reward size and delay

To examine how reward size and delay were encoded by individual STN units and how that encoding changed across time in a trial, we performed time-resolved linear regressions with single-unit neural activity as the dependent variable. For each task-related neuron (i.e., neurons encoding at least one task parameter for a least one-time bin, $n=124$), we tested whether the firing rate was modulated by the expected reward quantity and the delay to reward delivery (as expressed by Eq. 3). Because the STN contains an oculomotor territory (Matsumura et al., 1992), we included measures of eye

movements (i.e., gaze position and gaze velocity) in our model as nuisance variables. (Exclusion of eye parameters from this analysis produced very similar results – Fig.4 – figure supplement 1). As illustrated in Figure 3 with three example neurons, task parameters were encoded in the STN at different stages of the trial following different modalities. (Thresholds for significant regression coefficients were calculated relative to their values during the pre-instruction period using a one sample t -test, $df = 46$, $P < 0.05$.) Based on the polarity of the regression coefficients β_{Reward} , we found neurons whose activity transiently indexed reward size by increasing (e.g., neuron #1) or decreasing (e.g., neuron #2) their firing rate. Similarly, by detecting changes in the regression coefficients β_{Delay} , we found neurons that increased (e.g., neuron #2) or decreased (e.g., neuron #1) their activity as a function of the delay to reward. Neural activities were often influenced in opposite directions by the predicted amount of reward and the delay (positive β_{Reward} with negative β_{Delay} , or vice versa). The specific pattern of task encoding within individual cells, however, often changed over the course of the trial. For example, in the third exemplar unit activity shown in Figure 3 (right column), the influence of reward size on firing rate (i.e., β_{Reward}) reversed repeatedly in the post-instruction epoch. This type of variability in the regression coefficients impeded simple approaches for categorization of STN neurons via their pattern of encodings (e.g., positive reward encoding vs. negative).

Mixed encodings of reward size and delay

To gain deeper insight into how the reward and delay dimensions of the task were integrated by a neuron's activity, we projected the unit's time series of regression coefficients from Eq. 3 (β_{Reward} and β_{Delay}) into a space in which reward size and delay compose two orthogonal dimensions. For each neuron, vector time series were produced in this regression space for significant β values ($P < 0.05$) to capture the moment-by-moment mixture of encodings (Fig. 3C). In this space, vector angles indicated how a neuron's activity reflected the combined effects of reward size and delay, while vector magnitude captured the strength of the combined encoding. To determine the predominant encoding of these two characteristics (angle and magnitude of moment-by-moment vectors) and

their evolution during the post-instruction epoch, we summed across the time-resolved vectors across two consecutive phases of the waiting period [e.g., red dashed lines for phase 1 (0-2s post-instruction) and phase 2 (2-3.5s post-instruction) in Fig. 3C]. The angles (θ_1 and θ_2) of the resulting vector sums were used to identify consecutive patterns of activity consistent with, and those inconsistent with, encoding of a temporal discounting of reward value – that is, an encoding in which reward size and delay have opposing effects on firing rate (Fig. 4A). Over the two phases of the waiting period, some neurons exhibited a consistently positive encoding of reward combined with a negative encoding of delay (vector angles $-90^\circ < \theta < 0^\circ$; referred to as the ‘Discounting–’ pattern in Fig. 4A; see e.g., Fig. 3C *right*), while others modulated their activity in the converse pattern with a negative β_{Reward} and positive β_{Delay} value ($90^\circ < \theta < 180^\circ$; referred to as ‘Discounting+’ pattern; see e.g., Fig. 3C *middle*). Other neurons drastically changed their pattern between phases of the waiting period (Fig. 3C *left*). And some encoded reward size and delay in an additive fashion, inconsistent with a signal reflecting subjective value and referred to here as ‘Compounding+’ and ‘Compounding–’ patterns (Fig. 4A; see e.g., Fig. 3C *left*). Compounding signals like these are inconsistent with a temporal discounting of reward value and may instead be attributable to extraneous factors such as arousal or attentional engagement.

At the population level, STN neurons exhibited variable mixed signals over the course of the waiting period with, on average, a change in angle of 34 degrees measured between vector sums of phases 1 and 2 (Fig. 4B-D). First, in phase 1, the neural encoding of reward size and delay parameters predominantly followed a Discounting pattern as evidenced by the fraction of neurons with Discounting– type encodings ($\chi^2=17.1$ df=3 $P=0.007$; Fig. 4E), and the longer mean vector magnitude of Discounting– units (One-way ANOVA, $F_{(3,121)}=4.54$ $P=0.005$; Fig. 4G). Of the 124 task-related neurons, 60 (48%) increased firing rates as a function of reward size while they decreased according to the temporal delay to reward delivery (i.e, consistent with a Discounting– pattern). The remaining neurons were distributed across the other three encoding patterns. Vector magnitudes for neurons with a Discounting– firing pattern were longer, on average, than those for neurons with either type

of Compounding pattern, while the vector magnitude for neurons with a Discounting+ pattern fell in-between. Notably, the angle of the mean vector across all neurons ($\theta_1 = -12^\circ$, Fig. 4C) showed that, despite the wide diversity of encoding patterns across individual neurons in phase 1, the whole neural ensemble encoded information about reward size and delay in a pattern that was strongly consistent with a temporal discounting of value (i.e, Discounting-). The significance of this bias in the population encoding was supported further by the observation that the angles of individual vectors were distributed in a markedly non-uniform fashion (Rayleigh's test, $z=12.5$ $P<0.001$; Fig.4C). Hence, during the first two seconds of the post-instruction period, the neural ensemble combined information related to reward size and delay into a coherent population-scale signal that reflected subjective value according to a Discounting- pattern.

Then, in phase 2, the population encoding of reward size and delay parameters changed drastically to a predominantly Compounding pattern (Fig. 4D). The fraction of neurons with Compounding+ type encodings increased markedly ($\chi^2=9.9$ $df=3$ $P=0.02$; Fig. 4F), and mean vector magnitude of Compounding+ units was longer than those of units with other types of encoding (One-way ANOVA, $F_{(3,119)}=3.91$ $P=0.01$; Fig. 4H). Of the 124 task-related neurons, 42 (34%) increased firing rates during this period as a function of both reward size and delay (i.e, consistent with a Compounding+ pattern), while only 24 (19%) were still categorized as Discounting- type neurons. Inconsistent with a temporal discounting of reward value, the angle of the mean vector across all neurons ($\theta_2 = 65^\circ$, Fig. 4D) showed that the whole neural ensemble encoded significantly a signal correlated positively with both the reward size and the time delay (Rayleigh's test, $z=40.7$ $P<0.001$). This dynamic shift in the encoding of task-relevant information (i.e, rotation of the vectors between phases 1 and 2) was detected not only at the population level but also in the activities of 27 individual units (22%) in which the encoding pattern switched from Discounting- to Compounding+ in the transition from phase 1 to phase 2, respectively.

Remarkably, the neurons categorized as Discounting– in phase 1 and/or Compounding+ in phase 2 were located preferentially in the most posterior and dorsal portion of the STN (two sample t -test, $t_{(229)} > 3.04$ $P < 0.05$; Fig. 5). We did not observe a strict anatomic segregation of neurons with different encoding patterns, but rather an intermixed gradient of cell types along the antero-posterior axis. Discounting– neurons (Fig. 5A-G) and Compounding+ neurons (Fig. 5H-N) did not differ from other types of STN neurons with respect to tonic firing rate ($t_{(229)} < 1.87$ $P > 0.06$) or action potential shape ($t_{(229)} < 1.03$ $P > 0.3$). Aside from the dorso-posterior bias to their position in the STN, the neurons categorized as Discounting– in phase 1 and/or Compounding+ in phase 2 were indistinguishable from the others.

Dynamic encodings by the neural population ensemble

We then performed a population-based analysis using all recorded neurons ($n=231$) to identify the principal patterns of reward size and delay encoding within the neural ensemble and how they changed dynamically over the course of a trial. After projecting every unit's time series of regression coefficients (β_{Reward} and β_{Delay}) into the orthogonal space composed of the reward size and delay, we used a principal component analysis (PCA) to identify the predominant patterns (i.e., the Principal Components, "PCs") of encoding across the population. Unlike the single-unit analyses, the estimation of regression coefficients (β values) here was calculated in non-overlapped temporal windows (100-ms) to maintain the independence of β values across time bins. Given that the activity of individual neurons encoded diverse time-varying representations of task-relevant information (see e.g., Fig. 3), the PCA identified the patterns that accounted for the greatest variance within the neural population. In the resulting time series of eigenvectors, the angle of the eigenvectors indicated how reward size and delay parameters were integrated at the population level, while the magnitude of eigenvectors captured the relative strength of the encodings in the neural ensemble.

We identified the PCs that accounted for a significant fraction of variance relative to that accounted for by a population of control PCAs (Fig. 6A-B). Surrogate control PCAs were computed by shuffling neural activity across trials before application of the PCA, thereby scrambling the relationship between task conditions and the neural activity. We found that the first four PCs from the real data exceeded the 95% confidence interval of variances accounted for by PCs from the surrogate control PCAs. In total, the first four PCs explained 72.4% of the total variance (Fig. 6A). These four PCs were distinct from each other with respect to the pattern represented (Fig. 6E) and how it evolved across time in a trial (Fig. 6D). We identified the timing of significant encodings of reward size and delay in PCs by comparing the magnitude of eigenvectors to the 95% confidence interval of those calculated from the surrogate control PCAs (Fig. 6F). Consistent with our single-unit analyses, the first principal component (PC1, which accounted for 45.1% of variance) corresponded primarily to a temporal discounting of reward value (i.e. a signal that increased with increasing reward size and decreased with increasing delay; $-90^\circ < \theta < 0^\circ$) that later evolved to a Compounding+ pattern (i.e., a signal in which reward size and delay were integrated in an additive manner; $0^\circ < \theta < 90^\circ$). As evidenced by the eigenvector time series of PC1, the neural ensemble encoded a delay-discounted reward value (pointing toward -18°) in a stable manner for approximately 2-s after the presentation of the instruction cue and then rotated progressively toward a nominally-significant pattern (pointing toward 44°) in which activity increased with longer delays to reward. Thus, reward size and delay were combined into a common neural ensemble signal that evolved dynamically across the duration of the waiting period. The signal captured by PC2 (second column of Fig. 6; accounting for 13.9% of variance) differed markedly from that of PC1. In the 1-s after presentation of the instruction cue, the PC2 population signal was primarily sensitive to reward size only, with a transient encoding pointing toward 0° . Then, approximately 2-s after the instruction, PC2 deviated to a robust Discounting+ pattern ($90^\circ < \theta < 180^\circ$) in which reward size and delay had opposing effects on the signal. Consistent with an encoding of temporally discounted values, this Discounting+ pattern was maintained throughout the later half of the waiting period, with a slow progressive rotation of the vector toward

an increasing influence of delay the longer the delay period lasted (145 to 110°, Fig.6E). Once again, as with PC1, PC2 combined reward size and delay into an ensemble encoding of integrated value that evolved dynamically across the hold period.

Similarly, the eigenvector properties for PC3 (7.4% of total variance explained; third column of Fig. 6) revealed a population signal that involved two consecutive temporal phases. Between 0.5-1.5s after presentation of the instruction cue, PC3 approximated a transient Discounting– type pattern of value encoding (pointing toward -24°), after which, PC3 deviated to a relatively stable Compounding– pattern ($-90^\circ < \theta < -180^\circ$) in which neural signals correlated negatively with both the size of reward and the time delay. Finally, the signal extracted by PC4 (6% of variance explained) was characterized by multiple salient signals reflecting reward values only (close to 0°) and Discounting+ ($90^\circ < \theta < 180^\circ$) during relatively short periods after instruction. An additional signal in PC4 detected late in the hold period corresponded to a Discounting– pattern ($-90^\circ < \theta < 0^\circ$). Of interest, Fig.6 – figure supplement 1 shows that a PCA using sliding window analysis (200-ms test window stepped in 20-ms) in the estimation of regression coefficients (β values) produced results that were very similar to those from the non-overlapped temporal window analysis (see % of variance explained per PC), but with a higher number of time bins.

An anatomical gradient through STN

To test whether the predominant patterns of population encoding (i.e. PCs) varied as a function of location in the STN, we correlated the component scores of neurons with their position along variable anatomical axes (45° rotations along x-y-z axes). One anatomical axis correlated significantly with PC1 and PC2, i.e. the two major components consistent with a processing of temporally discounted value (Fig. 7). Because the scores for PC1 (Spearman's $Rho=0.35$ $P<0.001$; Figs. 7) and PC2 ($Rho=0.19$ $P=0.003$) increased along the ventro-anterior to dorso-posterior axis in the STN, this result indicates that reward size and delay were integrated into a common value most strongly in the dorso-posterior STN. No significant correlations were found in the spatial distribution of PC3 and PC4

scores ($\rho < 0.08$ $P > 0.21$), suggesting that these population signals were processed by neurons anatomically intermixed with others in the entire STN.

DISCUSSION

The present results reveal how STN neurons process temporally discounted value when a behavioral inhibition needs to be patiently sustained over time prior to delivery of a reward. At the single-neuron and population levels, we found a cost-benefit integration between the desirability of the expected reward and the imposed delay to delivery, with signals that dynamically combined both reward-related attributes to form a single integrated value estimate. The computation for such subjective value was increasingly observed along the antero-posterior axis of the STN, revealing that the most dorso-posterior-placed neurons are the most strongly involved in the representation of temporal discounting of value. These results expand our understanding concerning the involvement of STN in motivation and inhibitory control, by providing evidence for complex dynamical codings consistent with a continuous cost-benefit estimation promoting the goal pursuit and the willingness to bear the costs of time delays.

To determine if the STN conveys dynamic signals consistent with its predicted role in pursuing delayed gratification, we trained monkeys to perform a delayed reward task in which they were required to remain motionless for varying periods of time before the delivery of reward. We designed this task to investigate the ability of our animals to delay gratification and maintain a continuous self-control according to different levels of motivation. Previous studies have already tested monkeys' behavior for delay maintenance (Evans et al., 2012; Evans and Beran, 2007; Freeman et al., 2012, 2009; Perdue et al., 2015; Szalda-Petree et al., 2004), but none to our knowledge have investigated the underlying neural mechanisms. Our analysis of animals' performance confirmed that different levels of motivation or subjective value influenced their

patience and willingness to sustain behavioral inhibition over time (Fig. 1). As evidenced by the more frequent rejection of trials offering small rewards and long delays, both cost-benefit parameters were integrated by monkeys to complete trials and finally obtain rewards. Consistent with previous findings (Fujimoto et al., 2019; Minamimoto et al., 2009), the likelihood of staying engaged in the trials for delayed rewards was well approximated by a model that calculates the subjective value with a hyperbolic delay discounting. This shows that monkeys estimated the cost-benefit conditions properly, and that active motivational processes were recruited during the time course of the post-instruction waiting period to maintain their behavioral inhibition.

In both human and non-human animal research, temporal discounting has been traditionally studied via the intertemporal choice task, which pits a small reward available sooner against a large reward available later (Frederick et al., 2002). As the delay to the large reward becomes longer, agents tend to start discounting the value of the large reward, biasing their preference toward the small reward available sooner. This choice behavior is considered impulsive and is referred to as a failure of self-control because it would be more economical to wait for the larger reward (Ainslie, 1975; Rachlin, 2004). To interpret individual choice behavior, the extent to which rewards are devalued over time have been inferred primarily using hyperbolic discounting models (Mazur, 2001; Vanderveldt et al., 2016). Based on that approach, neuroimaging studies have identified a network of brain areas – involving the striatum, medial prefrontal cortex and posterior cingulate cortex – activation of which correlates with the subjective value of delayed rewards during decision-making processes (Kable and Glimcher, 2007; McClure et al., 2007, 2004; Peters and Büchel, 2009; Pine et al., 2009). Consistent with a coding of cost-benefit integration for delays to reward, BOLD activity in these regions increases as the expected amount of a reward increases and, inversely, decreases as a function of the imposed delay to reward. At the single neuron level, however, electrophysiological results have been more mixed concerning the integration of both reward-related attributes into a common currency value attribution (Roesch and Bryden, 2011). While some studies have found strict dissociable

representations of both neural signals by different populations of neurons (Roesch et al., 2007, 2006; Roesch and Olson, 2005), two monkey studies have reported single-unit activities co-modulated by both reward size and delay (Cai et al., 2011; Kim et al., 2008). Our results confirm and extend to the STN those findings by showing that the cost and benefit dimensions of the task are represented in different ways in different sub-groups of neurons, either independently or in an integrated manner during the waiting period. Respectively, 14% and 5% of STN neurons were exclusively modulated by the expectation of reward size and delay without any co-modulation by the other parameter, maintaining distinct dynamic encodings split between these subsets of cells. Signals related to reward size were represented preferentially shortly after cue presentation, while signals related to delay cost occurred primarily later as animals endured the waiting period (Fig. 2). The fact we were able to dissociate in the STN these two types of neurons suggests that temporal discounting originates likely from different neural circuits than those that signal expected reward value. Alternatively, 34% of our STN neurons exhibited a dual coding of size and delay information, engendering a representation of both cost and benefit into a common dimension. Among those cells, the most prevalent activity pattern was characterized by opposing effects of reward magnitude and delay on their firing (Fig. 4). Specifically, STN neurons that fired more strongly for larger anticipated rewards tended to fire less strongly for longer delays (referred as Discounting-), consistent with the computation of temporally discounted value. Despite a high variability across individual neurons, we found that the whole neural ensemble sampled from the STN encoded and combined both reward size and delay into a strong population signal that corresponded with a temporal discounting of reward value (Fig. 6). As evidenced by the eigenvector time series of different PCs, reward size and delay were integrated into a common signal that evolved dynamically while animals remained immobile, waiting for the reward. Such dynamics appear consistent with the hypothesized role for STN in self-control and perseverance.

Unlike most studies of temporal discounting that focus on decision-related neural processes at just one time point in the task, we investigated activity during the post-instruction delay period, during which animals were required to exert sustained commitment to an action with a continuous behavioral inhibition. It is quite likely that distinct components of self-control are measured during those two time periods (Addessi et al., 2013). Our approach offers the opportunity to investigate the dynamic processes engaged during the post-instruction period to support an animal's ability to delay gratification. An important avenue for future research will be to determine how STN signals, such as those described here, change when animals run out of patience and finally decide to stop waiting. To do this, however, smaller reward sizes and longer delays might be used to promote more escape behaviors during the delay interval. Rejection rates were relatively low in the present version of our task. Given current models where STN is thought to prevent hasty decisions by elevating its activity to pause cortical commands via pallido-thalamocortical circuits (Cavanagh et al., 2011; Frank, 2006; Mansfield et al., 2011), signals underlying a steeper temporal discounting may be observed shortly before behavioral disinhibition if this nucleus effectively drives this type of inhibitory function. In addition, further studies relying on simultaneous multisite recordings are still necessary to clarify the neural origins of the information transmitted by the signals identified here. Although our results show that STN neurons process a dual dynamic coding of size and delay information, it remains undetermined whether the integration between these two reward-related attributes occurs primarily within this nucleus or upstream in other brain areas such as the dorsolateral prefrontal cortex (Kim et al., 2008) or the anterior striatum (Cai et al., 2011), for instance.

A growing number of animal studies have examined the involvement of STN in motivational functions using single-unit recordings. Neurons with phasic responses evoked by reward-predictive cues and the reward itself have been reported in different species and tasks (Breyse et al., 2015; Darbaky et al., 2005; Espinosa-Parrilla et al., 2013; Lardeux et al., 2013, 2009; Matsumura et al., 1992; Nougaret et al., 2022; Teagarden and Rebec, 2007). However, it has been unclear how this reward processing

495 relates to the known organization of STN into anatomically and functionally distinct territories
496 (Alexander et al., 1990; Nambu et al., 2002; Parent and Hazrati, 1995). The STN receives
497 topographically-organized inputs from most regions of the frontal cortex (Haynes and Haber, 2013;
498 Nambu et al., 2002), the pallidum (Karachi et al., 2005; Shink et al., 1996), and the parafascicular
499 nucleus of the thalamus (Sadikot et al., 1992). Together, tracing studies have indicated that the
500 posterior-dorsal-lateral STN is interconnected with circuits devoted to sensorimotor function,
501 whereas associative- and limbic-related territories are found in progressively more anterior, ventral
502 and medial regions of the STN (Haynes and Haber, 2013; Mettler and Stern, 1962; Monakow et al.,
503 1978; Shink et al., 1996). Based on this, a widely accepted tripartite model divides STN into
504 segregated motor, associative, and limbic regions that are thought to play distinct roles in motor
505 control, cognition, and emotion (Parent and Hazrati, 1995). Surprisingly, in previous primate studies,
506 reward-responsive neurons were found to be scattered throughout all parts of the STN (Darbaky et
507 al., 2005; Espinosa-Parrilla et al., 2013; Nougaret et al., 2022), rather than showing a preferential
508 location in the anterior portion which is held to be the zone with strongest connectivity to limbic
509 structures (Haynes and Haber, 2013; Karachi et al., 2005). That lack of anatomic specificity has been
510 confirmed in human recordings for which reward-modulated neurons were also identified in
511 sensorimotor regions of the STN (Justin Rossi et al., 2017; Sieger et al., 2015). To date, these
512 paradoxical observations were attributed to a bias in data collection and the inherent observational
513 bias in data collected as part of DBS implantation surgeries due to the posterior-dorsal-lateral
514 location of the target in those surgeries. By accumulating 231 neurons distributed across all portions
515 of the STN, our study provides a more complete survey of this nucleus for regional variations in the
516 representation of reward-related information. At the single-neuron and population levels (Figs. 5,7),
517 we found that the valuation of delayed rewards was represented preferentially in the most dorso-
518 posterior portion of the STN, thereby challenging predictions from the classical tripartite model of
519 STN functional topography. While earlier studies suggested segregated functional subdivisions, now
520 more recent evidence points toward overlapping territories (Alkemade and Forstmann, 2014; Emmi

et al., 2020). Consistent with a functional convergence within this structure, our results show that motivational signals were conveyed by neurons located in areas traditionally identified as part of the sensorimotor circuits. Because muscle activities were not altered by reward contingencies during the task (Figs. 1H-I), we assume that these dynamic reward-related signals were primarily processed by cognitive circuits that monitor and manage goal achievement by translating motivational drives into behavioral perseverance. This role in self-control extends our understanding concerning the involvement of STN in the control of impulsive behaviors (Aron et al., 2016; Jahanshahi et al., 2015; Zavala et al., 2015) and the willingness to work for food (Baunez et al., 2005, 2002). By providing evidence for dynamic cost-benefit valuation by STN, our results imply that this structure promotes the pursuit of motivated behavior by computing which costs are acceptable for the reward at stake. Further research is needed to determine whether the neural signals identified here causally drive animals' behavior or rather just participate to reflect or evaluate the current situation.

Impatience for reward and lack of perseverance are major facets of many psychiatric disorders. Numerous clinical studies have shown that these maladaptive behaviors are characterized by a disruption in the ability to weigh appropriately the amount of reward against the cost of delay (Kirby et al., 1999; Madden et al., 1997; Mitchell, 1999; Reynolds, 2006; Vuchinich and Simpson, 1998). For instance, patients with self-control issues such as in gambling disorder (Alessi and Petry, 2003), drug addiction (Kirby and Petry, 2004; Madden et al., 1997; Washio et al., 2011), schizophrenia (Ahn et al., 2011; Heerey et al., 2007; MacKillop and Tidey, 2011), depression (Dombrovski et al., 2012; Pulcu et al., 2014; Takahashi et al., 2008), mania (Mason et al., 2012), attention deficit hyperactivity disorder (Barkley et al., 2001; Scheres et al., 2010; Tripp and Alsop, 2001), and anxiety disorder (Rounds et al., 2007) have higher delay discounting rates than normal subjects. Our results merge existing lines of evidence that implicate the STN in motivation and inhibitory control, positioning this structure as a potential hub to regulate aberrant reward processing and the capability to postpone. This view agrees with the beneficial effects of DBS-STN on drug addiction reported in recent animal studies

(Pelloux et al., 2018; Rouaud et al., 2010; Wade et al., 2017), and with the battery of psychiatric side-effects observed in parkinsonian patients after electrode implantation within the STN (Castrìoto et al., 2014). However, an open question remains on how the STN contributes in these pathological states to distort defective valuation processes in terms of cost-benefit trade-off.

MATERIALS AND METHODS

Animals

Two rhesus monkeys (monkey C, 8 kg, male; and monkey H, 6 kg, female) were used in this study. Procedures were approved by the Institutional Animal Care and Use Committee of the University of Pittsburgh (protocol number: 12111162) and complied with the Public Health Service Policy on the humane care and use of laboratory animals (amended 2002). When animals were not in active use, they were housed in individual primate cages in an air-conditioned room where water was always available. The monkeys' access to food was regulated to increase their motivation to perform the task. Throughout the study, the animals were monitored daily by animal care staff and veterinary technicians for evidence of disease or injury and body weight was documented weekly. If a body weight <90% of baseline was observed, the food regulation was stopped.

Behavioral task

Monkeys were trained to perform the delayed reward task with the left arm using a torquable exoskeleton (KINARM, BKIN Technologies, Kingston, Ontario, Canada). This device had hinge joints aligned with the monkey's shoulder and elbow and allowed the animal to make arm movements in the horizontal plane. Visual cues and cursor feedback of hand position were presented in the horizontal plane of hand movements by a virtual-reality system. A detailed description of the apparatus can be found in our previous studies (Pasquereau and Turner, 2015, 2013).

In our delayed reward task (Fig. 1A), the monkey was required to align the cursor on a visual target (radius, 1.8 cm) and to maintain this position for varying periods of time before delivery of the food reward. In total, six combinations of reward size and waiting delay were used. A trial began when a gray-filled target appeared (the same location for all trials) and the animal made the appropriate joint movements to place the cursor in this circle. Maintenance of the cursor within the target required the animal to actively stabilize the posture of both shoulder and elbow joints in the horizontal plane. While the monkey maintained its arm position, an instruction cue was displayed over the gray-filled target for 0.5 s. After a variable interval (1.2-2.8 s), the gray fill disappeared from the circle, cueing the animal to remain motionless during an additional delay until the reward delivery. For the instruction cues, cue colors indicated the size of reward (1, 2, and 3 drops of food) and symbols indicated the delay duration that the animal would have to wait before reward delivery [short delay (3.5-5.6s) and long delay (5.2-7.3s)]. The gaussian distributions of these two delay ranges overlapped for 9% of trials. Cue colors were calibrated to have the same physical brightness (~30 cd/m²). The six unique cue types (3 reward sizes x 2 delay ranges) were presented in pseudo-random order across trials with equal probability. At the end of each successful trial, food reward was delivered via a sipper tube attached to a computer-controlled peristaltic pump (1 drop = ~0.5 ml, puree of fresh fruits and protein biscuits). The trials were separated by 1.5–2.5 s intertrial intervals, during which the screen was black. Failures to maintain the cursor in the start position (radius = 1.8 cm) during the interval between instruction cue and reward delivery were counted as errors. After an error, the rejected trial was aborted and a blank screen appeared (1s), followed by an intertrial interval.

To confirm the ability of our monkeys to evaluate appropriately the different instruction cues, we also trained them to perform a decision task in which they were allowed to choose freely between two alternate reward size or delay conditions. The overall structure of this task was similar to the delayed reward task except two instruction cues were presented simultaneously. The location of the

cues alternated randomly between left and right sides. The animal chose its preferential option by positioning and maintaining the cursor on one of the targets. The pair of cues presented on a single trial differed only along one dimension, presenting either a reward-based decision (different colors but the same symbol) or a delay-based decision (different symbols but the same color).

Before the start of data collection, we trained the monkeys to perform these two behavioral tasks for more than 6 months. Neuronal data were not collected during performance of the decision task.

Surgery

After reaching asymptotic task performance, animals were prepared surgically for recording using aseptic surgery under Isoflurane inhalation anesthesia. An MRI-compatible plastic chamber (custom-machined PEEK, 28 x 20 mm) was implanted with stereotaxic guidance over a burr hole allowing access to the STN. The chamber was positioned in the parasagittal plane with an anterior-to-posterior angle of 20°. The chamber was fixed to the skull with titanium screws and dental acrylic. A titanium head holder was embedded in the acrylic to allow fixation of the head during recording sessions. Prophylactic antibiotics and analgesics were administered post-surgically.

Localization of the recording site

The anatomical location of the STN and proper positioning of the recording chamber to access it were estimated from structural MRI scans (Siemens 3 T Allegra Scanner, voxel size of 0.6 mm). An interactive 3D software system (Cicerone) was used to visualize MRI images, define the target location and predict trajectories for microelectrode penetrations (Miocinovic et al., 2007). Electrophysiological mapping was performed with penetrations spaced 1 mm apart. The boundaries of brain structures were identified based on standard criteria including relative location, neuronal spike shape, firing pattern, and responsiveness to behavioral events (e.g. movement, reward). By aligning microelectrode mapping results (electrophysiologically characterized X-Y-Z locations) with

structural MRI images and high resolution 3-D templates of individual nuclei derived from an atlas (Martin and Bowden, 1996), we were able to gauge the accuracy of individual microelectrode penetrations and determine chamber coordinates for the STN.

Recording and data acquisition

During recording sessions, a single glass-coated tungsten microelectrode (impedance: 0.7–1 MOhm measured at 1000 Hz) was advanced into the target nucleus using a hydraulic manipulator (MO-95, Narishige). Neuronal signals were amplified with a gain of 10K, bandpass filtered (0.3–10 kHz) and continuously sampled at 25 KHz (RZ2, Tucker-Davis Technologies, Alachua FL). Individual spikes were sorted using Plexon off-line sorting software (Plexon Inc., Dallas TX). The timing of detected spikes and of relevant task events was sampled digitally at 1 kHz. Horizontal and vertical components of eye position were recorded using an infrared camera system (240 Hz; ETL-200, ISCAN, Woburn, MA). For electromyographic recordings (EMG, in monkey H only), pairs of Teflon-insulated multi-stranded stainless-steel wires were implanted into multiple muscles during 12 training sessions. EMG signals were differentially amplified (gain = 10 K), band-pass filtered (200 Hz to 5 kHz), rectified and then low-pass filtered (100 Hz).

Analysis of behavioral data

We analyzed the way the animals performed the delayed reward task to test whether the behavior varied according to the levels of reward and delay. As the monkey maintained its arm position for varying periods of time before to obtain rewards, the sum of errors per session was the main variable of interest. Rejection rates were calculated by dividing the number of errors by the total number of trials for each task condition. Two-way ANOVAs were used to test these rejections for interacting effects of reward size (1, 2, or 3 drops) and delay (short or long delay duration). Modulation in rejection rates reflected the monkey's motivation to stay engaged in the task and to fully complete the different types of trials by maintaining the correct arm position. In the present study, the

motivational or subjective value linked to each task condition was estimated by integrating the forthcoming reward size and the delay discounting. The subjective value of delayed reward is commonly formulated as a hyperbolic discounting model (Green and Myerson, 2004; Mazur, 1987) as follows:

$$SV = \frac{R}{1+kD} \quad (1)$$

where SV is the subjective value (i.e. the temporally discounted value), R is the reward size, k is a discount factor that reflects an individual animal's sensitivity to the waiting cost, and D is the delay to the reward. In previous studies (Hori et al., 2021; Kobayashi and Schultz, 2008; Minamimoto et al., 2009), this hyperbolic discounting model has been shown consistently to fit to monkey behavior better than an exponential function. Because the number of errors in task performance is inversely related to the subjective value (Fujimoto et al., 2019; Minamimoto et al., 2009), we have inferred the subjective value in each monkey by fitting the average rejection rate to the following model:

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

where E is the rejection rate, R is the reward size, D is the delay, k is a discount factor, and a is a monkey-specific free parameter. We estimated the best pair of free parameters (k and a) with the Matlab function "fminsearch" that provided the maximum-likelihood fit. Goodness of fit was evaluated by the coefficient of determination (R^2). Because of the limited number of task conditions (only two delay ranges), we assumed a linearity in the estimation of reward value.

As monkeys were not required to control their gaze while performing the task, we tested whether eye positions varied according to the levels of reward and delay. For this analysis, we combined horizontal and vertical components of eye position to obtain tangential coordinates, and the

potential interaction with task parameters was examined using a two-way ANOVA combined with a sliding window procedure (200-ms test window stepped in 20-ms). The threshold for significance was corrected for multiple comparisons ($P < 0.05/ n\text{-time bins}$; Bonferroni correction). The same statistical method was used to analyze EMGs, with series of two-way ANOVAs performed to test for effects of reward size and delay. All of the data analyses were performed using custom scripts in the MATLAB environment (The MathWorks, MA).

Neuronal data analysis

Neuronal recordings were accepted for analysis based on electrode location, recording quality (signal/noise ratio of >3 SD) and duration (>120 trials). The width of action potential waveforms was calculated as the interval from the beginning of the first negative inflection (>2 SD) to the subsequent positive peak, and the magnitude of the biphasic spike waveforms was measured between maxima and minima. Trials with errors were excluded from the analysis of neuronal data. Continuous neuronal activation functions [spike density functions (SDFs)] were generated around instruction cues (-1 to 3.5 s) by convolving each discriminated action potential with a Gaussian kernel (20-ms variance). Mean peri-event SDFs (averaged across trials) for each of the reward and delay conditions were constructed. A neuron's baseline firing rate was calculated as the mean of the SDFs across the 1 s epoch preceding cue instruction. The Fano factor, defined as the variance-to-mean ratio of firing rates, was used to measure the variability of neuronal activities across the six trial conditions. For each single-unit activity, we tested for effects of reward size and delay using two-way ANOVAs combined with a sliding window procedure (200-ms test window stepped in 20-ms). Specifically, we extracted single-trial spike counts across a series of 200-ms windows and we investigated for each step whether the neuronal activity was influenced by 1) reward size, 2) delay to reward, or 3) both task parameters. The ANOVA identified any interacting effects of reward size and delay. The threshold for significance in those ANOVAs was corrected for multiple comparisons using the Bonferroni correction ($P < 0.05/ n\text{-time bins}$). Application of the same ANOVA analysis and statistical

threshold ($p < 0.05/n$ -time bins) to neuronal data during the 1 s period of the pre-instruction control epoch (i.e., from a time period when the animal had no information about the upcoming reward and delay) resulted in 0.12% of Type 1 (false-positive) errors, thereby confirming that the threshold for statistical significance was appropriate. A neuron was judged to be task-related if its firing rate reflected a significant encoding of the reward size and/or delay for a least one-time bin.

To test how individual task-related neurons dynamically encoded the forthcoming reward and delay discounting through trials, we used time-resolved multiple linear regressions. We tested whether trial-to-trial neuronal activity was modulated simultaneously by the size of reward (R), the delay duration (D), and measures of eye movements made in the task [tangential position (P) and tangential velocity (V)] to control for possible relationships of STN activity with gaze control. For each task-related neuron, we counted spikes (SC) trial-by-trial within a 200-ms test window that was stepped in 20-ms increments across the 3.5-s period following onset of the instruction cue. For each bin, we applied the following model:

$$SC_i = \beta_0 + \beta_R R_i + \beta_D D_i + \beta_P P_i + \beta_V V_i, \quad (3)$$

Where all regressors for the i^{th} trial were normalized to obtain standardized regression coefficients (Z-scored in standard deviation units). The β coefficients for each task-related unit were estimated using the 'glmfit' function in Matlab. The threshold for significance of individual test β values was determined by comparing them against a population of 46 control β values calculated by applying the same sliding window linear regression approach to spike counts from a 1 s pre-instruction control epoch (one sample t -test, $df=46$, $P < 0.05$).

To characterize how individual neurons integrate task parameters dynamically, time series of regression coefficients of Eq. 3 (β_R and β_D) were projected into an orthogonal space where reward

size and delay composed the two dimensions (axes) of interest. In this regression space, we used significant points (pairs of β_R - β_D) to produce vector time series originating from the control value of the pre-instruction epoch (averaged across time bins). Vectors were generated with significant regression coefficients calculated during the course of the hold period. In this context, vector angles (-180 to 180°) indicated how the neural activity encodes and combines reward size and delay, while vector magnitudes captured the strength of the signals transmitted. Considering these two characteristics (direction and magnitude of vectors), we summed all time-resolved vectors to identify the predominant encoding of task parameters for each neuron analyzed. Various patterns of neural encoding could be categorized from the angle (θ) of the resultant vector sum. For instance, θ indicated whether the neural activity was correlated (positive coding) or anticorrelated (negative coding) with the reward size, the delay to reward, or both. The angle of these vectors was interpreted as follows: 0° (positive β_R values) indicated an exclusive positive reward size coding; $\pm 180^\circ$ (negative β_R values) indicated an exclusive negative reward size coding; 90° (positive β_D values) indicated an exclusive positive delay coding; and -90° (negative β_D values) indicated an exclusive negative delay coding (Fig. 4A). In this way, a θ between -90 and 0°, or between 90 and 180°, indicated a coding of both reward size and delay consistent with a temporally discounted value in which benefit and cost parameters have opposing effects on the subjective encoding (i.e, positive β_R with negative β_D , or vice versa). In contrast, a θ between 0 and 90°, or between -180 and -90°, indicated a coding of reward size and delay in which the two parameters are integrated in a compound signal inconsistent with a temporally discounted value (i.e, positive β_R and β_D , or negative β_R and β_D).

For population-based figures (Fig. 4), vectors were standardized between neurons by subtracting the mean values of the pre-instruction epoch and then dividing by 2 SD of this control period. The population encoding was considered significant if the distribution of the individual vector angles was non-uniform (Rayleigh's test, $P < 0.001$). The population vector was calculated by vectorially summing standardized cell vectors.

754

755 We then performed an alternative population-based analysis to characterize the predominant
756 patterns of neural encoding in the STN using all recorded neurons. We used a principal component
757 analysis (PCA) to identify patterns of encoding (dimensions) of the neuronal ensemble in the polar
758 orthogonal space composed of the reward size and delay. Our procedure was analogous to the
759 analyses performed by Yamada et al., 2021. A two-dimensional data matrix X of size $N_{(\text{neuron})} \times N_{(C \times T)}$
760 was prepared with regression coefficients of Eq. 3 (series of β_R and β_D), in which rows corresponded
761 to the total number of neurons and columns corresponded to the number of conditions (C , reward
762 sizes and delays) multiplied by the number of time bins (T). Unlike the single-unit analyses, the
763 estimation of regression coefficients (β values) here was calculated in non-overlapped temporal
764 windows (100-ms) to maintain the independence of β values across time bins. A series of
765 eigenvectors was obtained by applying PCA once to the data matrix X . In our analysis, the
766 eigenvectors represent vectors at different time bins in the orthogonal space composed of reward
767 size and delay. Principal components (PC) were estimated using the 'pca' function in Matlab. As was
768 the case for the vectors computed for single-units, vector angles (-180 to 180°) indicated how the
769 ensemble encodes and combines reward size and delay, while vector magnitudes captured the
770 strength of the signals transmitted. Adequate performance of PCA was estimated with the
771 percentages of variance explained by PCs. To test whether the PCA performance was significant, we
772 constructed a surrogate control population of PCs in which the neural activity (SC) was shuffled
773 across trials before application of Eq. 3. Consequently, in the surrogate data the linear projection of
774 neural activity into the regression subspace was randomized, eliminating any coherent modulation of
775 activity with task parameters (reward size and delay) in the matrix X . We built a surrogate control
776 population of PCs by repeating the shuffling procedure 1000 times and then compared percentages
777 of variance explained by actual PCs against the 95% confidence interval of variances accounted for by
778 the population of surrogate PCs. The percentage of variance accounted for by an actual PC was
779 deemed significant if its value fell outside of the 95% confidence interval of the population of

surrogate PCs. In addition, we identified the timing of significant encodings of reward size and delay in PCs by comparing the magnitude of eigenvectors to the 95% confidence interval of those calculated from the population of surrogate PCs.

To test how PCs mapped onto anatomical axes in STN, we correlated the component scores of the neurons with their locations (x-y-z coordinates). We performed this in standard stereotaxic space and also for a rotated set of anatomical axes (45° rotations along x-y-z axes). The rotated axis that best correlated with the scores of a component (maximal Spearman's Rho) was identified as the optimal axes for explaining that component's variance.

REFERENCES

- Addressi E, Paglieri F, Beran MJ, Evans TA, Macchitella L, De Petrillo F, Focaroli V. 2013. Delay choice versus delay maintenance: different measures of delayed gratification in capuchin monkeys (*Cebus apella*). *J Comp Psychol* **127**:392–398. doi:10.1037/a0031869
- Ahn W-Y, Rass O, Fridberg DJ, Bishara AJ, Forsyth JK, Breier A, Bussemeyer JR, Hetrick WP, Bolbecker AR, O'Donnell BF. 2011. Temporal discounting of rewards in patients with bipolar disorder and schizophrenia. *J Abnorm Psychol* **120**:911–921. doi:10.1037/a0023333
- Aiello M, Terenzi D, Furlanis G, Catalan M, Manganotti P, Eleopra R, Belgrado E, Rumiati RI. 2019. Deep brain stimulation of the subthalamic nucleus and the temporal discounting of primary and secondary rewards. *J Neurol* **266**:1113–1119. doi:10.1007/s00415-019-09240-0
- Ainslie G. 1975. Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin* **82**:463–496. doi:10.1037/h0076860
- Alessi SM, Petry NM. 2003. Pathological gambling severity is associated with impulsivity in a delay discounting procedure. *Behav Processes* **64**:345–354. doi:10.1016/s0376-6357(03)00150-5
- Alexander GE, Crutcher MD, DeLong MR. 1990. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions. *Prog Brain Res* **85**:119–146.
- Alkemade A, Forstmann BU. 2014. Do we need to revise the tripartite subdivision hypothesis of the human subthalamic nucleus (STN)? *Neuroimage* **95**:326–329. doi:10.1016/j.neuroimage.2014.03.010
- Aron AR, Herz DM, Brown P, Forstmann BU, Zaghoul K. 2016. Frontosubthalamic Circuits for Control of Action and Cognition. *J Neurosci* **36**:11489–11495. doi:10.1523/JNEUROSCI.2348-16.2016
- Baker PM, Ragozzino ME. 2014. The prelimbic cortex and subthalamic nucleus contribute to cue-guided behavioral switching. *Neurobiol Learn Mem* **107**:65–78. doi:10.1016/j.nlm.2013.11.006

813 Barkley RA, Edwards G, Laneri M, Fletcher K, Metevia L. 2001. Executive functioning, temporal
814 discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD)
815 and oppositional defiant disorder (ODD). *J Abnorm Child Psychol* **29**:541–556.

816 Baunez C, Amalric M, Robbins TW. 2002. Enhanced food-related motivation after bilateral lesions of
817 the subthalamic nucleus. *J Neurosci* **22**:562–568.

818 Baunez C, Christakou A, Chudasama Y, Forni C, Robbins TW. 2007. Bilateral high-frequency
819 stimulation of the subthalamic nucleus on attentional performance: transient deleterious effects and
820 enhanced motivation in both intact and parkinsonian rats. *Eur J Neurosci* **25**:1187–1194.
821 doi:10.1111/j.1460-9568.2007.05373.x

822 Baunez C, Dias C, Cador M, Amalric M. 2005. The subthalamic nucleus exerts opposite control on
823 cocaine and “natural” rewards. *Nat Neurosci* **8**:484–489. doi:10.1038/nn1429

824 Baunez C, Robbins TW. 1997. Bilateral lesions of the subthalamic nucleus induce multiple deficits in
825 an attentional task in rats. *Eur J Neurosci* **9**:2086–2099.

826 Berney A, Vingerhoets F, Perrin A, Guex P, Villemure J-G, Burkhard PR, Benkelfat C, Ghika J. 2002.
827 Effect on mood of subthalamic DBS for Parkinson’s disease: a consecutive series of 24 patients.
828 *Neurology* **59**:1427–1429. doi:10.1212/01.wnl.0000032756.14298.18

829 Berns GS, Laibson D, Loewenstein G. 2007. Intertemporal choice--toward an integrative framework.
830 *Trends Cogn Sci* **11**:482–488. doi:10.1016/j.tics.2007.08.011

831 Bonnevie T, Zaghloul KA. 2019. The Subthalamic Nucleus: Unravelling New Roles and Mechanisms in
832 the Control of Action. *Neuroscientist* **25**:48–64. doi:10.1177/1073858418763594

833 Breyse E, Pelloux Y, Baunez C. 2015. The Good and Bad Differentially Encoded within the
834 Subthalamic Nucleus in Rats(1,2,3). *eNeuro* **2**:ENEURO.0014-15.2015. doi:10.1523/ENEURO.0014-
835 15.2015

836 Cai X, Kim S, Lee D. 2011. Heterogeneous coding of temporally discounted values in the dorsal and
837 ventral striatum during intertemporal choice. *Neuron* **69**:170–182. doi:10.1016/j.neuron.2010.11.041

838 Castrioto A, Lhommée E, Moro E, Krack P. 2014. Mood and behavioural effects of subthalamic
839 stimulation in Parkinson’s disease. *Lancet Neurol* **13**:287–305. doi:10.1016/S1474-4422(13)70294-1

840 Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ. 2011.
841 Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat*
842 *Neurosci* **14**:1462–1467. doi:10.1038/nn.2925

843 Coulthard EJ, Bogacz R, Javed S, Mooney LK, Murphy G, Keeley S, Whone AL. 2012. Distinct roles of
844 dopamine and subthalamic nucleus in learning and probabilistic decision making. *Brain* **135**:3721–
845 3734. doi:10.1093/brain/aws273

846 Darbaky Y, Baunez C, Arecchi P, Legallet E, Apicella P. 2005. Reward-related neuronal activity in the
847 subthalamic nucleus of the monkey. *Neuroreport* **16**:1241–1244. doi:10.1097/00001756-200508010-
848 00022

849 Dombrovski AY, Siegle GJ, Szanto K, Clark L, Reynolds CF, Aizenstein H. 2012. The temptation of
850 suicide: striatal gray matter, discounting of delayed rewards, and suicide attempts in late-life
851 depression. *Psychol Med* **42**:1203–1215. doi:10.1017/S0033291711002133

852 Emmi A, Antonini A, Macchi V, Porzionato A, De Caro R. 2020. Anatomy and Connectivity of the
853 Subthalamic Nucleus in Humans and Non-human Primates. *Front Neuroanat* **14**:13.
854 doi:10.3389/fnana.2020.00013

855 Espinosa-Parrilla J-F, Baunez C, Apicella P. 2013. Linking reward processing to behavioral output:
856 motor and motivational integration in the primate subthalamic nucleus. *Front Comput Neurosci*
857 **7**:175. doi:10.3389/fncom.2013.00175

858 Evans TA, Beran MJ. 2007. Delay of gratification and delay maintenance by rhesus macaques (Macaca
859 mulatta). *J Gen Psychol* **134**:199–216. doi:10.3200/GENP.134.2.199-216

860 Evans TA, Perdue BM, Parrish AE, Menzel EC, Brosnan SF, Beran MJ. 2012. How is chimpanzee self-
861 control influenced by social setting? *Scientifica (Cairo)* **2012**:654094. doi:10.6064/2012/654094

862 Evens R, Stankevich Y, Dshemuchadse M, Storch A, Wolz M, Reichmann H, Schlaepfer TE, Goshke T,
863 Lueken U. 2015. The impact of Parkinson's disease and subthalamic deep brain stimulation on
864 reward processing. *Neuropsychologia* **75**:11–19. doi:10.1016/j.neuropsychologia.2015.05.005

865 Frank MJ. 2006. Hold your horses: a dynamic computational role for the subthalamic nucleus in
866 decision making. *Neural Netw* **19**:1120–1136. doi:10.1016/j.neunet.2006.03.006

867 Frank MJ, Samanta J, Moustafa AA, Sherman SJ. 2007. Hold your horses: impulsivity, deep brain
868 stimulation, and medication in parkinsonism. *Science* **318**:1309–1312. doi:10.1126/science.1146157

869 Frederick S, Loewenstein G, O'donoghue T. 2002. Time Discounting and Time Preference: A Critical
870 Review. *Journal of Economic Literature* **40**:351–401. doi:10.1257/jel.40.2.351

871 Freeman KB, Green L, Myerson J, Woolverton WL. 2009. Delay discounting of saccharin in rhesus
872 monkeys. *Behav Processes* **82**:214–218. doi:10.1016/j.beproc.2009.06.002

873 Freeman KB, Nonnemacher JE, Green L, Myerson J, Woolverton WL. 2012. Delay discounting in
874 rhesus monkeys: equivalent discounting of more and less preferred sucrose concentrations. *Learn*
875 *Behav* **40**:54–60. doi:10.3758/s13420-011-0045-3

876 Fujimoto A, Hori Y, Nagai Y, Kikuchi E, Oyama K, Suhara T, Minamimoto T. 2019. Signaling Incentive
877 and Drive in the Primate Ventral Pallidum for Motivational Control of Goal-Directed Action. *J*
878 *Neurosci* **39**:1793–1804. doi:10.1523/JNEUROSCI.2399-18.2018

879 Fumagalli M, Rosa M, Giannicola G, Marceglia S, Lucchiari C, Servello D, Franzini A, Pacchetti C,
880 Romito L, Albanese A, Porta M, Pravettoni G, Priori A. 2015. Subthalamic involvement in monetary
881 reward and its dysfunction in parkinsonian gamblers. *J Neurol Neurosurg Psychiatry* **86**:355–358.
882 doi:10.1136/jnnp-2014-307912

883 Green L, Myerson J. 2004. A discounting framework for choice with delayed and probabilistic
884 rewards. *Psychol Bull* **130**:769–792. doi:10.1037/0033-2909.130.5.769

885 Haynes WIA, Haber SN. 2013. The organization of prefrontal-subthalamic inputs in primates provides
886 an anatomical substrate for both functional specificity and integration: implications for Basal Ganglia
887 models and deep brain stimulation. *J Neurosci* **33**:4804–4814. doi:10.1523/JNEUROSCI.4674-12.2013

888 Heerey EA, Robinson BM, McMahon RP, Gold JM. 2007. Delay discounting in schizophrenia. *Cogn*
889 *Neuropsychiatry* **12**:213–221. doi:10.1080/13546800601005900

890 Hori Y, Mimura K, Nagai Y, Fujimoto A, Oyama K, Kikuchi E, Inoue K-I, Takada M, Suhara T, Richmond
891 BJ, Minamimoto T. 2021. Single caudate neurons encode temporally discounted value for formulating
892 motivation for action. *Elife* **10**:e61248. doi:10.7554/eLife.61248

893 Jahanshahi M, Obeso I, Baunez C, Alegre M, Krack P. 2015. Parkinson's disease, the subthalamic
894 nucleus, inhibition, and impulsivity. *Mov Disord* **30**:128–140. doi:10.1002/mds.26049

895 Janakiraman N, Meyer RJ, Hoch SJ. 2011. The Psychology of Decisions to Abandon Waits for Service.
896 *Journal of Marketing Research* **48**:970–984. doi:10.1509/jmr.10.0382

897 Jimura K, Chushak MS, Braver TS. 2013. Impulsivity and self-control during intertemporal decision
898 making linked to the neural dynamics of reward value representation. *J Neurosci* **33**:344–357.
899 doi:10.1523/JNEUROSCI.0919-12.2013

900 Justin Rossi P, Peden C, Castellanos O, Foote KD, Gunduz A, Okun MS. 2017. The human subthalamic
901 nucleus and globus pallidus internus differentially encode reward during action control. *Hum Brain*
902 *Mapp* **38**:1952–1964. doi:10.1002/hbm.23496

903 Kable JW, Glimcher PW. 2007. The neural correlates of subjective value during intertemporal choice.
904 *Nat Neurosci* **10**:1625–1633. doi:10.1038/nn2007

905 Karachi C, Yelnik J, Tandé D, Tremblay L, Hirsch EC, François C. 2005. The pallidosubthalamic
906 projection: an anatomical substrate for nonmotor functions of the subthalamic nucleus in primates.
907 *Mov Disord* **20**:172–180. doi:10.1002/mds.20302

908 Kim S, Hwang J, Lee D. 2008. Prefrontal coding of temporally discounted values during intertemporal
909 choice. *Neuron* **59**:161–172. doi:10.1016/j.neuron.2008.05.010

910 Kirby KN. 1997. Bidding on the future: Evidence against normative discounting of delayed rewards.
911 *Journal of Experimental Psychology: General* **126**:54–70. doi:10.1037/0096-3445.126.1.54

912 Kirby KN, Petry NM. 2004. Heroin and cocaine abusers have higher discount rates for delayed
913 rewards than alcoholics or non-drug-using controls. *Addiction* **99**:461–471. doi:10.1111/j.1360-
914 0443.2003.00669.x

915 Kirby KN, Petry NM, Bickel WK. 1999. Heroin addicts have higher discount rates for delayed rewards
916 than non-drug-using controls. *J Exp Psychol Gen* **128**:78–87. doi:10.1037//0096-3445.128.1.78

917 Kobayashi S, Schultz W. 2008. Influence of reward delays on responses of dopamine neurons. *J*
918 *Neurosci* **28**:7837–7846. doi:10.1523/JNEUROSCI.1600-08.2008

919 Lardeux S, Paleressompoulle D, Pernaud R, Cador M, Baunez C. 2013. Different populations of
920 subthalamic neurons encode cocaine vs. sucrose reward and predict future error. *J Neurophysiol*
921 **110**:1497–1510. doi:10.1152/jn.00160.2013

922 Lardeux S, Pernaud R, Paleressompoulle D, Baunez C. 2009. Beyond the reward pathway: coding
923 reward magnitude and error in the rat subthalamic nucleus. *J Neurophysiol* **102**:2526–2537.
924 doi:10.1152/jn.91009.2008

925 Loewenstein GF, Elster J, Loewenstein G, editors. 1992. Choice over time. New York: Russell Sage
926 Foundation.

927 Loewenstein GF, Prelec D. 1993. Preferences for sequences of outcomes. *Psychological Review*
928 **100**:91–108. doi:10.1037/0033-295X.100.1.91

929 MacKillop J, Tidey JW. 2011. Cigarette demand and delayed reward discounting in nicotine-
930 dependent individuals with schizophrenia and controls: an initial study. *Psychopharmacology (Berl)*
931 **216**:91–99. doi:10.1007/s00213-011-2185-8

932 Madden GJ, Petry NM, Badger GJ, Bickel WK. 1997. Impulsive and self-control choices in opioid-
933 dependent patients and non-drug-using control participants: drug and monetary rewards. *Exp Clin*
934 *Psychopharmacol* **5**:256–262. doi:10.1037//1064-1297.5.3.256

935 Mansfield EL, Karayanidis F, Jamadar S, Heathcote A, Forstmann BU. 2011. Adjustments of response
936 threshold during task switching: a model-based functional magnetic resonance imaging study. *J*
937 *Neurosci* **31**:14688–14692. doi:10.1523/JNEUROSCI.2390-11.2011

938 Martin RF, Bowden DM. 1996. A stereotaxic template atlas of the macaque brain for digital imaging
939 and quantitative neuroanatomy. *Neuroimage* **4**:119–150. doi:10.1006/nimg.1996.0036

940 Mason L, O’Sullivan N, Blackburn M, Bentall R, El-Deredy W. 2012. I want it now! Neural correlates of
941 hypersensitivity to immediate reward in hypomania. *Biol Psychiatry* **71**:530–537.
942 doi:10.1016/j.biopsych.2011.10.008

943 Matsumura M, Kojima J, Gardiner TW, Hikosaka O. 1992. Visual and oculomotor functions of monkey
944 subthalamic nucleus. *J Neurophysiol* **67**:1615–1632.

945 Mazur JE. 2001. Hyperbolic value addition and general models of animal choice. *Psychol Rev* **108**:96–
946 112. doi:10.1037/0033-295x.108.1.96

947 Mazur JE. 1987. An adjusting procedure for studying delayed reinforcementThe Effect of Delay and of
948 Intervening Events on Reinforcement Value, Quantitative Analyses of Behavior, Vol. 5. Hillsdale, NJ,
949 US: Lawrence Erlbaum Associates, Inc. pp. 55–73.

950 McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD. 2007. Time discounting for primary
951 rewards. *J Neurosci* **27**:5796–5804. doi:10.1523/JNEUROSCI.4246-06.2007

952 McClure SM, Laibson DI, Loewenstein G, Cohen JD. 2004. Separate neural systems value immediate
953 and delayed monetary rewards. *Science* **306**:503–507. doi:10.1126/science.1100907

954 McGuire JT, Kable JW. 2015. Medial prefrontal cortical activity reflects dynamic re-evaluation during
955 voluntary persistence. *Nat Neurosci* **18**:760–766. doi:10.1038/nn.3994

956 Mettler FA, Stern GM. 1962. Somatotopic localization in rhesus subthalamic nucleus. *Arch Neurol*
957 **7**:328–329.

958 Minamimoto T, La Camera G, Richmond BJ. 2009. Measuring and modeling the interaction among
959 reward size, delay to reward, and satiation level on motivation in monkeys. *J Neurophysiol* **101**:437–
960 447. doi:10.1152/jn.90959.2008

961 Miocinovic S, Noecker AM, Maks CB, Butson CR, McIntyre CC. 2007. Cicerone: stereotactic
962 neurophysiological recording and deep brain stimulation electrode placement software system. *Acta*
963 *Neurochir Suppl* **97**:561–567.

964 Mitchell SH. 1999. Measures of impulsivity in cigarette smokers and non-smokers.
965 *Psychopharmacology (Berl)* **146**:455–464. doi:10.1007/pl00005491

966 Monakow KH, Akert K, Künzle H. 1978. Projections of the precentral motor cortex and other cortical
967 areas of the frontal lobe to the subthalamic nucleus in the monkey. *Exp Brain Res* **33**:395–403.

968 Nambu A, Tokuno H, Takada M. 2002. Functional significance of the cortico-subthalamo-pallidal
969 “hyperdirect” pathway. *Neurosci Res* **43**:111–117.

970 Nougaret S, Baunez C, Ravel S. 2022. Neurons in the Monkey’s Subthalamic Nucleus Differentially
971 Encode Motivation and Effort. *J Neurosci* **42**:2539–2551. doi:10.1523/JNEUROSCI.0281-21.2021

972 Parent A, Hazrati LN. 1995. Functional anatomy of the basal ganglia. II. The place of subthalamic
973 nucleus and external pallidum in basal ganglia circuitry. *Brain Res Brain Res Rev* **20**:128–154.

974 Pasquereau B, Turner RS. 2017. A selective role for ventromedial subthalamic nucleus in inhibitory
975 control. *Elife* **6**. doi:10.7554/eLife.31627

976 Pasquereau B, Turner RS. 2015. Dopamine neurons encode errors in predicting movement trigger
977 occurrence. *J Neurophysiol* **113**:1110–1123. doi:10.1152/jn.00401.2014

978 Pasquereau B, Turner RS. 2013. Limited encoding of effort by dopamine neurons in a cost-benefit
979 trade-off task. *J Neurosci* **33**:8288–8300. doi:10.1523/JNEUROSCI.4619-12.2013

980 Pelloux Y, Degoulet M, Tiran-Cappello A, Cohen C, Lardeux S, George O, Koob GF, Ahmed SH, Baunez
981 C. 2018. Subthalamic nucleus high frequency stimulation prevents and reverses escalated cocaine
982 use. *Mol Psychiatry* **23**:2266–2276. doi:10.1038/s41380-018-0080-y

983 Perdue BM, Bramlett JL, Evans TA, Beran MJ. 2015. Waiting for what comes later: capuchin monkeys
984 show self-control even for nonvisible delayed rewards. *Anim Cogn* **18**:1105–1112.
985 doi:10.1007/s10071-015-0878-9

986 Peters J, Büchel C. 2009. Overlapping and distinct neural systems code for subjective value during
987 intertemporal and risky decision making. *J Neurosci* **29**:15727–15734. doi:10.1523/JNEUROSCI.3489-
988 09.2009

989 Pine A, Seymour B, Roiser JP, Bossaerts P, Friston KJ, Curran HV, Dolan RJ. 2009. Encoding of marginal
990 utility across time in the human brain. *J Neurosci* **29**:9575–9581. doi:10.1523/JNEUROSCI.1126-
991 09.2009

992 Pulcu E, Trotter PD, Thomas EJ, McFarquhar M, Juhasz G, Sahakian BJ, Deakin JFW, Zahn R, Anderson
993 IM, Elliott R. 2014. Temporal discounting in major depressive disorder. *Psychol Med* **44**:1825–1834.
994 doi:10.1017/S0033291713002584

995 Rachlin H. 2004. The science of self-control, 2nd ed. ed. Cambridge (Mass.): Harvard university press.

996 Reynolds B. 2006. A review of delay-discounting research with humans: relations to drug use and
997 gambling. *Behav Pharmacol* **17**:651–667. doi:10.1097/FBP.0b013e3280115f99

998 Roesch MR, Bryden DW. 2011. Impact of size and delay on neural activity in the rat limbic
999 corticostriatal system. *Front Neurosci* **5**:130. doi:10.3389/fnins.2011.00130

1000 Roesch MR, Calu DJ, Burke KA, Schoenbaum G. 2007. Should I stay or should I go? Transformation of
1001 time-discounted rewards in orbitofrontal cortex and associated brain circuits. *Ann N Y Acad Sci*
1002 **1104**:21–34. doi:10.1196/annals.1390.001

1003 Roesch MR, Olson CR. 2005. Neuronal activity dependent on anticipated and elapsed delay in
1004 macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *J*
1005 *Neurophysiol* **94**:1469–1497. doi:10.1152/jn.00064.2005

1006 Roesch MR, Taylor AR, Schoenbaum G. 2006. Encoding of time-discounted rewards in orbitofrontal
1007 cortex is independent of value representation. *Neuron* **51**:509–520.
1008 doi:10.1016/j.neuron.2006.06.027

1009 Rouaud T, Lardeux S, Panayotis N, Paleressompoulle D, Cador M, Baunez C. 2010. Reducing the desire
1010 for cocaine with subthalamic nucleus deep brain stimulation. *Proc Natl Acad Sci USA* **107**:1196–1200.
1011 doi:10.1073/pnas.0908189107

1012 Rounds JS, Beck JG, Grant DM. 2007. Is the delay discounting paradigm useful in understanding social
1013 anxiety? *Behav Res Ther* **45**:729–735. doi:10.1016/j.brat.2006.06.007

1014 Sadikot AF, Parent A, François C. 1992. Efferent connections of the centromedian and parafascicular
1015 thalamic nuclei in the squirrel monkey: a PHA-L study of subcortical projections. *J Comp Neurol*
1016 **315**:137–159. doi:10.1002/cne.903150203

1017 Scheres A, Tontsch C, Thoeny AL, Kaczurkin A. 2010. Temporal reward discounting in attention-
1018 deficit/hyperactivity disorder: the contribution of symptom domains, reward magnitude, and session
1019 length. *Biol Psychiatry* **67**:641–648. doi:10.1016/j.biopsych.2009.10.033

1020 Seinstra M, Wojtecki L, Storzer L, Schnitzler A, Kalenscher T. 2016. No Effect of Subthalamic Deep
1021 Brain Stimulation on Intertemporal Decision-Making in Parkinson Patients. *eNeuro* **3**:ENEURO.0019-
1022 16.2016. doi:10.1523/ENEURO.0019-16.2016

1023 Seymour B, Barbe M, Dayan P, Shiner T, Dolan R, Fink GR. 2016. Deep brain stimulation of the
1024 subthalamic nucleus modulates sensitivity to decision outcome value in Parkinson’s disease. *Sci Rep*
1025 **6**:32509. doi:10.1038/srep32509

1026 Shink E, Bevan MD, Bolam JP, Smith Y. 1996. The subthalamic nucleus and the external pallidum: two
1027 tightly interconnected structures that control the output of the basal ganglia in the monkey.
1028 *Neuroscience* **73**:335–357.

1029 Sieger T, Serranová T, Růžička F, Vostatek P, Wild J, Štastná D, Bonnet C, Novák D, Růžička E, Urgošík
1030 D, Jech R. 2015. Distinct populations of neurons respond to emotional valence and arousal in the
1031 human subthalamic nucleus. *Proc Natl Acad Sci U S A* **112**:3116–3121. doi:10.1073/pnas.1410709112

1032 Szalda-Petree AD, Craft BB, Martin LM, Deditius-Island HK. 2004. Self-control in rhesus macaques
1033 (*Macaca mulatta*): controlling for differential stimulus exposure. *Percept Mot Skills* **98**:141–146.
1034 doi:10.2466/pms.98.1.141-146

1035 Takahashi T, Oono H, Inoue T, Boku S, Kako Y, Kitaichi Y, Kusumi I, Masui T, Nakagawa S, Suzuki K,
1036 Tanaka T, Koyama T, Radford MHB. 2008. Depressive patients are more impulsive and inconsistent in
1037 intertemporal choice behavior for monetary gain and loss than healthy subjects--an analysis based
1038 on Tsallis’ statistics. *Neuro Endocrinol Lett* **29**:351–358.

1039 Tan H, Pogosyan A, Ashkan K, Cheeran B, Fitzgerald JJ, Green AL, Aziz T, Foltynie T, Limousin P, Zrinzo
1040 L, Brown P. 2015. Subthalamic nucleus local field potential activity helps encode motor effort rather
1041 than force in parkinsonism. *J Neurosci* **35**:5941–5949. doi:10.1523/JNEUROSCI.4609-14.2015

1042 Tanaka D, Aoki R, Suzuki S, Takeda M, Nakahara K, Jimura K. 2020. Self-Controlled Choice Arises from
1043 Dynamic Prefrontal Signals That Enable Future Anticipation. *J Neurosci* **40**:9736–9750.
1044 doi:10.1523/JNEUROSCI.1702-20.2020

- Teagarden MA, Rebec GV. 2007. Subthalamic and striatal neurons concurrently process motor, limbic, and associative information in rats performing an operant task. *J Neurophysiol* **97**:2042–2058. doi:10.1152/jn.00368.2006
- Tripp G, Alsop B. 2001. Sensitivity to reward delay in children with attention deficit hyperactivity disorder (ADHD). *J Child Psychol Psychiatry* **42**:691–698.
- Uslaner JM, Robinson TE. 2006. Subthalamic nucleus lesions increase impulsive action and decrease impulsive choice - mediation by enhanced incentive motivation? *Eur J Neurosci* **24**:2345–2354. doi:10.1111/j.1460-9568.2006.05117.x
- Vanderveldt A, Oliveira L, Green L. 2016. Delay discounting: Pigeon, rat, human--does it matter? *J Exp Psychol Anim Learn Cogn* **42**:141–162. doi:10.1037/xan0000097
- Voon V, Droux F, Morris L, Chabardes S, Bougerol T, David O, Krack P, Polosan M. 2017. Decisional impulsivity and the associative-limbic subthalamic nucleus in obsessive-compulsive disorder: stimulation and connectivity. *Brain* **140**:442–456. doi:10.1093/brain/aww309
- Voon V, Kubu C, Krack P, Houeto J-L, Tröster AI. 2006. Deep brain stimulation: neuropsychological and neuropsychiatric issues. *Mov Disord* **21 Suppl 14**:S305–327. doi:10.1002/mds.20963
- Vuchinich RE, Simpson CA. 1998. Hyperbolic temporal discounting in social drinkers and problem drinkers. *Exp Clin Psychopharmacol* **6**:292–305. doi:10.1037//1064-1297.6.3.292
- Wade CL, Kallupi M, Hernandez DO, Breyse E, de Guglielmo G, Crawford E, Koob GF, Schweitzer P, Baunez C, George O. 2017. High-Frequency Stimulation of the Subthalamic Nucleus Blocks Compulsive-Like Re-Escalation of Heroin Taking in Rats. *Neuropsychopharmacology* **42**:1850–1859. doi:10.1038/npp.2016.270
- Washio Y, Higgins ST, Heil SH, McKerchar TL, Badger GJ, Skelly JM, Dantona RL. 2011. Delay discounting is associated with treatment response among cocaine-dependent outpatients. *Exp Clin Psychopharmacol* **19**:243–248. doi:10.1037/a0023617
- Wiener M, Magaro CM, Matell MS. 2008. Accurate timing but increased impulsivity following excitotoxic lesions of the subthalamic nucleus. *Neurosci Lett* **440**:176–180. doi:10.1016/j.neulet.2008.05.071
- Winstanley CA, Baunez C, Theobald DEH, Robbins TW. 2005. Lesions to the subthalamic nucleus decrease impulsive choice but impair autoshaping in rats: the importance of the basal ganglia in Pavlovian conditioning and impulse control. *Eur J Neurosci* **21**:3107–3116. doi:10.1111/j.1460-9568.2005.04143.x
- Yamada H, Imaizumi Y, Matsumoto M. 2021. Neural Population Dynamics Underlying Expected Value Computation. *J Neurosci* **41**:1684–1698. doi:10.1523/JNEUROSCI.1987-20.2020
- Zavala B, Zaghoul K, Brown P. 2015. The subthalamic nucleus, oscillations, and conflict. *Mov Disord* **30**:328–338. doi:10.1002/mds.26072
- Zénon A, Duclos Y, Carron R, Witjas T, Baunez C, Régis J, Azulay J-P, Brown P, Eusebio A. 2016. The human subthalamic nucleus encodes the subjective value of reward and the cost of effort during decision-making. *Brain* **139**:1830–1843. doi:10.1093/brain/aww075

FIGURE LEGENDS

Figure 1. Delayed reward task and behavioral performance. **(A)** Temporal sequence of task events. After the monkey initiated a trial by positioning a cursor (+) within a visual target (gray circle), an instruction cue was presented briefly signaling the reward size (1, 2, or 3 drops of food) and the delay-to-reward (short or long). The animal was required to maintain the cursor position over the waiting period to successfully obtain reward. **(B-C)** Rejection rates (mean \pm SEM) were calculated and averaged for the six possible reward contingencies across sessions. Measures were affected by both reward size and delay (two-way ANOVA). Size: *** $P < 0.001$ ** $P < 0.01$; Delay: ### $P < 0.001$ ## $P < 0.01$. **(D-G)** For each animal, a temporal discount factor (k) was found that yielded the best fit between averaged rejection rates and the hyperbolic model expressed by Eq. 2. Goodness of fit was evaluated by the coefficient of determination (R^2). **(H-I)** EMG signals collected in monkey H were aligned on the presentation of cues. The effects of reward size and delay were examined using a series of two-way ANOVAs. Red lines indicate the statistical threshold ($P < 0.05/173$; Bonferroni correction).

Figure 2. STN neurons were modulated by reward size and delay. **(A-B)** Reconstruction of a trajectory used for STN recordings with a structural MRI and high resolution 3-D templates of individual nuclei derived from (Martin and Bowden, 1996)). Globus pallidus (GP), Substantia nigra (SN), Zona incerta (ZI), Thalamus (THL). **(C)** Sample of action potential waveforms emitted by STN neurons. **(D)** Color map histograms of neuronal activities recorded from the STN. Each horizontal line indicates neural activity aligned to instruction cues averaged across trial types. Neuronal firing rates were Z-score normalized. **(E)** Population-averaged activities of STN neurons, and **(F)** Fano factors that showed the variability of the neural population ensemble across the six possible reward contingencies. The width of the curves indicates the population SEM. **(G-L)** Influence of reward size and delay on individual neural activities was detected by a series of two-way ANOVAs ($P < 0.05/173$, Bonferroni correction). The time course of encoding of task-relevant information (*left column*) and the fractions of neurons modulated by reward size and/or delay (*right column*) were represented for each time bin. Pie charts show the total fraction of STN neurons influenced by reward size (*blue*), delay (*green*), or both task parameters simultaneously (*gray*).

Figure 3. Response of STN neurons to the six possible reward contingencies. **(A)** The activity of three exemplar neurons that were classified as task-related cells. Spike density functions and raster plots were constructed separately around the presentation of instruction cues for the different cost-benefit conditions. **(B)** A sliding window regression analysis compared firing rates between trial types

(as expressed by Eq. 3). The regression coefficients (yellow-to-black lines) were used to characterize the dynamic encoding of reward size (β_{Reward}) and delay (β_{Delay}). The horizontal dashed lines indicate the statistical threshold for significant β values (calculated from the pre-instruction period with a one sample t -test, $df = 46$, $P < 0.05$). (C) Time series of regression coefficients projected into an orthogonal space where reward size and delay composed the two dimensions. Vector time series were produced for significant β values. Black dashed lines indicate statistical thresholds. The angle (θ) of the vector sum (*red dashed lines*) was calculated to identify how neurons integrated cost-benefit conditions during the two consecutive phases of the waiting period (phase 1: 0-2s, phase 2: 2-3.5s).

Figure 4. STN neurons exhibit mixed signals in phase 1 (0-2s post-instruction) and phase 2 (2-3.5s post-instruction). (A) Schematic depiction of the regression subspace composed of reward size and delay. Various patterns of neural encoding could be categorized depending on the angle (θ) of vectors: Discounting- (between -90° and 0°); Discounting+ (between 90° and 180°); Compounding+ (between 0° and 90°); Compounding- (between -180° and -90°). (B) The angle differences calculated between phases 1 and 2 ($\theta_2 - \theta_1$) shows how the neural encodings were modified during the course of the hold period. A positive angle difference corresponds to a turn anticlockwise, while a negative one corresponds to a turn clockwise. (C-D) Vectorial encoding of reward size and delay for all task-related neurons in phases 1 (C) and 2 (D). Vector sums were calibrated by subtracting the mean β values of the pre-instruction epoch and then dividing by 2 SD of this control period. The red dashed lines indicate the population vectors. (E-F) Fractions of task-related neurons categorized as Discounting cells (Disc-, Disc+) or Compounding cells (Comp+, Comp-). (G-H) Vector magnitudes (mean \pm SEM) were compared between different categories of task-related neurons (one-way ANOVA $*F > 2.1$ $P < 0.05$). The central line of the box plots represents the median, the edges of the box show the interquartile range, and the edges of the whiskers show the full extent of the overall distributions.

Figure 5. Topography of STN neurons categorized as Discounting- in phase 1 and Compounding+ in phase 2. (A) Three-dimensional plots of cell type distributions based on coordinates from the recording chamber. The template of STN (*gray surface*) is derived from the atlas Martin and Bowden (1996). (B-D) Comparisons of cell positions show that Discounting- neurons (Disc-) were located more dorsal and posterior than other categories of cell type (two sample t -test). (E-G) No differences were found in the tonic rate (E), spike magnitude (F), and spike duration (G) between Discounting- neurons and other cells. The central line of the box plots represents the median, the edges of the box show

the interquartile range, and the edges of the whiskers show the full extent of the overall distributions. **(H-N)** Neurons categorized as Compounding+ in phase 2 were located more dorsal and posterior than other cell types.

Figure 6. Neural population ensemble provides dynamic integration of reward size and delay. **(A)** Cumulative variance explained by PCA in the population of all recorded neurons. Dashed lines indicate the percentages of variance explained by the first four PCs: PC1 (*cyan*), PC2 (*red*), PC3 (*green*) and PC4 (*yellow*). **(B)** Cumulative variance explained by a control shuffled procedure (data shuffled 1000 times). **(C)** Series of eigenvectors produced for the first four PCs. Eigenvectors capture the moment-by-moment signals in the subspace composed of reward size and delay. Colors indicated the eigenvectors with a significant magnitude relative to those calculated from the surrogate control PCAs. **(D-E)** For each PC, eigenvector magnitudes captured the extent of the signals dynamically transmitted **(D)**, while angles (-180 to 180°) indicated how the ensemble integrates reward size and delay **(E)**. **(F)** Examples of eigenvectors produced by the control shuffled procedure. Percentages of variance explained by the first four PCs were significantly higher than chance (permutation test, $P > 0.05$).

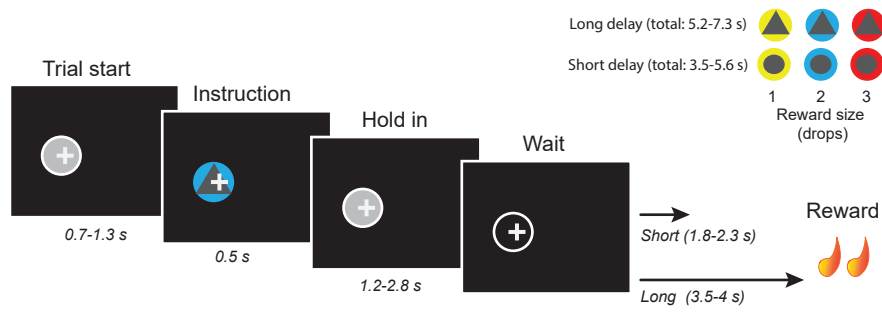
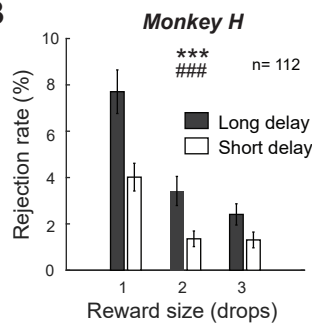
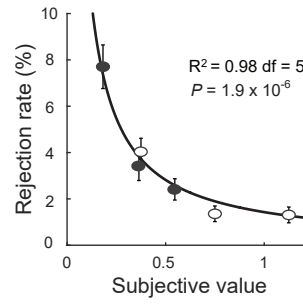
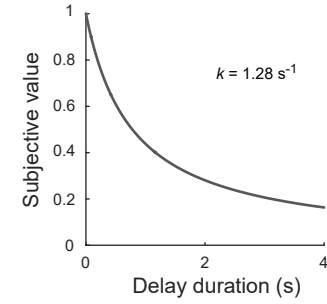
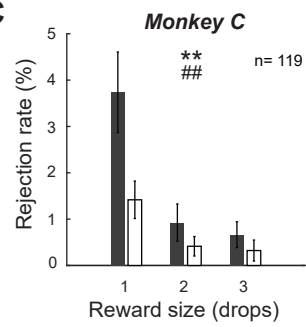
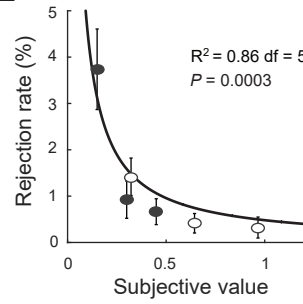
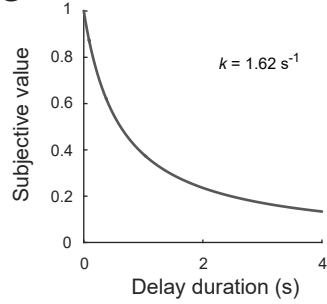
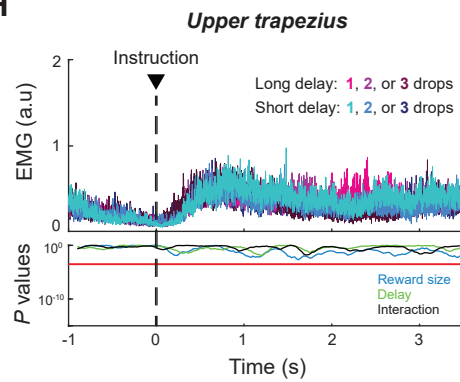
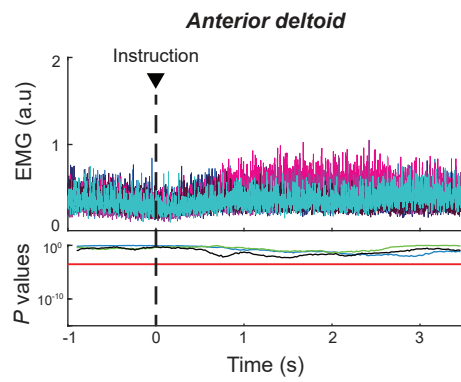
Figure 7. Signals vary along the antero-posterior axis of the STN. **(A-B)** Correlations between component scores [PC1 (*cyan*) and PC2 (*red*)] of the neurons and their anatomical position. **(C-D)** Anatomical axis corresponding the most to PC1 **(C)** and PC2 **(D)**. Arrows show direction of increasing component scores.

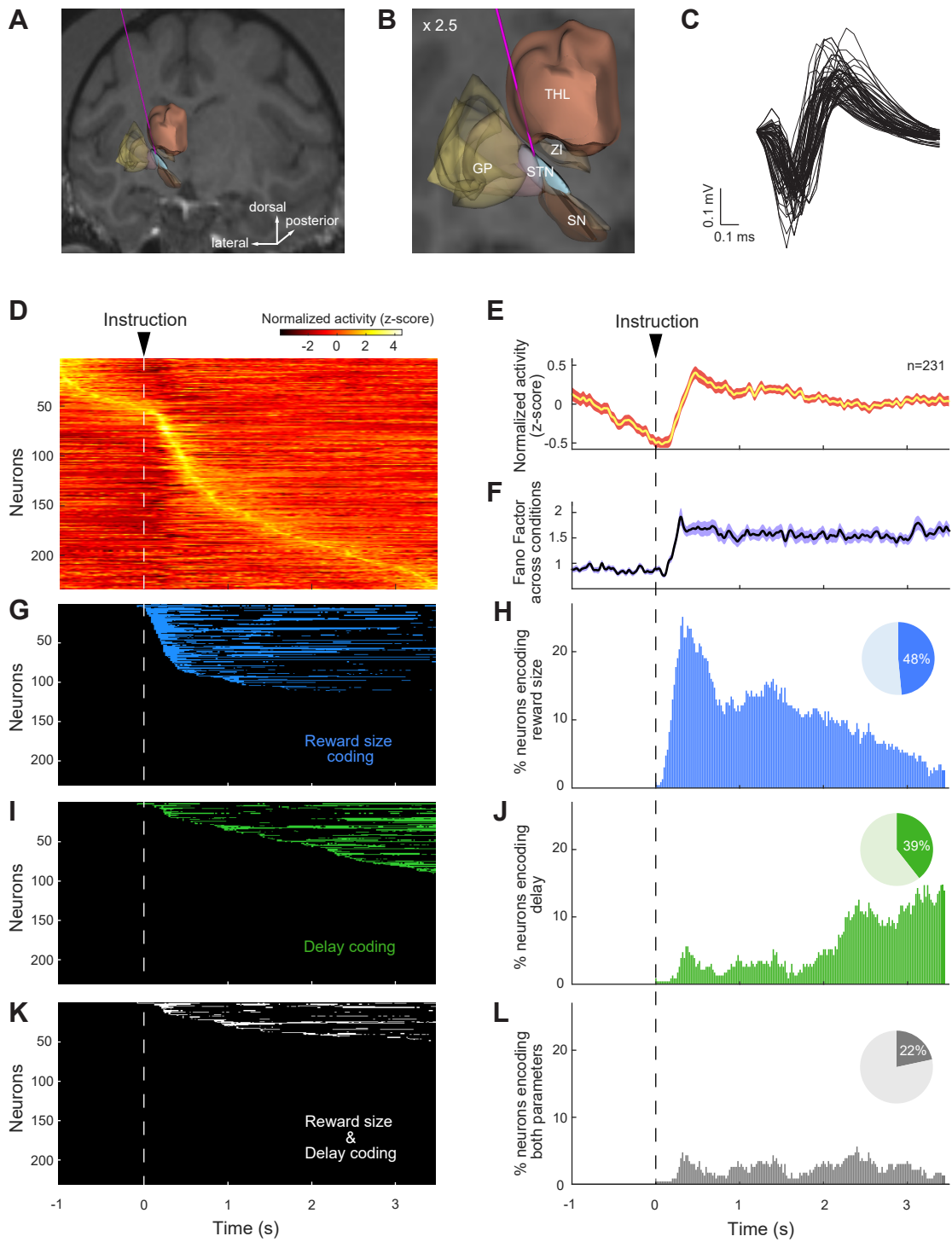
SUPPLEMENTAL FIGURE LEGEND

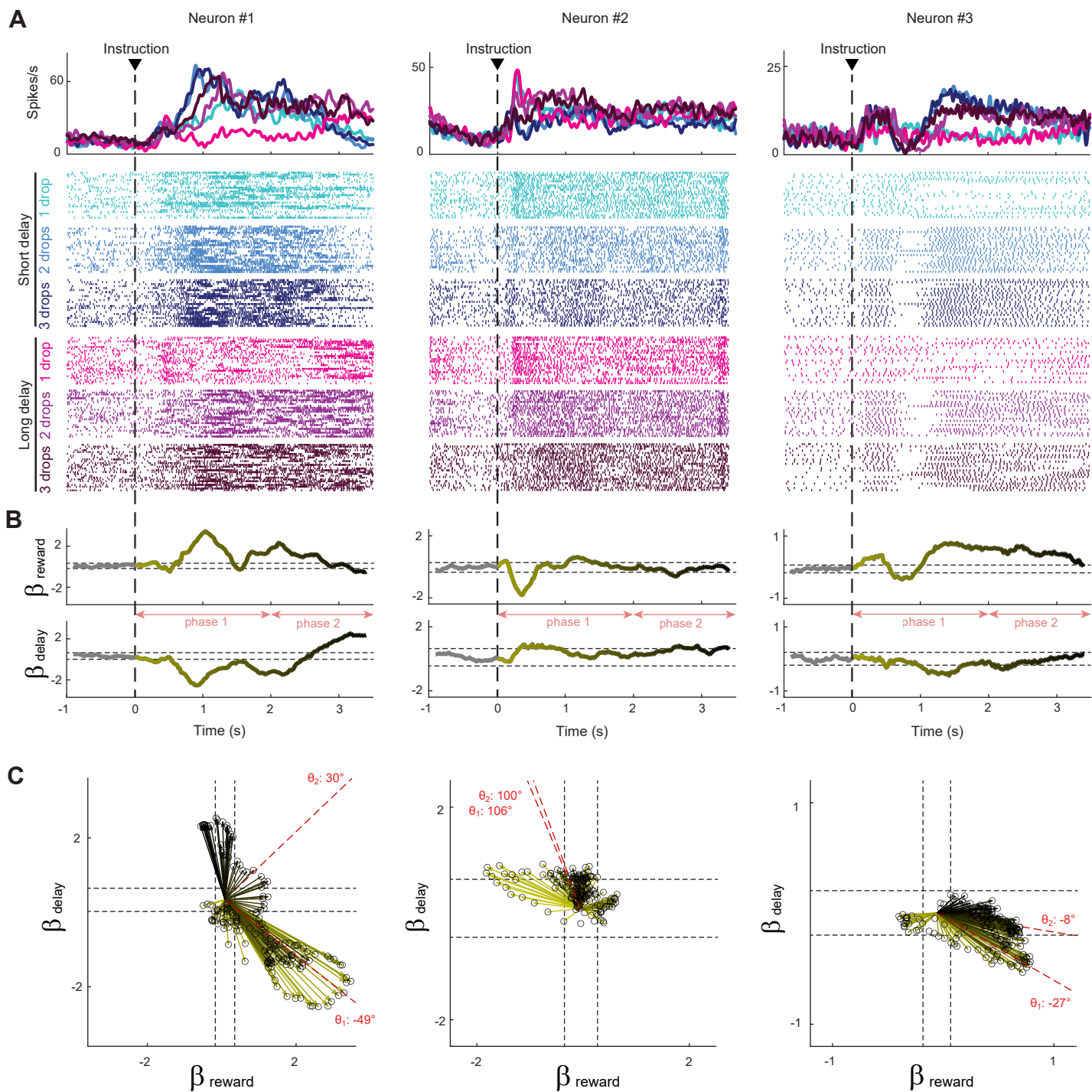
Figure 1 – supplement 1. Eye positions varied according to the levels of reward and delay. For each task condition, horizontal and vertical components were combined to obtain tangential coordinates (averaged across sessions; \pm SEM). Potential interaction with task parameters was tested using a series of two-way ANOVAs (Reward size x Delay). Red lines indicate the statistical threshold ($P < 0.05/173$, Bonferroni correction).

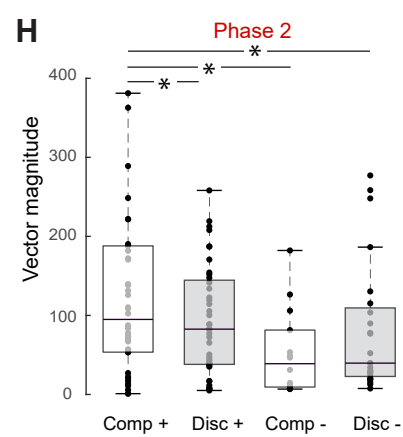
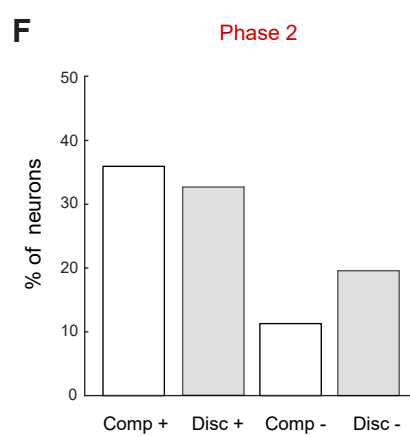
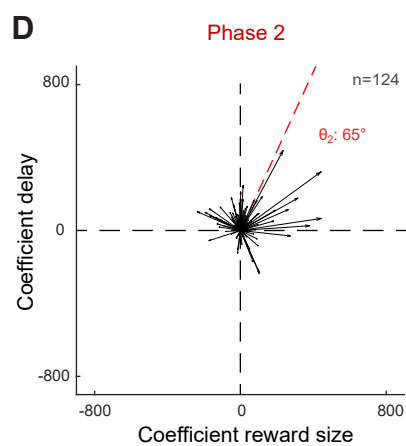
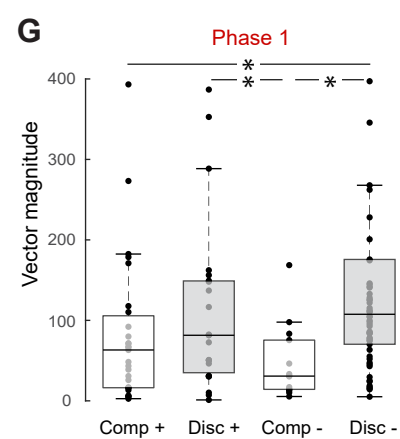
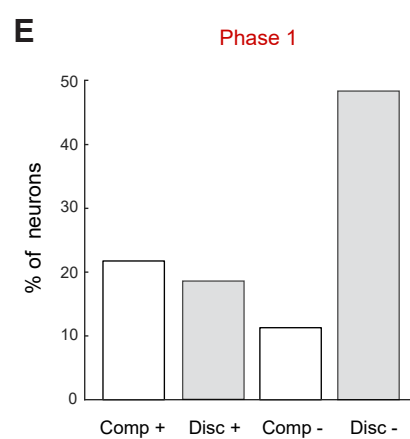
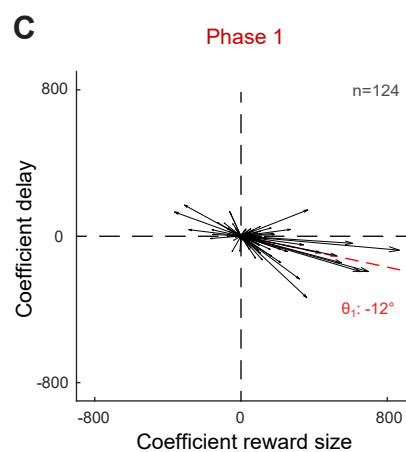
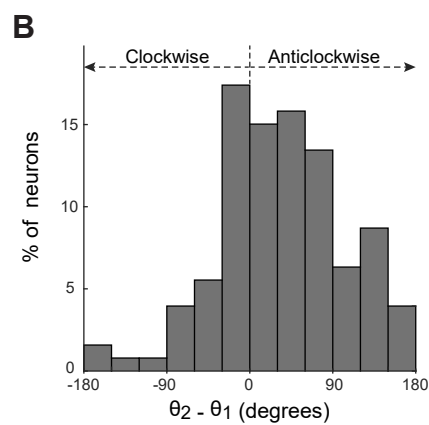
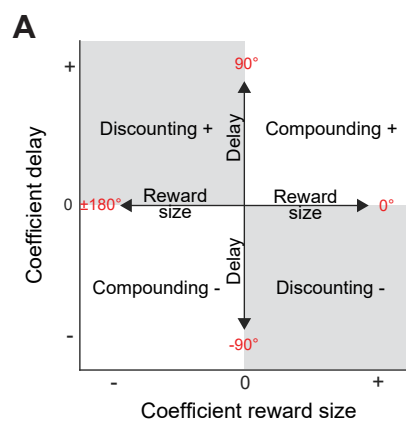
Figure 4 – supplement 1. Population vectors of the encoding of reward size and delay in phase 1 (**A**) and phase 2 (**B**) using a regression model that excluded eye movement variables ($SC_i = \beta_0 + \beta_R R_i + \beta_D D_i$). These figures follow the conventions of Figure 4C-D. Individual vectors and mean population vectors produced here were very similar to results of Figure 4 (differences $<3^\circ$).

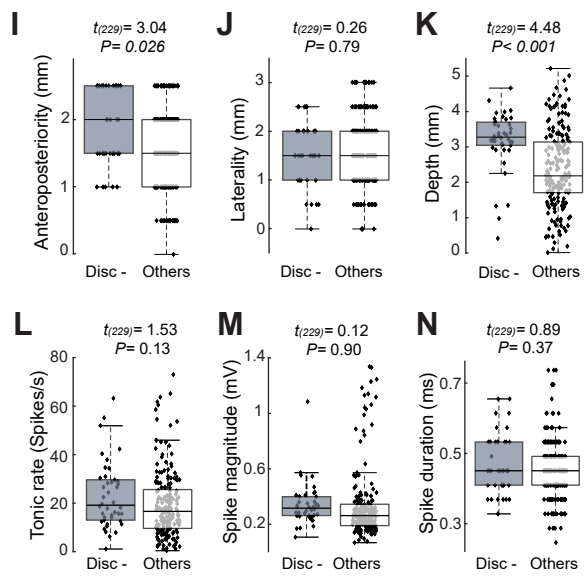
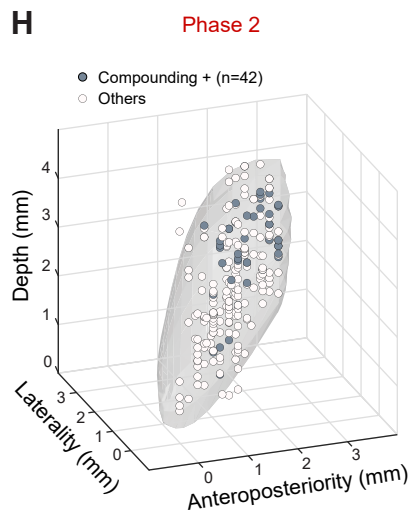
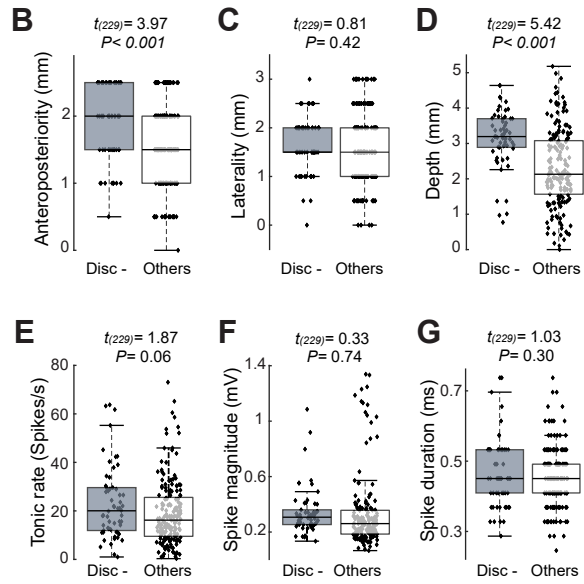
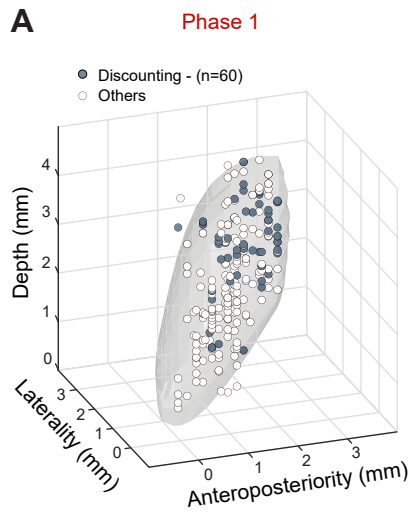
Figure 6 – supplement 1. Population-based analysis performed using a sliding window method in the estimation of regression coefficients (200-ms test window stepped in 20-ms). These figures follow the conventions of Figure 6. (**A-B**) Cumulative variance explained by PCA in the population of all recorded neurons (**A**), or by the control shuffled procedure (**B**). (**C**) Series of eigenvectors produced for the first four PCs: PC1 (cyan), PC2 (red), PC3 (green) and PC4 (yellow). Eigenvectors capture the moment-by-moment signals in the subspace composed of reward size and delay. Colors indicated the eigenvectors with a significant magnitude relative to those calculated from the population of surrogate control PCs (data shuffled 1000 times). (**D-E**) For each PC, eigenvector magnitudes captured the strength of the signals transmitted (**D**), while angles (-180 to 180°) indicated how the ensemble integrates reward size and delay (**E**). (**F**) Examples of eigenvectors produced by the control shuffled procedure. Note that the encoding patterns and the percents of variance explained here were very similar to results shown in Figure 6.

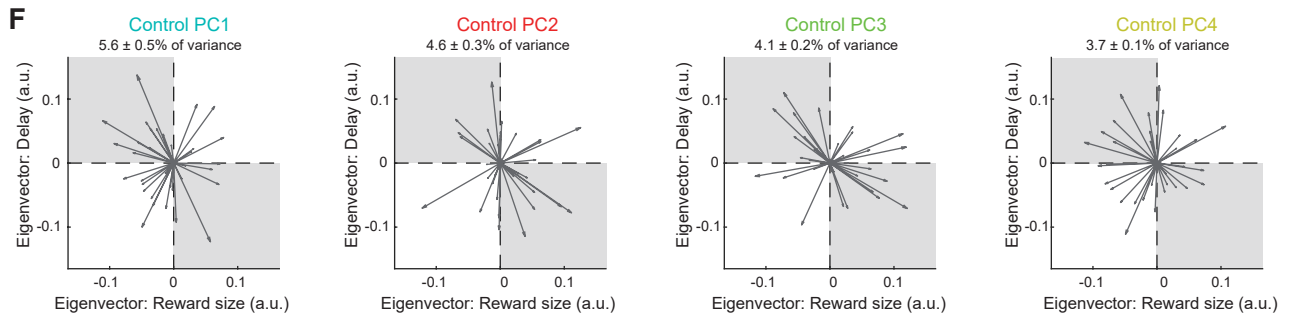
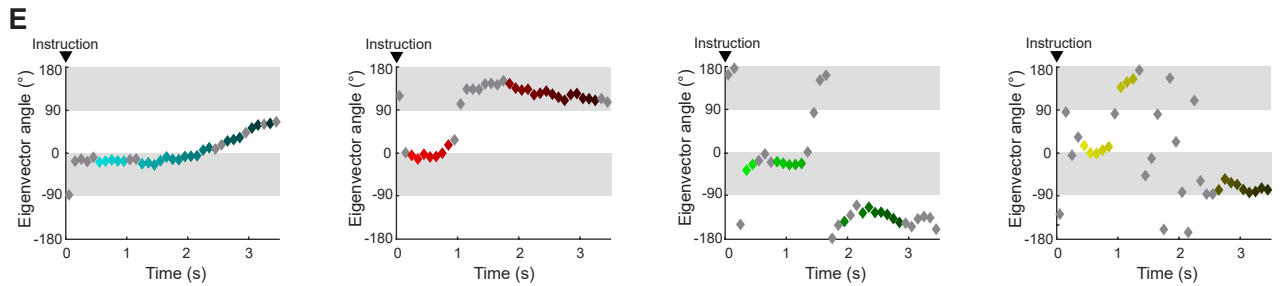
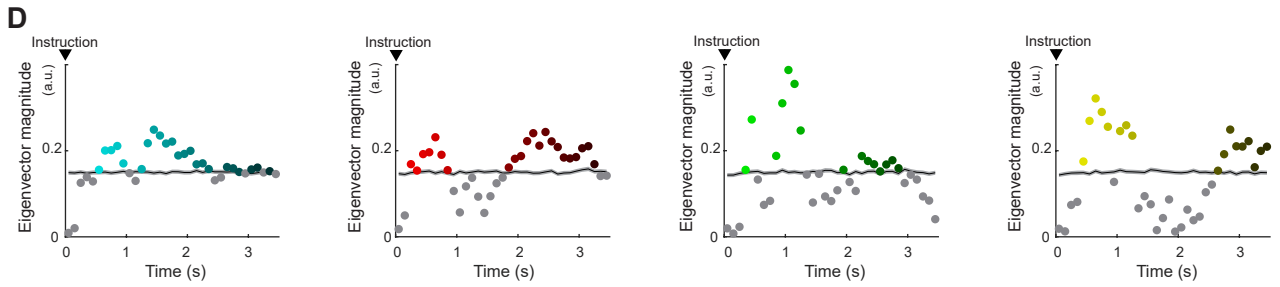
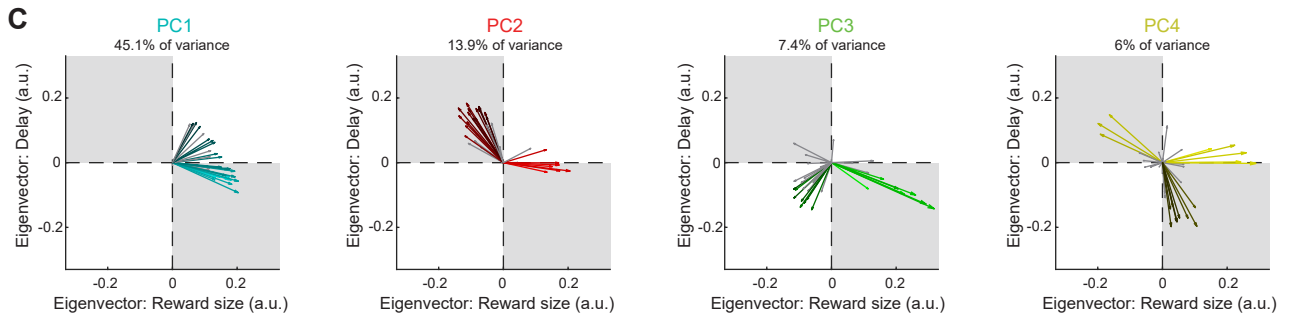
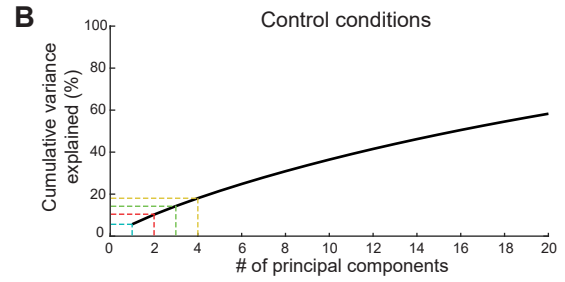
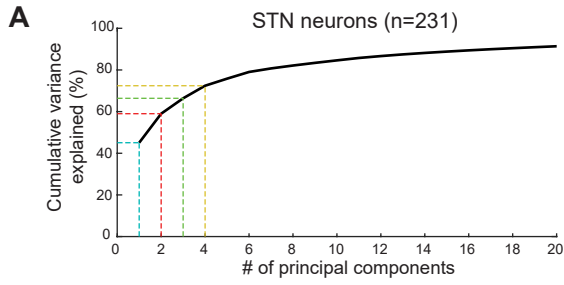
A**B****D****F****C****E****G****H****I**

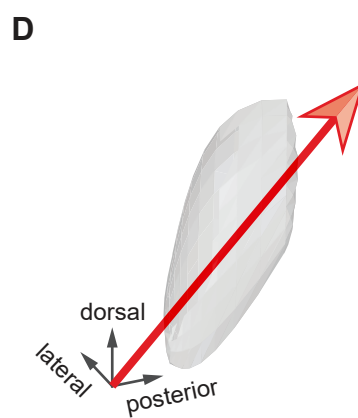
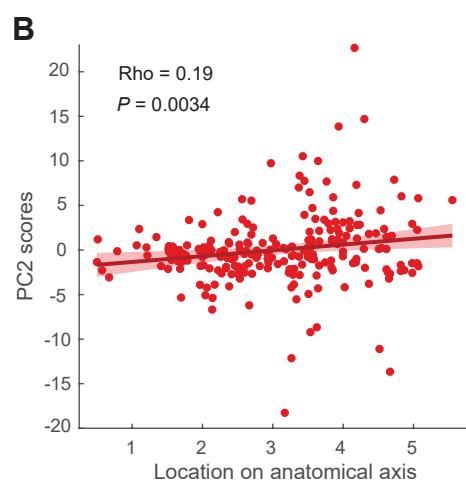
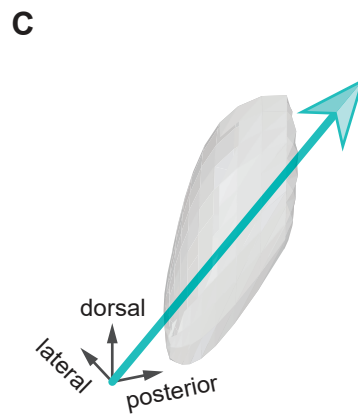
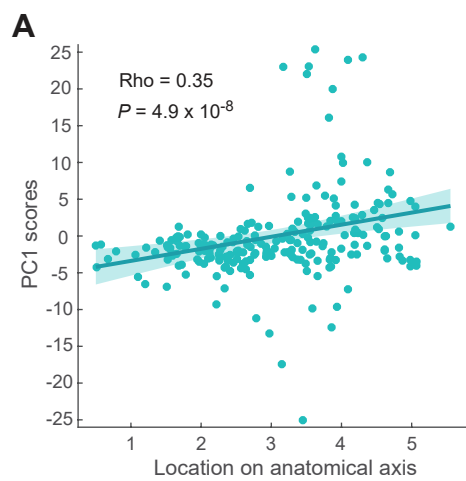




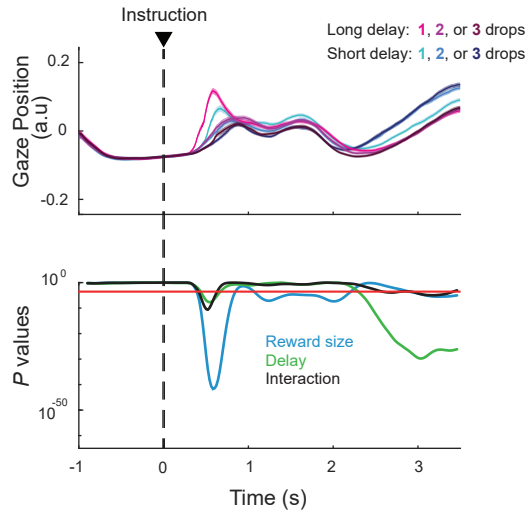




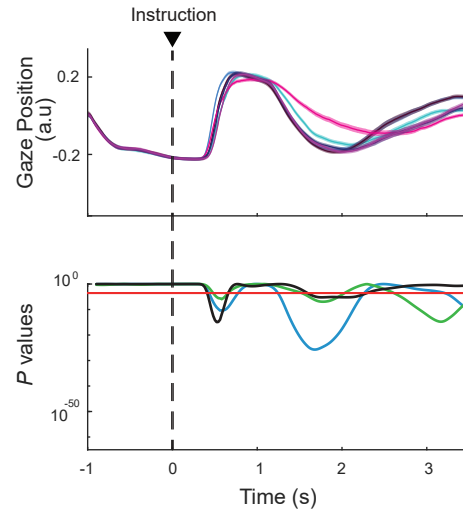


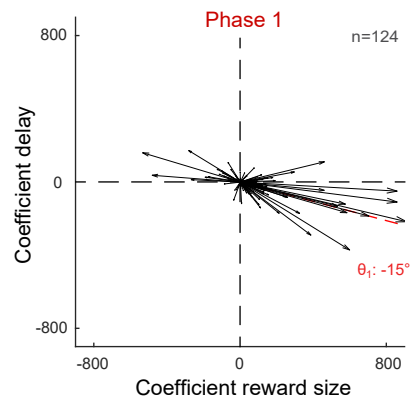


Monkey H
n= 112



Monkey C
n= 119



A**B**