The subthalamic nucleus contributes causally to perceptual decision-making in monkeys

Kathryn Rogers, Joshua I. Gold, Long Ding[*]

Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104

* Corresponding author (lding@pennmedicine.upenn.edu)

Abstract

The subthalamic nucleus (STN) plays critical roles in the motor and cognitive function of the basal ganglia (BG), but the exact nature of these roles is not fully understood, especially in the context of decision-making based on uncertain evidence. Guided by theoretical predictions of specific STN contributions, we used single-unit recording and electrical microstimulation in the STN of healthy monkeys to assess its causal, computational roles in visual-saccadic decisions based on noisy evidence. The recordings identified subpopulations of STN neurons with distinct task-related activity patterns that related to different theoretically predicted functions. Microstimulation caused changes in behavioral choices and response times that reflected multiple contributions to an "accumulate-to-bound"-like decision process, including modulation of decision bounds and evidence accumulation, and to non-perceptual processes. These results provide new insights into the multiple ways that the STN can support higher brain function.
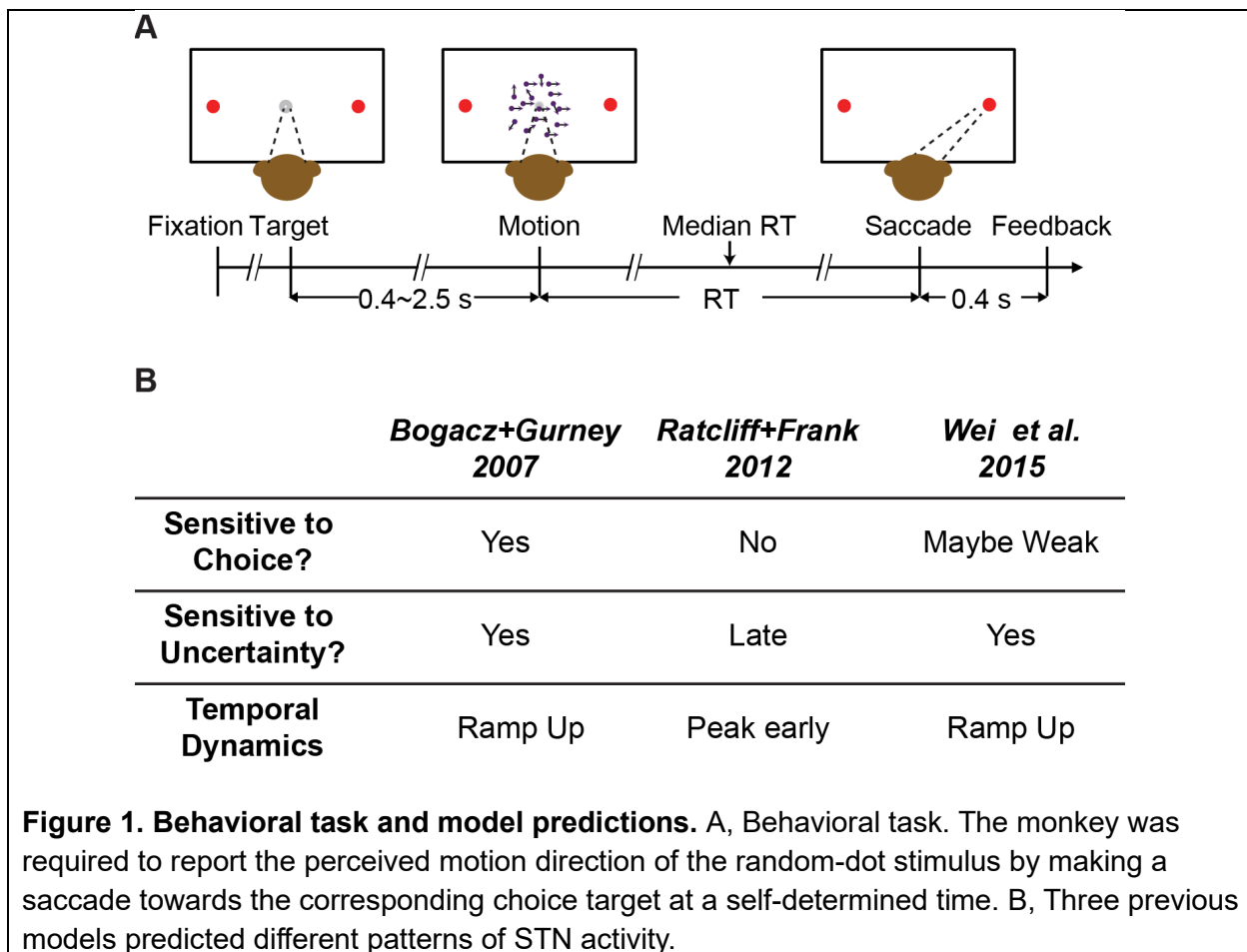
Introduction

The subthalamic nucleus (STN) is a critical junction in both the indirect and hyperdirect pathways of the basal ganglia (BG). It receives inputs from the external segment of the globus pallidum (GPe) and cortex and sends diffuse excitation to pallidal output nuclei of the BG. The STN has well-recognized functions in movement control. For example, in humans and monkeys, lesions of the STN cause involuntary movements of contralateral body parts (Martin, 1927; Martin and Alcock, 1934; Whittier and Mettler, 1949; Carpenter et al., 1950). In monkeys with experimentally induced Parkinsonism, STN lesions and inactivation can reverse abnormal BG output activity and alleviate both akinesia and rigidity (Bergman et al., 1990, 1994; Wichmann et al., 1994a). In Parkinsonian human patients, deep brain stimulation (DBS) of the STN has become a common treatment option to alleviate movement abnormalities (DeLong and Wichmann, 2001).

Recognizing that motor symptoms associated with STN damage are often accompanied by emotional and cognitive deficits, recent work has also begun to examine the roles of the STN in cognition. For example, the STN has been shown to contribute to cued, goal-driven action inhibition (Baunez et al., 2001; Desbonnet et al., 2004; Witt et al., 2004; Aron and Poldrack, 2006; Frank et al., 2007; Isoda and Hikosaka, 2008; Schmidt et al., 2013; Pasquereau and Turner, 2017). STN activity can also be sensitive to task complexity and decision conflict, as measured in imaging studies and human patients undergoing DBS (Lehericy et al., 2004; Aron et al., 2007; Fumagalli et al., 2011; Brittain et al., 2012; Zaghloul et al., 2012; Zavala et al., 2017). These findings have led to the idea that STN may also contribute to resolving difficult decisions based on uncertain evidence. This idea has been formalized in several computational models, which posit three, not mutually exclusive, functions for STN: 1) in coordination with the medial prefrontal cortex, STN adjusts decision bounds (i.e., thresholds on accumulated evidence that govern decision termination and commitment) to control impulsivity in responding (Frank, 2006; Cavanagh et al., 2011; Ratcliff and Frank, 2012; Zavala et al., 2014; Herz et al., 2016, 2017; Pote

41 et al., 2016); 2) through its interaction with GPe, STN computes a normalization signal to
42 calibrate how the available, alternative options are assessed (Bogacz and Gurney, 2007;
43 Coulthard et al., 2012; Green et al., 2013); and 3) by maintaining the balance between the direct
44 and indirect pathways of the BG, STN helps to implement a nonlinear computation that improves
45 the efficacy with which the BG adjusts decision bounds (Lo and Wang, 2006; Wei et al., 2015).

46 Guided by predictions of these models (Figure 1B), we assessed the role of the STN in decisions
47 made by monkeys performing a random-dot visual motion direction discrimination task (Figure.
48 1A). We recorded from individual STN neurons while monkeys performed the task and found
49 activity patterns that were highly heterogenous across neurons. Nevertheless, these patterns
50 could be sorted into three prominent clusters with functional properties that, in principle, could
51 support each of the three theoretically predicted STN functions from previous modeling studies.
52 In addition, we tested STN's causal contribution to the decision process using electrical
53 microstimulation. These perturbations of STN activity affected both choice and reaction time
54 (RT) performance in multiple ways that could be ascribed to particular computational
55 components of an "accumulate-to-bound" decision process. As detailed below, these results show
56 that STN can play multiple, causal roles in the formation of a deliberative perceptual decision,
57 likely reflecting its diverse contributions to the many cognitive and motor functions that depend
58 on the BG.

	Bogacz+Gurney 2007	Ratcliff+Frank 2012	Wei et al. 2015
Sensitive to Choice?	Yes	No	Maybe Weak
Sensitive to Uncertainty?	Yes	Late	Yes
Temporal Dynamics	Ramp Up	Peak early	Ramp Up

Figure 1. Behavioral task and model predictions. A, Behavioral task. The monkey was required to report the perceived motion direction of the random-dot stimulus by making a saccade towards the corresponding choice target at a self-determined time. B, Three previous models predicted different patterns of STN activity.

Results

STN neurons show diverse response profiles

We recorded 203 neurons while the monkeys were performing a random-dot motion discrimination task ($n = 115$ and 88 for monkeys C and F, respectively). The behavioral performance of both monkeys has been documented extensively (Ding and Gold, 2010, 2012a; Fan et al., 2018). Their performance in three example sessions are shown in Figure 4A–C (black data points). In general, both monkeys made more contralateral choices with increasing signed motion strength (positive for motion toward the contralateral target) and had lower RTs (i.e., faster responses) for higher absolute motion strength.

STN neurons showed diverse response profiles. Figure 2A shows average activity patterns of three example neurons. The top neuron showed an initial suppression of activity after motion onset and became active, in a choice-dependent manner, before saccade onset. The middle neuron showed choice- and motion coherence-dependent activation during the motion-viewing period before saccade onset. The bottom neuron exhibited activation after motion onset that was similar for both choices and all coherence levels, which then decayed in a choice- and coherence-dependent manner around saccade onset.
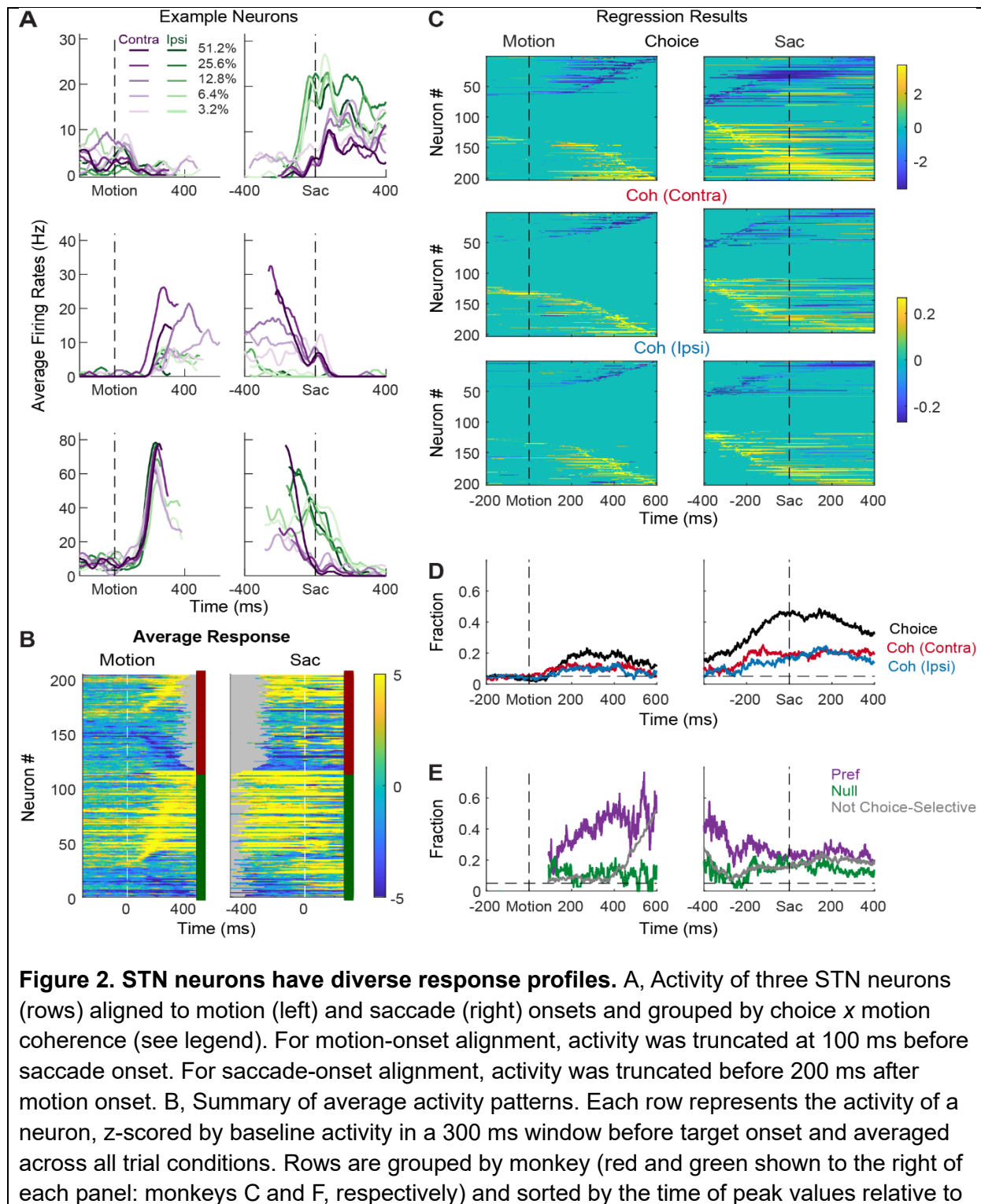
The diversity of response profiles can be seen in the summary heatmaps for the population (Figure 2B). When activity was averaged across all trial types, STN neurons can become activated or suppressed (warm vs. cool colors, respectively), relative to pre-stimulus baseline, during motion viewing and around saccade onset. The timing of peak modulation also spanned the entire motion-viewing period and extended beyond saccade generation, including a substantial fraction of neurons that also responded to target onset before the motion stimulus appeared. These diverse spatiotemporal response profiles suggest that the STN as a whole may serve multiple functions in perceptual decision-making.

Across the population, a substantial fraction of neurons was sensitive to choice, motion coherence, and RT (Figure 2C-E, Supplemental Figure 1). We performed multiple linear regressions, separately for coherence and RT (Eqs. 1 and 2), for each neuron and used the regression coefficients to measure these decision-related sensitivities. For choice sensitivity (Figure 2C, first row), both contralateral and ipsilateral preferences were commonly observed.

The overall fraction of neurons showing choice sensitivity increased after motion onset and peaked at saccade onset (Figure 2D). For coherence sensitivity, modulations were observed for trials with contralateral or ipsilateral choices and with similar tendencies for positive and negative coefficients (Figure 2C, rows 2 and 3). The fraction of neurons showing reliable coherence sensitivity was also higher around saccade onset (Figure 2D).

Despite the diverse distributions of regression coefficients, there were systematic patterns in when and how these forms of selectivity were evident in the neural responses. Notably, neurons showing choice sensitivity were more likely to show coherence modulation during early motion viewing, especially for trials when the monkey chose the neuron's preferred choice (Figure 2E, purple). In contrast, coherence modulation emerged later for neurons that did not show choice

4

98   sensitivity (Figure 2E, gray lines). These systematic interactions in modulation types suggest that
99   the STN population does not simply reflect a random mix of selectivity for decision-related
100   quantities. Instead, there appears to exist subpopulations with distinct decision-related
101   modulation patterns, which we detail below.

Figure 2. STN neurons have diverse response profiles. A, Activity of three STN neurons (rows) aligned to motion (left) and saccade (right) onsets and grouped by choice *x* motion coherence (see legend). For motion-onset alignment, activity was truncated at 100 ms before saccade onset. For saccade-onset alignment, activity was truncated before 200 ms after motion onset. B, Summary of average activity patterns. Each row represents the activity of a neuron, z-scored by baseline activity in a 300 ms window before target onset and averaged across all trial conditions. Rows are grouped by monkey (red and green shown to the right of each panel: monkeys C and F, respectively) and sorted by the time of peak values relative to

5

motion onset. Only correct trials were included. C, Heatmaps of linear regression coefficients for choice (top), coherence for trials with contralateral choices (middle), and coherence for trials with ipsilateral choices (bottom), for activity aligned to motion (left) and saccade (right) onsets. Regression was performed in running windows of 300 ms. Regression coefficients that were not significantly different from zero (*t*-test, p>0.05) were set to zero (green) for display purposes. Neurons were sorted in rows by the time of peak coefficient magnitude. Only correct trials were included. D, Time courses of the fractions of regression coefficients that were significantly different from zero (*t*-test, p<0.05), for choice (black), coherence for trials with contralateral choices (red), and coherence for trials with ipsilateral choices (blue). Dashed line indicates chance level. E, Time courses of the fractions of non-zero regression coefficients for coherence. Separate fractions were calculated for trials with the preferred (purple) and null (green) choices from choice-selectivity activity and for all trials from activity that was not choice selective (gray). Only time points after motion onset with fractions > 0.05 for choice-selective activity were included. Dashed line: chance level.
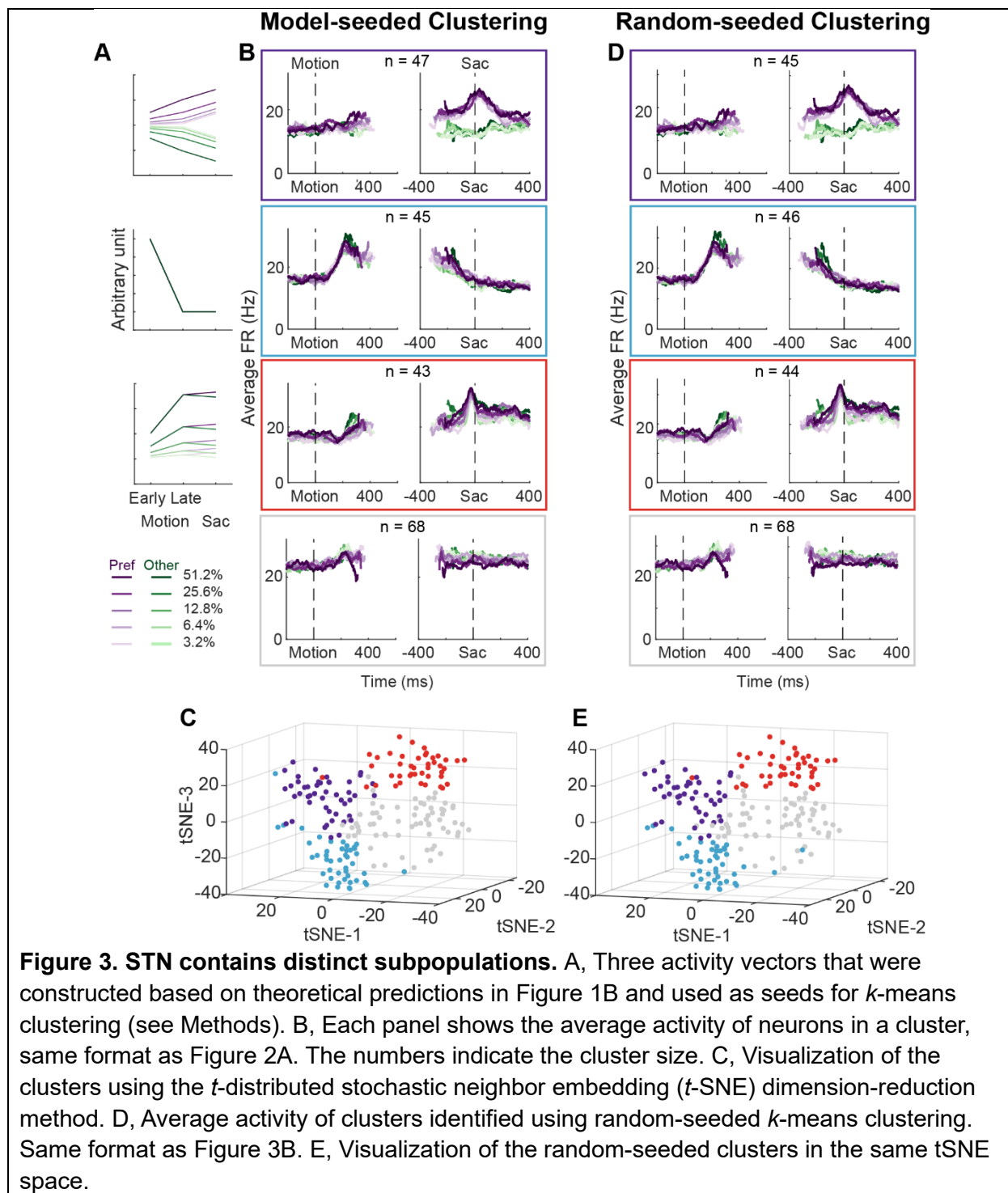
102

STN subpopulations can support previously theorized functions

104 Using two forms of cluster analysis, we identified three subpopulations of neurons in the STN
105 with distinct activity patterns that conform to predictions of each of the three previously
106 published sets of models. For the first analysis, we represented a neuron's activity pattern with a
107 30-dimension vector, consisting of normalized average activity associated with two choices, five
108 coherence levels, and three task epochs. We generated three artificial vectors based on the
109 predicted activity patterns of each model, as follows (Figure 3A). Bogacz and Gurney (2007)
110 posited that STN neurons, through their reciprocal connections with the external segment of
111 globus pallidus, pool and normalize evidence-related signals, leading to the prediction of choice
112 and coherence-modulated activity during motion viewing (Figure 3A, top). Ratcliff and Frank
113 (2012) posited that STN neurons, through their direct innervation by cortical regions, provide an
114 early signal to suppress immature choices, leading to the prediction of a choice-independent
115 signal that appears soon after motion onset and dissipates over time (Figure 3A, middle). Wei
116 and colleagues (2015) posited that the STN balances evidence-related signals in the GPe until
117 near decision time, leading to the prediction of coherence-dependent ramping activity with no or
118 weak choice selectivity (Figure 3A, bottom). We performed k-means clustering using these three
119 vectors and another arbitrary vector as the seeds to group the population into four clusters.

120 Figure 3B shows the average activity from each of the resulted clusters. Consistent with the
121 design of this analysis, the average activity of the first cluster shows choice- and coherence-
122 dependent activity that also ramps up during motion viewing. The average activity of the second
123 cluster shows an early, sharper rise in activity that is modulated by neither choice or coherence
124 during motion viewing and this activity gradually decreases toward saccade onset. The average
125 activity of the third cluster shows choice- and coherence-dependent ramping activity during
126 motion viewing and a short burst of activity for one choice just before saccade onset. The last
127 cluster shows little task-related modulation. The first three clusters contained similar numbers of
128 neurons. When visualized using the T-distributed stochastic neighbor embedding technique, these
129 clusters did not form a single continuum but instead reflected separable features between clusters

130   (Figure 3C). In other words, the clustering did not simply force a uniform distribution with
131   random-mixed selectivity into four groups.

132   For the second cluster analysis, we used random seeds without considering any of the model
133   predictions and obtained almost identical clusters. As detailed in Methods, we explored a wide
134   range of settings for clustering, including: 1) using directly the 30-D vectors or their principal

Figure 3. STN contains distinct subpopulations. A, Three activity vectors that were constructed based on theoretical predictions in Figure 1B and used as seeds for *k*-means clustering (see Methods). B, Each panel shows the average activity of neurons in a cluster, same format as Figure 2A. The numbers indicate the cluster size. C, Visualization of the clusters using the *t*-distributed stochastic neighbor embedding (*t*-SNE) dimension-reduction method. D, Average activity of clusters identified using random-seeded *k*-means clustering. Same format as Figure 3B. E, Visualization of the random-seeded clusters in the same tSNE space.

7

135 component projections, 2) basing the clustering on three different distance metrics, and 3)
136 varying the number of presumed clusters. To identify the best setting, we assessed the goodness
137 of clustering using the silhouette score, which quantifies for each member the relative distance
138 between its average within-cluster distance and distance to those in its closest neighboring
139 cluster (a higher score indicates better cluster separation). The silhouette plots favored the
140 combination of using the 30-D vector directly and correlation distance (Supplemental Figure
141 2A), which generated less variability across clusters and few/small-magnitude negative
142 silhouette scores (negative scores indicate that a member is closer to its neighboring cluster than
143 its own cluster).

144 We assessed the stability of clustering using the Rand index (Rand, 1971), which measures how
145 consistently two members are assigned to the same clusters from different iterations of clustering
146 (a high index indicates greater stability). The Rand index was generally high (above 0.95 out of a
147 max of 1) except for the combination of using the 30-D vector and cosine distance
148 (Supplemental Figure 2B). Finally, using the raw vector-correlation combination, an assumption
149 of four clusters resulted in the highest average Rand index and assuming 4-6 clusters generally
150 resulted in higher mean silhouette score and lower number of negative scores (Supplemental
151 Figure 2C). We thus considered that the raw-vector-correlation combination and an assumption
152 of four clusters produced the most stable and plausible results.

153 As shown in Figure 3D and E, the four clusters thus identified closely matched those obtained
154 using model predicted seeds, in terms of the average activity, the cluster sizes, and their locations
155 in the tSNE space. Increasing the assumed number of clusters caused changes mostly in the gray
156 cluster, with some changes in the blue cluster, and little effects on the red and purple clusters
157 (Supplemental Figure 3). Together these results suggest that, absent the ground truth on the
158 number of subpopulations in STN, there exist at least three subpopulations that each corresponds
159 to the predictions of one of three previous published models. As a consequence, STN appears in
160 principle to be able to support multiple decision-related functions.

161

Perturbation of STN activity affects choice and RT

163 To better understand STN's functional roles in the decision process, we perturbed STN activity
164 using electrical microstimulation while monkeys performed the task. Specifically, we applied a
165 train of current pulses at identified STN sites during decision formation, lasting from motion
166 onset to saccade onset. Figure 4 shows microstimulation effects on choices and RTs in three
167 example sessions. In the first example session, STN microstimulation caused a leftward
168 horizontal shift (more contralateral choices) and slope reduction (more variable choices) in the
169 psychometric curve (Figure 4A, top), as well as a substantial flattening of RT curves (faster
170 responses that depended less on motion coherence) for both choices (bottom). In the second
171 example session, STN microstimulation induced a minor leftward shift in the psychometric curve
172 and asymmetric changes in RT for the two choices (Figure 4B). In the third example session,
173 STN microstimulation did not change the psychometric curve but caused reductions in RT for
174 both choices (Figure 4C).

Figure 4. STN microstimulation affects monkeys' choice and RT. A-C, Monkey's choice (top) and RT (bottom) performance for trials with (red) and without (black) microstimulation for three example sessions (A,B: monkey C; C: monkey F). Lines: DDM fits. D, Distributions of microstimulation effects on bias and slope terms of the logistic function. Filled bars in histograms indicate sessions with significant modulation of the specific term (bootstrap method). Triangles indicate the median values. Filled triangle: Wilcoxon sign-rank test for H_0: median=0, $p < 0.05$. E and F, Summary of microstimulation effects on the offset (E) and slope (F) terms of a linear regression fit to RT data. Two separate linear regressions were performed for the two choices (Ipsi/Contra, as indicated). Triangles indicate the median values. Filled triangles: Wilcoxon sign-rank test, $p < 0.05$.

175   Across 54 different STN sites, microstimulation caused variable choice biases and tended to
176   reduce the dependence of choice on motion strength. We fitted the choice data to a logistic

9

177 function and measured choice bias (horizontal shift) and motion strength-dependence (slope). In
178 23 sessions, microstimulation induced a reliable choice bias (Figure 4D). The induced bias was
179 toward the contralateral or ipsilateral choice in 15 and 8 sessions, respectively, and the median
180 value for bias was not significantly different from zero (Wilcoxon sign-rank test, $p = 0.15$). In 18
181 sessions, microstimulation induced a change in the slope. The slope was reduced in 15 sessions
182 and the median value was negative ($p = 0.008$). These tendencies were robust across different
183 variants of logistic functions (Supplemental Figure 4). Of the sessions where inclusion of lapse
184 terms for the control and microstimulation trials produced lower AICs, very few showed
185 significant microstimulation-induced changes in lapses (2 sessions each for the "Symmetric
186 Lapse" and "Asymmetric Lapse" variants). Thus, based on fitting results using logistic functions,
187 STN microstimulation most consistently reduced the choice dependence on motion strength,
188 caused session-specific choice biases, and had minimal effects on lapses.

189 Microstimulation also tended to reduce RT. We fitted linear functions separately for RTs
190 associated with the two choices, in which offset and slope terms measure coherence-independent
191 and -dependent changes in RTs, respectively. Microstimulation caused changes in RT offsets in
192 25 sessions for contralateral choices (18 were reductions in RT, with a median change across all
193 sessions of -36 ms; Wilcoxon sign-rank test for H_0: median change=0, $p <0.0001$) and 18
194 sessions for ipsilateral choices (15 reductions, mean change = -23 ms, p<0.0001; Figure 4E).
195 Microstimulation caused changes in RT slopes in 6 sessions for contralateral choices (5 were
196 positive, implying a weaker coherence dependence; $p < 0.0001$) and 4 sessions for ipsilateral
197 choices (all 4 were positive; Figure 4F). Thus, based on fitting results using linear functions,
198 STN microstimulation can induce choice-specific changes in RT, with overall tendencies to
199 reduce both the coherence-independent component and the RT's dependence on coherence for
200 the contralateral choice.

201

Microstimulation effects reflected changes in multiple computational components

203 To infer STN's computational roles in the decision process, we examined the microstimulation
204 effects using a drift-diffusion model (DDM) framework. This framework has been widely used in
205 studies of perceptual decision-making and can provide a unified, computational account of both
206 choice and RT (Gold and Shadlen, 2007). It assumes that noisy evidence is accumulated over
207 time and a decision is made when the accumulated evidence reaches a certain decision bound.
208 The overall RT is the sum of the time needed to reach the bound and non-decision times
209 reflecting perceptual and motor latencies. Previous theoretical models also made predictions
210 about the effects of perturbing STN activity that can be interpreted in the DDM framework. The
211 model by Bogacz and Gurney (2007) predicted that the perturbation would reduce the effect of
212 task difficulty on decision performance by eliminating a nonlinear transformation that is needed
213 for appropriate evidence accumulation (Green et al., 2013). The model by Ratcliff and Frank
214 (2012) predicted that the perturbation, by causing changes in the STN's influence onto the
215 substantia nigra pars reticulata (SNr), would change temporal dynamics of the decision bound
216 and influence non-decision time. The model by Wei and colleagues (2015) predicted that the
217 perturbation would result in a reduction in the decision bound.

Figure 5. STN microstimulation affected multiple computational components in the DDM. A, Illustration of the DDM. Red/black lines represent across-trial mean/single-trial example of the evidence (top) and drift rate (bottom). Blue lines represent the collapsing decision bounds. B, Distribution of the difference in AIC between the None and Full models. Red dashed line indicates the criterion for choosing the full model: AIC difference = 3. C, Histograms of microstimulation effects on DDM parameters. Each histogram included only sessions in which the Full model outperformed the corresponding reduced model (e.g., the histogram for parameter a included only sessions in which $AIC_{NoA} - AIC_{Full} > 3$ and $AIC_{None} - AIC_{Full} > 3$). Triangles indicate median values. Filled triangles: Wilcoxon sign rank test, $p < 0.05$. D, Summary of microstimulation effects on all parameters, for sessions in which at least one significant effect was present. Sessions were sorted by the prevalence and sign of the effects.

To test whether these predictions, and/or other effects, were present in our microstimulation data, we fitted a DDM to choice and RT data simultaneously (Figure 5A). We performed AIC-based model selection and found that, in 40 of 54 sessions, the Full model, which included microstimulation effects on any model parameters, outperformed the None model, which assumed that there was no microstimulation effect on any parameters (Figure 5B). This result implies that, in these sessions, STN microstimulation affected one or more computational

11

224   components of the decision process. To better characterize these effects, we compared AICs

225   between the Full model and six reduced models to identify sessions with reliable

226   microstimulation-induced changes in particular model parameters (Supplemental Figure 5A).

227   We found that STN microstimulation resulted in reliable changes in several model parameters

228   over different subsets of sessions (Figure 5C and D). Consistent with model predictions from

229   Bogacz and Gurney (2007), microstimulation reduced the scale factor for evidence

230   accumulation, $k,$ in 14 sessions. This effect contributed to a decreased motion coherence

231   dependence of choice and RT (Figure 5C, first histogram; Wilcoxon sign-rank test for H_0: zero

232   median effect, $p = 0.021$). Consistent with model predictions from Ratcliff and Frank (2012) and

233   Wei and colleagues (2015), microstimulation affected parameters that controlled the decision

234   bound ($a, B\_collapse, B\_t$) in 16 sessions each (not necessarily in the same sessions for each

235   parameter, see Figure 5D). The changes in the maximal decision bound (a) were variable across

236   sessions ($p = 0.68$). The changes in the collapsing bound dynamics ($B\_collapse, B\_t$) tended to

237   indicate faster and earlier decreases in bounds ($p = 0.039$ and 0.088, respectively). Consistent

238   with model predictions from Ratcliff and Frank (2012), microstimulation caused changes in non-

239   decision times in 30 sessions ($t0\_Contra$ and $t0\_Ipsi$). These changes varied from session to

240   session ($p = 0.28$ and 0.75, respectively). Statistical tests on fitted parameters of all sessions,

241   regardless of whether a microstimulation effect was necessary to account for the behavioral data,

242   showed similar trends (Supplemental Figure 5B).

243   STN microstimulation had two additional effects beyond those predicted by the previous

244   modeling studies. First, consistent with the above-demonstrated microstimulation-induced choice

245   biases (Figure 4), microstimulation induced offsets in momentary (*me; n* = 16 sessions; *p* = 0.61)

246   and accumulated (*z; n* = 12 sessions; *p* = 0.016) evidence. Second, the microstimulation effects

247   involved changes in more than one model parameter in the majority of sessions (Figure 5D). We

248   did not observe any dominant combinations of effects. These results suggest that the STN is

249   causally involved in multiple decision-related functions, including those mediating the

250   dependence on evidence, choice biases, and bound dynamics.

251

Distribution of microstimulation effects reflected intermingled neuron activity patterns

253   The multi-faceted microstimulation effects, combined with the fact that the kind of

254   microstimulation we used tends to activate not just one neuron, but rather groups of neurons near

255   the tip of the electrode (Tehovnik, 1996), suggested that STN neurons with different functional

256   roles are located close to one another. Consistent with this idea, neurons that were classified as

257   belonging to different clusters tended to be intermingled (Figure 6). We did not observe any

258   consistent topographical organization patterns within or between the two monkeys. At certain

259   locations, neurons belonging to different clusters were recorded using the same electrode. We

260   calculated silhouette scores to quantify whether different the activity pattern-based neuron

261   clusters also formed clusters in the 3D physical space. The mean values were -0.09 and -0.11 for

262   the two monkeys, respectively, indicating that neurons were often closer to others from a

263   different cluster than those within the same cluster. In other words, STN subpopulations did not

264 segregate from each other and instead tended to be intermingled, and thus microstimulation
265 likely activated multiple neurons with different functional properties.

Figure 6. Different STN subpopulations are intermingled. Locations of STN neurons, color-coded by clusters based on random-seed clustering (same as Figure 3D). The Medial-Lateral (ML) values were jittered for better visualization of neurons recorded along the same track and at similar depths. Anterior-Posterior (AP) levels were relative to the anterior commissure. ML and depth levels were relative to the center of the recording chambers.

266 Although the intermingled organization of STN subpopulations, defined based on their task-
267 related activity patterns (Figure 3), made it challenging to relate a specific microstimulation
268 effect to a specific subpopulation, we did observe certain trends that could contribute to the site-
269 specific microstimulation effects. We assigned the single or multi-unit activity at the stimulation
270 sites according to the clusters identified using the random-seeded clustering (Figure 7A). We
271 then grouped the sites by neuron clusters (Figure 7C). When neurons of different clusters were
272 recorded at the same site, the same microstimulation effects were assigned to each cluster. We
273 found that the second cluster was associated with lower overall likelihood of observing
274 microstimulation effects compared to other clusters (Figure 7B; Chi-square test, H_0: the
275 likelihood is the same for the first cluster and the other clusters; $p = 0.003$), while the third
276 cluster had higher overall likelihood ($p = 0.035$). For the first three clusters, no microstimulation
277 effect dominated ($p > 0.3$ for all), whereas it was more likely to observe effects on non-decision
278 times for the fourth cluster (i.e., with neural activity patterns not related to the three models; $p =$
279 0.001).

280 The sign of microstimulation effects depended weakly on neuron clusters. For example, it was
281 more likely to observe an increase in maximal bound height ("a") for the third neuron cluster
282 (Chi-square test, H_0: same fractions of increase/decrease for all clusters; $p = 0.073$; Chi-square
283 test, H_0: equal fractions of increase/decrease within the cluster; $p = 0.036$). Microstimulation
284 decreased the scale factor ("k") for the third and fourth clusters but caused variable changes for
285 the first cluster ($p = 0.070$ and 0.021, respectively). Microstimulation effects on the non-decision
286 time for the contralateral choices were dominated by increases for the fourth cluster ($p = 0.04$
287 and 0.007, respectively). Together, these results suggested that microstimulation effects reflected

13

288   multiple contributions of intermingled STN subpopulations to decision- and non-decision-related
289   processes.
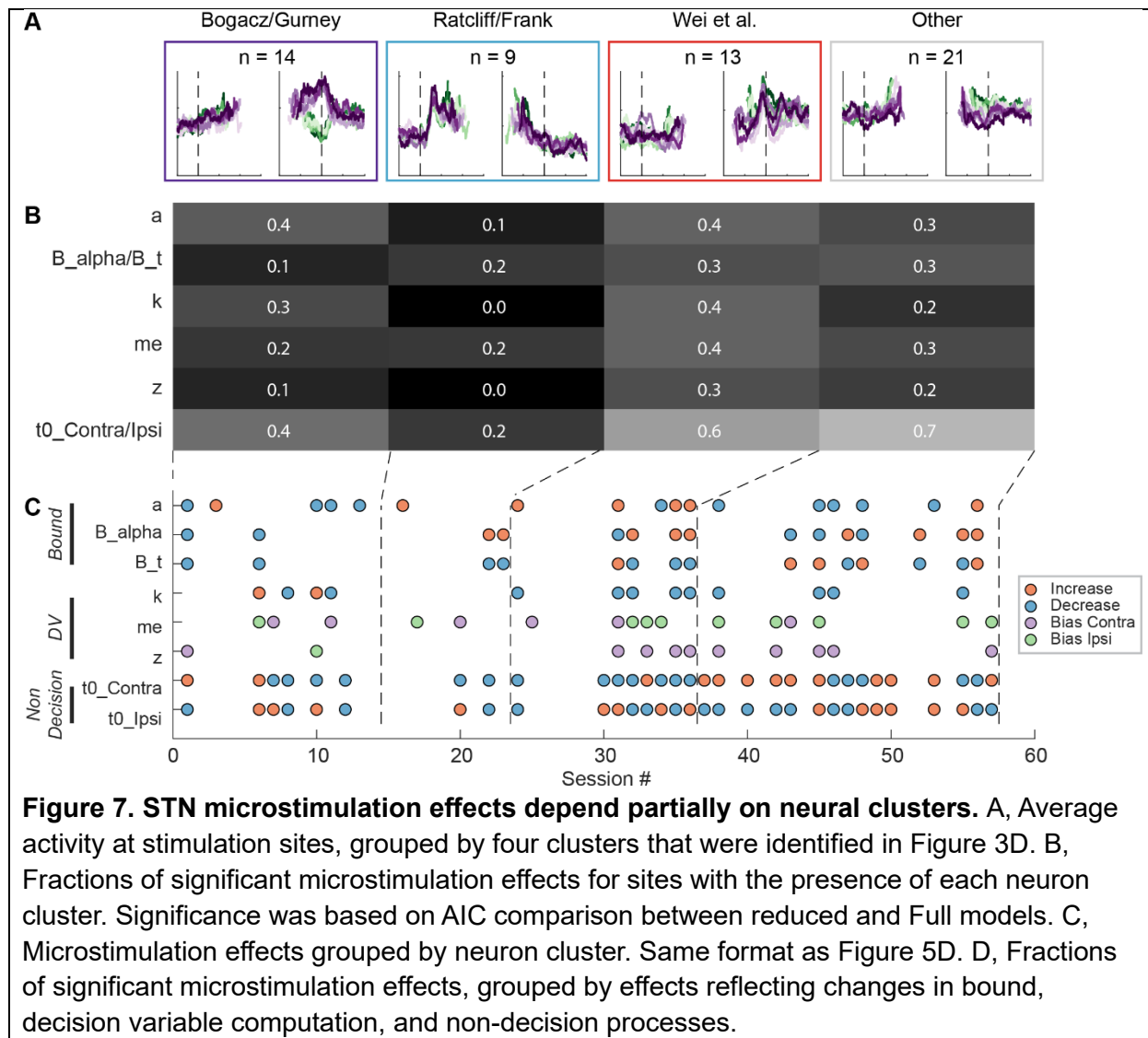
Figure 7. STN microstimulation effects depend partially on neural clusters. A, Average activity at stimulation sites, grouped by four clusters that were identified in Figure 3D. B, Fractions of significant microstimulation effects for sites with the presence of each neuron cluster. Significance was based on AIC comparison between reduced and Full models. C, Microstimulation effects grouped by neuron cluster. Same format as Figure 5D. D, Fractions of significant microstimulation effects, grouped by effects reflecting changes in bound, decision variable computation, and non-decision processes.

290

Discussion

292   We provide the first characterization of single-unit recordings and electrical microstimulation in
293   the STN of monkeys performing a demanding perceptual-decision task. We show that: 1) STN
294   neurons are heterogeneous in their response profiles; 2) different STN subpopulations, with
295   distinct decision-related activity modulation patterns and intermingled within the region, can
296   support previously-theorized functions; and 3) electrical microstimulation in STN causes
297   changes in choice and RT behaviors, reflecting effects on multiple computational components of
298   an accumulate-to-bound decision process. These results indicate that the STN plays important
299   and complex roles in perceptual decision formation, both supporting and extending existing
300   views of STN function.

14

301 Our study was motivated by the differing predictions of STN activity patterns from several
302 theoretical studies that were based on STN cellular physiology, connectivity, and/or response
303 patterns in non-perceptual decision-making contexts (Bogacz and Gurney, 2007; Ratcliff and
304 Frank, 2012; Wei et al., 2015). Remarkably, we found three clusters of STN activity that are
305 consistent with each of these predictions. The three clusters were robust and stable, emerging
306 when we used two different clustering methods (one with model-based seeds, the other with
307 random seeds). Interestingly, Zavala and colleagues (2017) have reported two types of STN
308 responses in human patients performing a flanker task. The "early" response they identified may
309 correspond to our second cluster, while the "late" response may reflect a combination of our first
310 and third clusters. Together these results suggest that the primate STN contains distinct
311 subpopulations with different functional roles. Combined with the microstimulation results, the
312 presence of these subpopulations suggests that the STN can both contribute to the conversion of
313 sensory evidence into an appropriately formatted/calibrated decision variable and modulate the
314 dynamics of decision bound. Future studies of BG function should strive to better understand
315 how these subpopulations interact with each other, as well as with other neurons in the BG and
316 larger decision network to support decision making.

317 Despite the general agreements between our observations and previous theoretical predictions,
318 there were also differences that could be informative for developing future BG models. Most
319 notably, the models focused on the period of evidence accumulation and less on neural activity
320 patterns at or after decision commitment. In contrast, our data show interesting modulations
321 around saccade onset (see Figure 3) that raise several intriguing possibilities for STN's
322 contributions to the decision process. In particular, one subpopulation showed a broad peak with
323 strong choice modulation and little coherence modulation. These modulations may reflect bound-
324 crossing in an accumulate-to-bound process (but see below). A second subpopulation returned to
325 the baseline level independent of choice or coherence. The relatively constant trajectory of this
326 modulation may reflect a collapsing bound or urgency signal that is dependent only on elapsed
327 time and not the sensory evidence. A third subpopulation maintained coherence-dependent
328 activity until very close to saccade onset and showed a sharp peak with little choice or coherence
329 modulation. This sharp peak may signal the end of decision deliberation, without specifying
330 which decision is made, to direct the network to a post-decision state for decision evaluation.

331 The diverse activity patterns and their intermingled distribution in the STN underscore the
332 challenge of identifying specific, causal contributions of a particular neural subpopulation. In
333 many sessions, we observed effects that have been predicted theoretically and observed
334 experimentally in human PD patients undergoing DBS. These effects included a reduction in RT,
335 a weaker dependence on evidence, and changes in the maximal value and trajectories of the
336 decision bound (Frank et al., 2007; Cavanagh et al., 2011; Coulthard et al., 2012; Green et al.,
337 2013; Zavala et al., 2014; Herz et al., 2016; Pote et al., 2016). In addition to these predicted
338 effects, microstimulation also changed choice biases, measured as horizontal shifts of
339 psychometric functions and as two different types of biases in the DDM framework. This
340 departure from previous DBS studies may arise from different task designs (button press versus
341 eye movement), health status of the subjects, and experience level (minimally versus extensively
342 trained). The lateralized bias suggests that the STN may be involved in flexible decision

15

343   processes that adapt to environments with asymmetric prior probability and/or reward outcomes
344   for different alternatives, in addition to modulating speed-accuracy tradeoff. Consistent with this
345   idea, DBS can affect the threshold for deliberations over uncertain sensory inputs or motivational
346   factors such as reward and effort (Pagnier et al., 2024), suggesting that the STN may be part of a
347   general selection machinery that can incorporate sensory evidence with information about the
348   task environment (Redgrave et al., 1999).

349   Our findings also suggest that STN's role in decision formation differs in important ways from
350   other oculomotor regions that have been examined under similar conditions. First, in the frontal
351   eye field (FEF), lateral intraparietal area (LIP), and superior colliculus (SC), decision-related
352   neural activity is dominated by a choice- and coherence-dependent "ramp-to-bound" pattern
353   (Roitman and Shadlen, 2002; Ding and Gold, 2012a; Crapse et al., 2018; Cho et al., 2021; Jun et
354   al., 2021; Stine et al., 2023), with additional multiplexing of decision-irrelevant signals (Meister
355   et al., 2013). In contrast, different STN subpopulations can carry distinct signals that may all be
356   relevant to decision formation. Moreover, these signals include patterns not evident in the other
357   regions, such as a choice- and coherence-independent activation in early motion viewing (blue
358   cluster in Figure 3B and D), that may signal a unique role for the STN.

359   Second, choice-selective ramping activity has been identified in LIP, FEF, SC, the caudate
360   nucleus, and now two STN subpopulations (Ding and Gold, 2010; Fan et al., 2020). However,
361   such activity differs among these oculomotor regions just before saccade onset for the preferred
362   choice. In LIP, FEF, and SC, when the ramping activity is aligned to saccade onset, it shows
363   negative coherence modulation and positive RT modulation before converging to a common,
364   higher level, consistent with an accumulate-to-bound process. In the caudate nucleus, the
365   ramping activity does not converge to a common, higher level. For the first STN subpopulation
366   (Figure 3), the ramping activity showed on average positive coherence modulation and negative
367   RT modulation (opposite to predictions of an accumulate-to-bound process) before converging to
368   a common, higher level. The second STN subpopulation did not show choice-selective activity
369   before saccade onset. These differences suggest that the caudate and STN neurons participate in
370   decision deliberation but do not directly mediate decision termination (bound crossing).

371   Third, whereas unilateral perturbations in LIP and SC tend to induce contralateral choice biases
372   (Hanks et al., 2006; Jun et al., 2021; Jeurissen et al., 2022; Stine et al., 2023), unilateral STN
373   (and caudate) microstimulation can induce both contralateral and ipsilateral choice biases,
374   depending on the stimulation site (Ding and Gold, 2012b; Doi et al., 2020). At many sites, STN
375   microstimulation effects on RT were often bilateral and of the same polarity. Moreover, STN
376   microstimulation seems to have a particularly strong effect on the overall dependence of choice
377   and RT on evidence, which was not the case for other oculomotor regions. These differences
378   suggest that the STN has unique roles in choice-independent computations, likely including
379   those involving evidence pooled for all alternatives or general bound dynamics (Bogacz and
380   Gurney, 2007; Ratcliff and Frank, 2012).

381   In summary, we characterized single-neuron activity and the effects of local perturbations in the
382   STN of monkeys performing a deliberative visual-oculomotor decision task. Our results
383   validated key aspects of previous theoretical predictions, including specific roles for the STN in

modulating decision deliberation and commitment. Our results also identified other features of decision-related processing in STN that differ from both theoretical predictions and known properties of other brain areas that contribute to these kinds of decisions. These differences can help guide future investigations that aim to delineate how cortical-subcortical interactions in general, and interactions involving the STN in particular, support decision-making and other aspects of higher brain function.

Methods

For this study, we used two adult male rhesus monkeys (*Macaca mulatta*) that have been extensively trained on the direction-discrimination (dots) task. All training, surgery, and experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of Pennsylvania Institutional Animal Care and Use Committee (protocol # 804726).

Task design and electrophysiology

The behavioral task (Figure 1A), general surgical procedure, and data acquisition methods have been described in detail previously (Ding and Gold, 2010, 2012b). Briefly, the monkey was required to report the perceived motion direction of the random-dot stimulus with a saccade at a self-determined time. Trials with different motion coherences (drawn from five levels) and directions were interleaved randomly. The monkey's eye position was monitored with a video-based eye tracker and provided reward/error feedback online based on comparisons between the monkey's eye position and task-relevant locations. Saccade reaction time (RT) was measured offline with established velocity and acceleration criteria. Neural activity was recorded using glass-coated tungsten electrodes (Alpha-Omega) or polyamide-coated tungsten electrodes (FHC, Inc.), using a grid system through a recording chamber with access to the STN. For microstimulation sessions, lower-impedance FHC electrodes were used to record and stimulate at the same sites. Single units were identified by offline spike sorting (Offline Sorter, Plexon, Inc.). Electrical microstimulation was delivered using Grass S88 stimulator as a train of negative-leading bipolar current pulses (250 µs pulse duration, 200 Hz) from motion onset to saccade onset. For most sessions, a current intensity of 50 µA was used. In other sessions, we lowered the intensity to ensure that microstimulation did not abolish the monkey's ability to complete the trials. We randomly interleaved trials with and without microstimulation at a 1:1 ratio.

Localizing the STN

We obtained structural MRI scans using T1- MPRAGE and/or T2-SPACE sequences. We estimated the likely chamber coordinates with access to the STN from these images (and 3D reconstruction using BrainSight from Rogue Research, Inc) and mapped the surrounding areas electrophysiologically. Specifically, we identified several putative landmark regions, including 1) thalamus, which showed characteristic bursts of activity in a low-firing background while the monkey dozed off; 2) reticular nucleus of the thalamus, where neurons exhibited high baseline firing rates (with bursts sometimes > 100Hz); 3) zona incerta, where neurons exhibited low, tonic

17

423 baseline firing and briefly paused their activity around saccades (Ma, 1996); 4) substantia nigra,
424 pars reticulata, where some neurons showed high baseline firing rates and suppression in activity
425 around visual stimulus or saccade onset (Hikosaka and Wurtz, 1983); and 5) substantia nigra,
426 pars compacta, where neurons showed low baseline firing and responded to unexpected reward.
427 Based on a macaque brain atlas (Saleem and Logothetis, 2007) and previously reported STN
428 activity patterns (Matsumura et al., 1992; Wichmann et al., 1994b; Isoda and Hikosaka, 2008),
429 we defined STN as the area that: 1) was surrounded by these landmark regions, 2) was separated
430 from them by gaps with minimal activity (white matter), and 3) exhibited irregular firing patterns
431 with occasional short bursts. The baseline firing rate, measured within 50 ms before fixation
432 point onset, had a mean±SD magnitude of 15.4±12.4 spikes/s in our sample.

433 <u>Neural-activity analysis</u>

434 We measured the firing rates for each neuron and trial condition in running windows (300 ms)
435 aligned to motion and saccade onsets. To visualize the overall activation/suppression, we
436 averaged the firing rates across trial conditions and computed the z-scores using a 300 ms
437 window before motion onset as the baseline. To visualize the overall choice preferences, we
438 averaged the firing rates for each choice, computed the difference between choices, and z-scored
439 the difference using the same baseline window. To quantitatively measure each neuron's
440 sensitivity to choice and motion coherence, we performed two multiple linear regressions for
441 each running window:

442 $Spike\ count = \beta_0 + \beta_{Choice} \times I_{Choice} + + \beta_{Coh-Contra} \times I_{Coh-Contra} + \beta_{Coh-Ipsi} \times I_{Coh-Ipsi}$
443 (Eq. 1)

444 $Spike\ count = \beta_0 + \beta_{Choice} \times I_{Choice} + + \beta_{RT-Contra} \times I_{RT-Contra} + \beta_{RT-Ipsi} \times I_{RT-Ipsi}$ (Eq.
445 2)

446 where $I_{Choice} = \{1\ \text{for contralateral choice}, -1\ \text{for ipsilateral choice}\}$,

447 $I_{Coh-Contra} = \{\text{coherence for contralateral choice},\ 0\ \text{for ipsilateral choice}\}$,
448 $I_{Coh-Ipsi} = \{0\ \text{for contralateral choice}, \text{coherence for ipsilateral choice}\}$.
449 $I_{RT-Contra} = \{\text{RT for contralateral choice},\ 0\ \text{for ipsilateral choice}\}$,
450 $I_{RT-Ipsi} = \{0\ \text{for contralateral choice}, \text{RT for ipsilateral choice}\}$.
451
452 Significance of non-zero coefficients was assessed using a *t*-test (criterion: $p = 0.05$).

453 <u>Cluster analysis</u>

454 We converted each neuron's activity into a 30-D vector consisting of the average firing rate
455 within three 200-ms windows for all trial conditions (i.e., 2 choices *x* 5 coherence levels). The
456 windows were selected as early motion viewing (100 – 300 ms after motion onset), late motion
457 viewing (300 – 500 ms after motion onset), and peri-saccade (100 ms before to after saccade
458 onset). The choice identity was designated as either "preferred" and "other", based on the
459 relative average activity in the peri-saccade window. Note that this designation was used so that
460 neurons with similar general modulation patterns except for the polarity of their choice

18

461 selectivity would be grouped together. This designation was not based on any statistical test and
462 did not imply that the peri-saccade activity was reliably choice selective. The average firing rate
463 for each neuron was then z-scored based on baseline rates measured in a 300 ms window ending
464 at motion onset.

465 We explored multiple method variations using k-means clustering and present results from the
466 variation with the highest stability. These variations included: 1) whether or not the vectors were
467 projected onto 11 principal components that together explained at least 95% of total variance;
468 and 2) calculation of vector distance, including squared Euclidean, cosine, and correlation
469 metrics. We determined the best settings using: 1) the Rand index (Rand, 1971), which quantifies
470 the stability of clusters in repeated clustering; 2) Silhouette scores, which quantifies the quality
471 of grouping and separation between clusters; and 3) visual inspection of clustering results in
472 terms of both cluster distribution in a t-SNE space and average activity of the clusters. To
473 compute the Rand index, we performed 50 runs of clustering, assuming 3-9 clusters, for each
474 combination of variations. The Rand index was computed as the fraction of consistent grouping
475 between a pair of units between two clustering runs. For two runs of clustering results, Rand
476 index $= \frac{N_{same-same} + N_{different-different}}{N_{all\ pairs}}$, where $N_{same-same}$ counts the number of neuron pairs that
477 share clusters in both runs, $N_{different-different}$ counts the number of neuron pairs that do not
478 share clusters in either run, and $N_{all\ pairs}$ counts the total number of neuron pairs. To compute
479 the Silhouette scores, we chose the best of 100 repetitions of clustering for each combination of
480 variations. For each neuron, Silhouette score $= \frac{\max(D_{inter-cluster},\ D_{intra-cluster})}{D_{inter-cluster} - D_{intra-cluster}}$, where
481 $D_{inter-cluster}$ is the average distance to the neuron's nearest neighboring cluster, and
482 $D_{intra-cluster}$ is the average distance to other neurons in the same cluster. A positive score
483 implies that, for the given neuron, its activity was more similar to other neurons within the same
484 cluster than those in its nearest neighboring cluster. A negative score implies that the neuron's
485 activity was more similar to those outside its own cluster.

486 To classify activity recorded at a microstimulation site, we calculated the correlation between its
487 30-D vector and the centroids from random-seeded clustering. The centroid with the highest
488 correlation value determined the cluster identity of the activity.

489 <u>Microstimulation-effects analysis</u>

490 We analyzed microstimulation effects in several ways. To characterize the effects without
491 assumptions about the underlying decision process, we fitted logistic functions to the choice data
492 and linear functions to the RT data. We used three variants of the logistic functions that differed
493 in their use of lapse rates:

494 No Lapse: $p\ (contralateral\ choice) = \frac{1}{1 + e^{-(Slope_0 + Slope_{estim}) \times (Coh + Bias_0 + Bias_{estim})}}$ (Eq. 3)

495 Symmetric Lapse: $p\ (contralateral\ choice) = \lambda_0 + \lambda_{estim} +$

496 $\frac{1 - 2 \times (\lambda_0 + \lambda_{estim})}{1 + e^{-(Slope_0 + Slope_{estim}) \times (Coh + Bias_0 + Bias_{estim})}}$ (Eq. 4)

19

497 Asymmetric Lapse: $p\,(contralateral\ choice) = \lambda_{Ipsi0} + \lambda_{Ipsi-estim} +$

498 $\frac{1-\lambda_{Ipsi0}-\lambda_{Ipsi-estim}-\lambda_{Contra0}-\lambda_{Contra-estim}}{1+e^{-(Slope_0+Slope_{estim})\times(Coh+Bias_0+Bias_{estim})}}$ (Eq. 5)

499 where Coh is the signed coherence (positive/negative for motion toward the
500 contralateral/ipsilateral choice). To assess the significance of the "estim" terms, we used
501 bootstrap methods. Specifically, we generated 200 sets of data by shuffling the microstimulation
502 status of trials within each session. We fitted these artificial data using the same logistic
503 functions to estimate null distributions for each parameter and performed a one-tailed test to
504 determine if the actual fit value exceeded chance (criterion, $p < 0.05$).

505 We fitted linear functions to the RT data, separately for the two choices:

506 $RT = Offest_0 + Offset_{estim} + (Slope_0 + Slope_{estim}) \times Coh_{unsigned}$ (Eq. 6)

507 We assessed significance using t-tests (criterion, $p < 0.05$).

508 To infer microstimulation effects on decision-related computations, we fitted drift-diffusion
509 models to choice and RT data simultaneously. We used DDM variants with collapsing bounds
510 (DDM; Figure 7A), following previously established procedures (Fan et al., 2018; Doi et al.,
511 2020). Briefly, the DDM assumes that motion evidence is accumulated over time into a decision
512 variable (DV), which is compared to two collapsing choice bounds. A choice is made when the
513 DV crosses either bound, such that the time of crossing determines the decision time and the
514 identity of the bound determines the choice identity. The model has eight basic parameters
515 (presented here in six groups): 1) a, the maximal bound height; 2) $B\_collapse$ and B_t, the decay
516 speed and onset specifying the time course of the bound "collapse"; 3) k, a scale factor governing
517 the rate of evidence accumulation; 4) me, an offset specifying a bias in the rate of evidence
518 accumulation; 5) z, an offset specifying a bias in the DV, or equivalently, asymmetric offsets of
519 equal magnitude for the two choice bounds; and 6) $t0_{contra}$ and $t0_{ipsi}$, non-decision times for
520 the two choices that capture RT components that do not depend on evidence accumulation (e.g.,
521 visual latency and motor delay).

522 We used 8 variants of DDM. In the Full model, all eight parameters were allowed to change with
523 microstimulation. In the None model, all eight parameters did not change with microstimulation.
524 In six reduced models (NoA, NoCollapse, NoK, NoME, NoZ, NoT), the corresponding group of
525 parameters (specified above) were fixed while the other parameters were allowed to change with
526 microstimulation. We fitted each model using the maximum *a posteriori* estimate method and
527 previously established prior distributions (Wiecki et al., 2013). We performed five runs for each
528 fit and used the best run (highest likelihood) for analyses here. We used the Akaike Information
529 Criterion (AIC) for model selection. We considered an AIC difference >3 to indicate that the
530 smaller-AIC model significantly outperformed the larger-AIC model. For a given sessions, if the
531 Full model outperformed a reduced model and the None model, we considered that session to
532 show significant microstimulation effect(s) on the corresponding model parameter(s). For
533 example, we considered STN microstimulation to induce significant changes in k if the Full
534 model outperformed both None and NoK models for a given session.

20

Acknowledgements

Author contributions

Conceptualization, LD and JIG; methodology, KR and LD; investigation, KR and LD; visualization, LD and JIG; funding acquisition, LD and JIG; project administration, LD; supervision, LD; writing – original draft, LD; writing – review & editing, KR, JIG, and LD.

Declaration of interests

The authors declare no competing financial interests.

Figure legends

Figure 1. Behavioral task and model predictions. A, Behavioral task. The monkey was required to report the perceived motion direction of the random-dot stimulus by making a saccade towards the corresponding choice target at a self-determined time. B, Three previous models predicted different patterns of STN activity.

Figure 2. STN neurons have diverse response profiles. A, Activity of three STN neurons (rows) aligned to motion (left) and saccade (right) onsets and grouped by choice x motion coherence (see legend). For motion-onset alignment, activity was truncated at 100 ms before saccade onset. For saccade-onset alignment, activity was truncated before 200 ms after motion onset. B, Summary of average activity patterns. Each row represents the activity of a neuron, z-scored by baseline activity in a 300 ms window before target onset and averaged across all trial conditions. Rows are grouped by monkey (red and green shown to the right of each panel: monkeys C and F, respectively) and sorted by the time of peak values relative to motion onset. Only correct trials were included. C, Heatmaps of linear regression coefficients for choice (top), coherence for trials with contralateral choices (middle), and coherence for trials with ipsilateral choices (bottom), for activity aligned to motion (left) and saccade (right) onsets. Regression was performed in running windows of 300 ms. Regression coefficients that were not significantly different from zero (t-test, p>0.05) were set to zero (green) for display purposes. Neurons were sorted in rows by the time of peak coefficient magnitude. Only correct trials were included. D, Time courses of the fractions of regression coefficients that were significantly different from zero (t-test, p<0.05), for choice (black), coherence for trials with contralateral choices (red), and coherence for trials with ipsilateral choices (blue). Dashed line indicates chance level. E, Time courses of the fractions of non-zero regression coefficients for coherence. Separate fractions were calculated for trials with the preferred (purple) and null (green) choices from choice-selectivity activity and for all trials from activity that was not choice selective (gray). Only time points after motion onset with fractions > 0.05 for choice-selective activity were included. Dashed line indicates chance level.

Figure 3. STN contains distinct subpopulations. A, Three activity vectors that were constructed based on theoretical predictions in Figure 1B and used as seeds for k-means clustering (see Methods). B, Each panel shows the average activity of neurons in a cluster, same format as Figure 2A. The numbers indicate the cluster size. C, Visualization of the clusters using the t-distributed stochastic neighbor embedding (t-SNE) dimension-reduction method. D, Average activity of clusters identified using random-seeded k-means clustering. Same format as Figure 3B. E, Visualization of the random-seeded clusters in the same tSNE space.

Figure 4. STN microstimulation affects monkeys' choice and RT. A-C, Monkey's choice (top) and RT (bottom) performance for trials with (red) and without (black) microstimulation for three example sessions (A,B: monkey C; C: monkey F). Lines: DDM fits. D, Distributions of microstimulation effects on bias and slope terms of the logistic function. Filled bars in histograms indicate sessions with significant modulation of the specific term (bootstrap method). Triangles indicate the median values. Filled triangle: Wilcoxon sign-rank test for H_0: median=0, $p < 0.05$. E and F, Summary of microstimulation effects on the offset (E) and slope (F) terms of a

590   linear regression fit to RT data. Two separate linear regressions were performed for the two
591   choices (Ipsi/Contra, as indicated). Triangles indicate the median values. Filled triangles:
592   Wilcoxon sign-rank test, $p < 0.05$.

593   **Figure 5. STN microstimulation affected multiple computational components in the DDM.**
594   A, Illustration of the DDM. Red/black lines represent across-trial mean/single-trial example of
595   the evidence (top) and drift rate (bottom). Blue lines represent the collapsing decision bounds. B,
596   Distribution of the difference in AIC between the None and Full models. Red dashed line
597   indicates the criterion for choosing the full model: AIC difference = 3. C, Histograms of
598   microstimulation effects on DDM parameters. Each histogram included only sessions in which
599   the Full model outperformed the corresponding reduced model (e.g., the histogram for parameter
600   a included only sessions in which $AIC_{NoA} - AIC_{Full} > 3$ and $AIC_{None} - AIC_{Full} > 3$). Triangles
601   indicate median values. Filled triangles: Wilcoxon sign rank test, $p < 0.05$. D, Summary of
602   microstimulation effects on all parameters, for sessions in which at least one significant effect
603   was present. Sessions were sorted by the prevalence and sign of the effects.

604   **Figure 6. Different STN subpopulations are intermingled.** Locations of STN neurons, color-
605   coded by clusters based on random-seed clustering (same as Figure 3D). The Medial-Lateral
606   (ML) values were jittered for better visualization of neurons recorded along the same track and at
607   similar depths. Anterior-Posterior (AP) levels were relative to the anterior commissure. ML and
608   depth levels were relative to the center of the recording chambers.

609   **Figure 7. STN microstimulation effects depend partially on neural clusters.** A, Average
610   activity at stimulation sites, grouped by four clusters based on the clusters in Figure 3D. B,
611   Fractions of significant microstimulation effects for sites with the presence of each neuron
612   cluster. Significance was based on AIC comparison between reduced and Full models. C,
613   Microstimulation effects grouped by neuron cluster. Same format as Figure 5D. D, Fractions of
614   significant microstimulation effects, grouped by effects reflecting changes in bound, decision
615   variable computation, and non-decision processes.

616

Supplemental Information

Suppl. Figure 1. STN activity is modulated by choice and RT. Same format as Figure 2, except using choice and RT as regressors.

Suppl. Figure 2. Clustering parameters. A, Silhouette plots for clustering results using different combinations of settings. Silhouette scores for neurons are grouped by clusters and sorted. Red lines indicate the mean scores. Yellow shaded box indicates the chosen setting for results in Figure 3. B, Average Rand indices for different clustering settings. For each setting, the k-means algorithm was run 50 times, each time picking the best clusters out of 100 repetitions. Higher Rand index indicates greater cluster stability across different runs. C, Mean silhouette scores and the number of negative scores as a function of number of clusters, using the firing rate vectors and correlation distance. Higher mean score and fewer negative scores indicate better clustering.

Suppl. Figure 3. Clustering results using alternative numbers of clusters, visualized in tSNE space. Same format as Figure 3E.

Suppl Figure 4. Comparison of different logistic models. A, The No Lapse model was associated with the lowest AIC for most sessions. The Symmetric Lapse model was associated with lower AICs for 12 sessions. The Asymmetric Lapse model was associated with lower AICs for 8 sessions. B, Histograms of microstimulation effects on bias, slope, and lapse terms in the Symmetric Lapse model. C, Histograms of microstimulation effects on bias, slope, and two lapse (for each choice) terms in the Asymmetric Lapse model. Same format as the histograms in Figure 6A.

Suppl Figure 5. A, Differences in AIC between reduced and Full models. Filled circles indicate sessions for which $\text{AIC}_{\text{Reduced}} - \text{AIC}_{\text{Full}} > 3$ (red line). Note that for three sessions, the Full model outperformed the None model but not any of the reduced models. B, Histograms of difference in DDM parameters between trials with and without microstimulation. Filled bars represent sessions considered to show significant microstimulation effects on the given parameter, based on AIC comparisons. Triangles indicate median values. Filled triangles: Wilcoxon sign-rank test, $p < 0.05$.

References
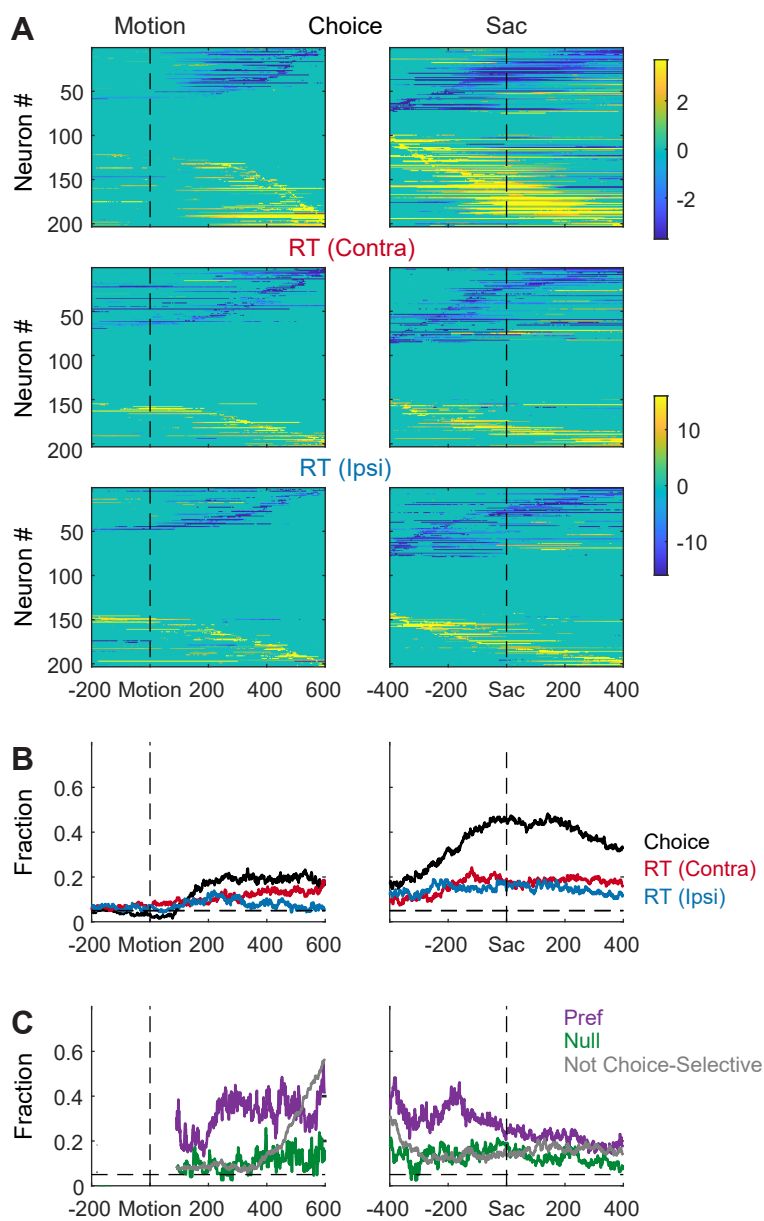
Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack RA (2007) Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. J Neurosci Off J Soc Neurosci 27:3743–3752.

Aron AR, Poldrack RA (2006) Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. J Neurosci 26:2424–2433.

Baunez C, Humby T, Eagle DM, Ryan LJ, Dunnett SB, Robbins TW (2001) Effects of STN lesions on simple vs choice reaction time tasks in the rat: preserved motor readiness, but impaired response selection. Eur J Neurosci 13:1609–1616.

Bergman H, Wichmann T, DeLong MR (1990) Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. Science 249:1436–1438.

Bergman H, Wichmann T, Karmon B, DeLong MR (1994) The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of parkinsonism. J Neurophysiol 72:507–520.

Bogacz R, Gurney K (2007) The basal ganglia and cortex implement optimal decision making between alternative actions. Neural Comput 19:442–477.

Brittain J-S, Watkins KE, Joundi RA, Ray NJ, Holland P, Green AL, Aziz TZ, Jenkinson N (2012) A role for the subthalamic nucleus in response inhibition during conflict. J Neurosci 32:13396–13401.

Carpenter MB, Whittier JR, Mettler FA (1950) Analysis of choreoid hyperkinesia in the Rhesus monkey; surgical and pharmacological analysis of hyperkinesia resulting from lesions in the subthalamic nucleus of Luys. J Comp Neurol 92:293–331.

Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ (2011) Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat Neurosci 14:1462–1467.

Cho S-H, Crapse T, Grimaldi P, Lau H, Basso MA (2021) Variable Statistical Structure of Neuronal Spike Trains in Monkey Superior Colliculus. J Neurosci 41:3234–3253.

Coulthard EJ, Bogacz R, Javed S, Mooney LK, Murphy G, Keeley S, Whone AL (2012) Distinct roles of dopamine and subthalamic nucleus in learning and probabilistic decision making. Brain.

Crapse TB, Lau H, Basso MA (2018) A Role for the Superior Colliculus in Decision Criteria. Neuron 97:181-194.e6.

DeLong MR, Wichmann T (2001) Deep brain stimulation for Parkinson's disease. Ann Neurol 49:142–143.

Desbonnet L, Temel Y, Visser-Vandewalle V, Blokland A, Hornikx V, Steinbusch HWM (2004) Premature responding following bilateral stimulation of the rat subthalamic nucleus is amplitude and frequency dependent. Brain Res 1008:198–204.

Ding L, Gold JI (2010) Caudate encodes multiple computations for perceptual decisions. J Neurosci 30:15747–15759.

Ding L, Gold JI (2012a) Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. Cereb Cortex 22:1052–1067.

Ding L, Gold JI (2012b) Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. Neuron 75:865–874.

Doi T, Fan Y, Gold JI, Ding L (2020) The caudate nucleus contributes causally to decisions that balance reward and uncertain visual information. Elife 9:e56694.

Fan Y, Gold JI, Ding L (2018) Ongoing, rational calibration of reward-driven perceptual biases. Elife 7:e36018.

Fan Y, Gold JI, Ding L (2020) Frontal eye field and caudate neurons make different contributions to reward-biased perceptual decisions. eLife 9:e60535.

Frank MJ (2006) Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw 19:1120–1136.

Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007) Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science 318:1309–1312.

Fumagalli M, Giannicola G, Rosa M, Marceglia S, Lucchiari C, Mrakic-Sposta S, Servello D, Pacchetti C, Porta M, Sassi M, Zangaglia R, Franzini A, Albanese A, Romito L, Piacentini S, Zago S, Pravettoni G, Barbieri S, Priori A (2011) Conflict-dependent dynamic of subthalamic nucleus oscillations during moral decisions. Soc Neurosci 6:243–256.

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574.

Green N, Bogacz R, Huebl J, Beyer AK, Kuhn AA, Heekeren HR (2013) Reduction of influence of task difficulty on perceptual decision making by STN deep brain stimulation. Curr Biol 23:1681–1684.

Hanks TD, Ditterich J, Shadlen MN (2006) Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. Nat Neurosci 9:682–689.

Herz DM, Tan H, Brittain JS, Fischer P, Cheeran B, Green AL, FitzGerald J, Aziz TZ, Ashkan K, Little S, Foltynie T, Limousin P, Zrinzo L, Bogacz R, Brown P (2017) Distinct mechanisms mediate speed-accuracy adjustments in cortico-subthalamic networks. Elife 6.

Herz DM, Zavala BA, Bogacz R, Brown P (2016) Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus. Curr Biol 26:916–920.

Hikosaka O, Wurtz RH (1983) Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. J Neurophysiol 49:1268–1284.

Isoda M, Hikosaka O (2008) Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. J Neurosci 28:7209–7218.

Jeurissen D, Shushruth S, El-Shamayleh Y, Horwitz GD, Shadlen MN (2022) Deficits in decision-making induced by parietal cortex inactivation are compensated at two timescales. Neuron 110:1924-1931.e5.

Jun E, Bautista A, Nunez M, Allen D, Tak J, Alvarez E, Basso M (2021) Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. Nat Neurosci 24:1121–1131.

Lehericy S, Ducros M, Krainik A, Francois C, Van de Moortele PF, Ugurbil K, Kim DS (2004) 3-D diffusion tensor axonal tracking shows distinct SMA and pre-SMA projections to the human striatum. Cereb Cortex 14:1302–1309.

Lo CC, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat Neurosci 9:956–963.

Ma TP (1996) Saccade-related omnivectoral pause neurons in the primate zona incerta. Neuroreport 7:2713–2716.

Martin JP (1927) HEMICHOREA RESULTING FROM A LOCAL LESION OF THE BRAIN. (THE SYNDROME OF THE BODY OF LUYS. Brain 50:637–649.

Martin JP, Alcock NS (1934) HEMICHOREA ASSOCIATED WITH A LESION OF THE CORPUS LUYSII. Brain 57:504–516.

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

Meister ML, Hennig JA, Huk AC (2013) Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. J Neurosci 33:2254–2267.

Pagnier GJ, Asaad WF, Frank MJ (2024) Double dissociation of dopamine and subthalamic nucleus stimulation on effortful cost/benefit decision making. Curr Biol CB 34:655-660.e3.

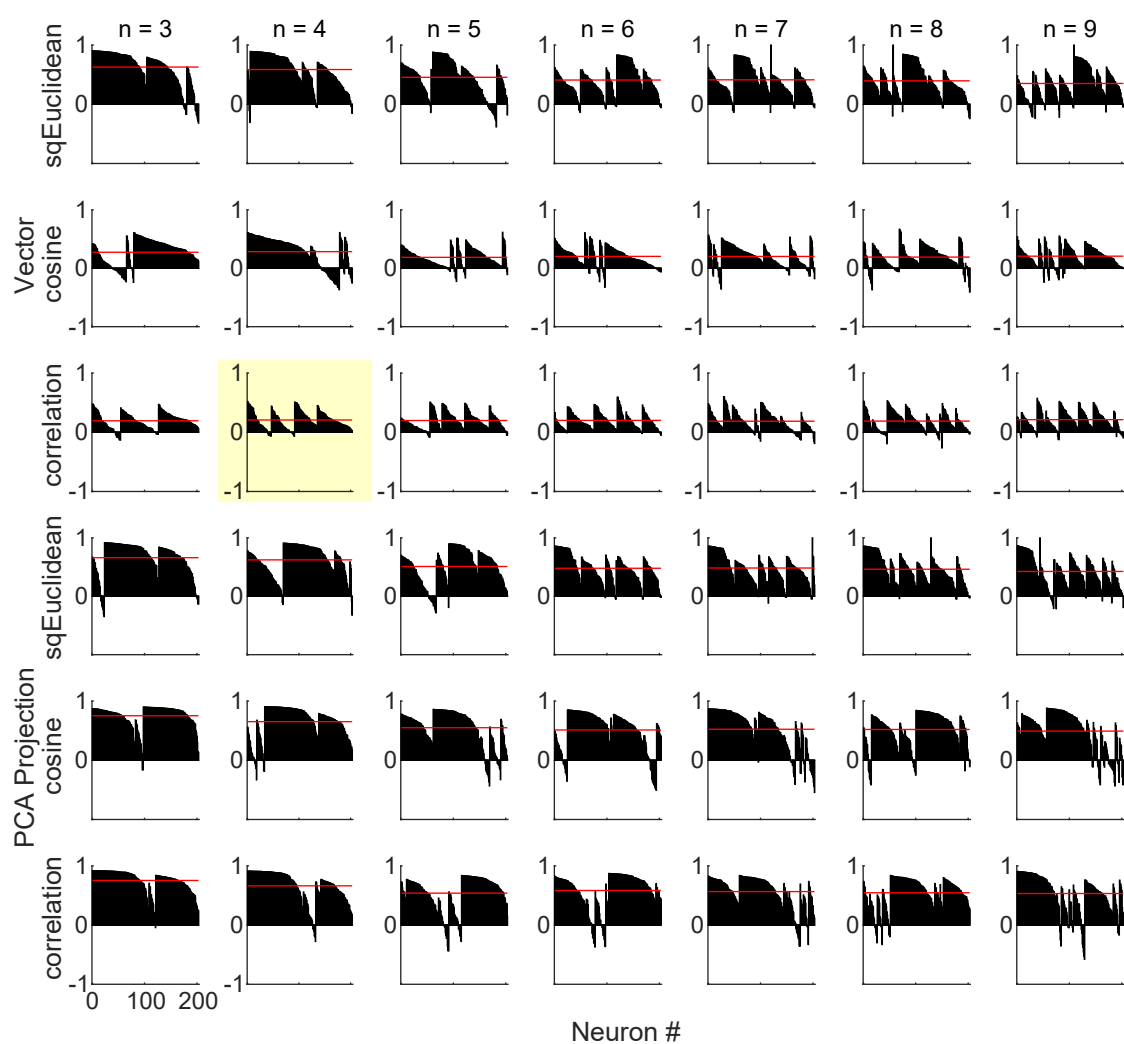Pasquereau B, Turner RS (2017) A selective role for ventromedial subthalamic nucleus in inhibitory control. Elife 6.

Pote I, Torkamani M, Kefalopoulou ZM, Zrinzo L, Limousin-Dowsey P, Foltynie T, Speekenbrink M, Jahanshahi M (2016) Subthalamic nucleus deep brain stimulation induces impulsive action when patients with Parkinson's disease act under speed pressure. Exp Brain Res 234:1837–1848.

Rand WM (1971) Objective Criteria for the Evaluation of Clustering Methods. J Am Stat Assoc 66:846–850.

Ratcliff R, Frank MJ (2012) Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. Neural Comput 24:1186–1229.

Redgrave P, Prescott TJ, Gurney K (1999) The basal ganglia: a vertebrate solution to the selection problem? Neuroscience 89:1009–1023.

Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci 22:9475–9489.

Saleem KS, Logothetis N (2007) A combined MRI and histology atlas of the rhesus monkey brain in stereotaxic coordinates. London ; Burlington, MA: Academic.

Schmidt R, Leventhal DK, Mallet N, Chen F, Berke JD (2013) Canceling actions involves a race between basal ganglia pathways. Nat Neurosci 16:1118–1124.

Stine GM, Trautmann EM, Jeurissen D, Shadlen MN (2023) A neural mechanism for terminating decisions. Neuron 111:2601-2613.e5.

Tehovnik EJ (1996) Electrical stimulation of neural tissue to evoke behavioral responses. J Neurosci Methods 65:1–17.

Wei W, Rubin JE, Wang XJ (2015) Role of the indirect pathway of the basal ganglia in perceptual decision making. J Neurosci 35:4052–4064.

Whittier JR, Mettler FA (1949) Studies on the subthalamus of the rhesus monkey; hyperkinesia and other physiologic effects of subthalamic lesions; with special reference to the subthalamic nucleus of Luys. J Comp Neurol 90:319–372.

Wichmann T, Bergman H, DeLong MR (1994a) The primate subthalamic nucleus. III. Changes in motor behavior and neuronal activity in the internal pallidum induced by subthalamic inactivation in the MPTP model of parkinsonism. J Neurophysiol 72:521–530.

Wichmann T, Bergman H, DeLong MR (1994b) The primate subthalamic nucleus. I. Functional properties in intact animals. J Neurophysiol 72:494–506.

Witt K, Pulkowski U, Herzog J, Lorenz D, Hamel W, Deuschl G, Krack P (2004) Deep brain stimulation of the subthalamic nucleus improves cognitive flexibility but impairs response inhibition in Parkinson disease. Arch Neurol 61:697–700.

781 Zaghloul KA, Weidemann CT, Lega BC, Jaggi JL, Baltuch GH, Kahana MJ (2012) Neuronal
782   Activity in the Human Subthalamic Nucleus Encodes Decision Conflict during Action
783   Selection. J Neurosci 32:2453–2460.

784 Zavala B, Damera S, Dong JW, Lungu C, Brown P, Zaghloul KA (2017) Human Subthalamic
785   Nucleus Theta and Beta Oscillations Entrain Neuronal Firing During Sensorimotor
786   Conflict. Cereb Cortex 27:496–508.

787 Zavala BA, Tan H, Little S, Ashkan K, Hariz M, Foltynie T, Zrinzo L, Zaghloul KA, Brown P
788   (2014) Midline Frontal Cortex Low-Frequency Activity Drives Subthalamic Nucleus
789   Oscillations during Conflict. J Neurosci 34:7322.

790

Suppl. Figure 1. STN activity is modulated by choice and RT. Same format as Figure 2, except using choice and RT as regressors.
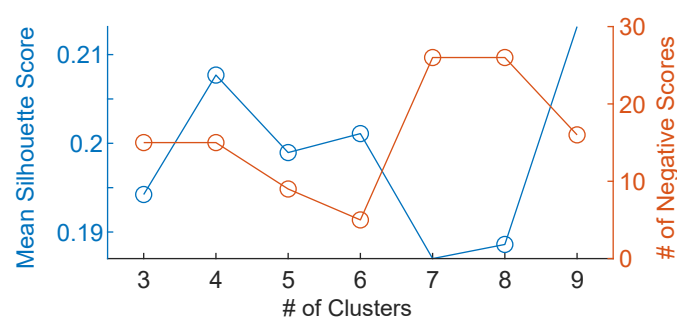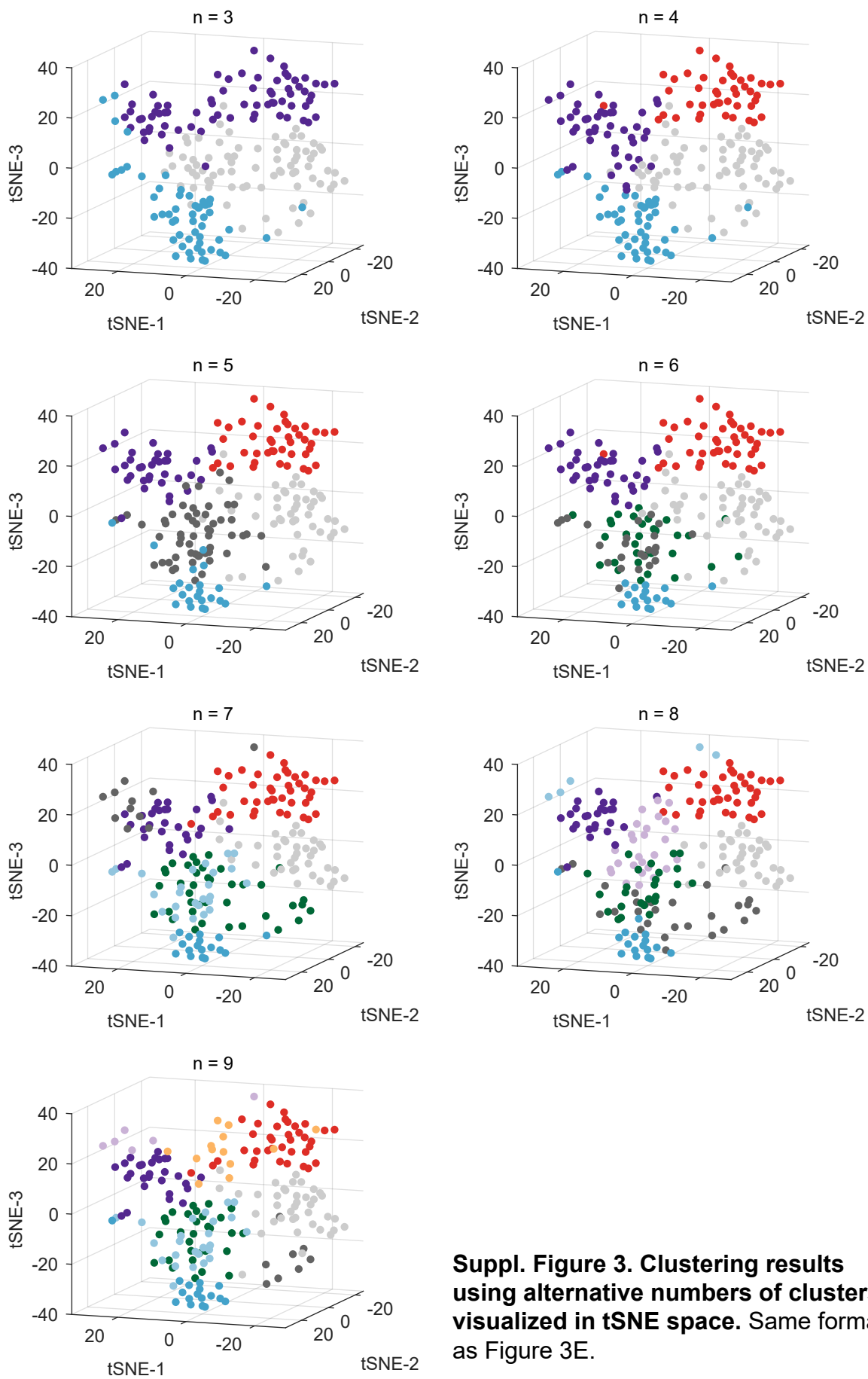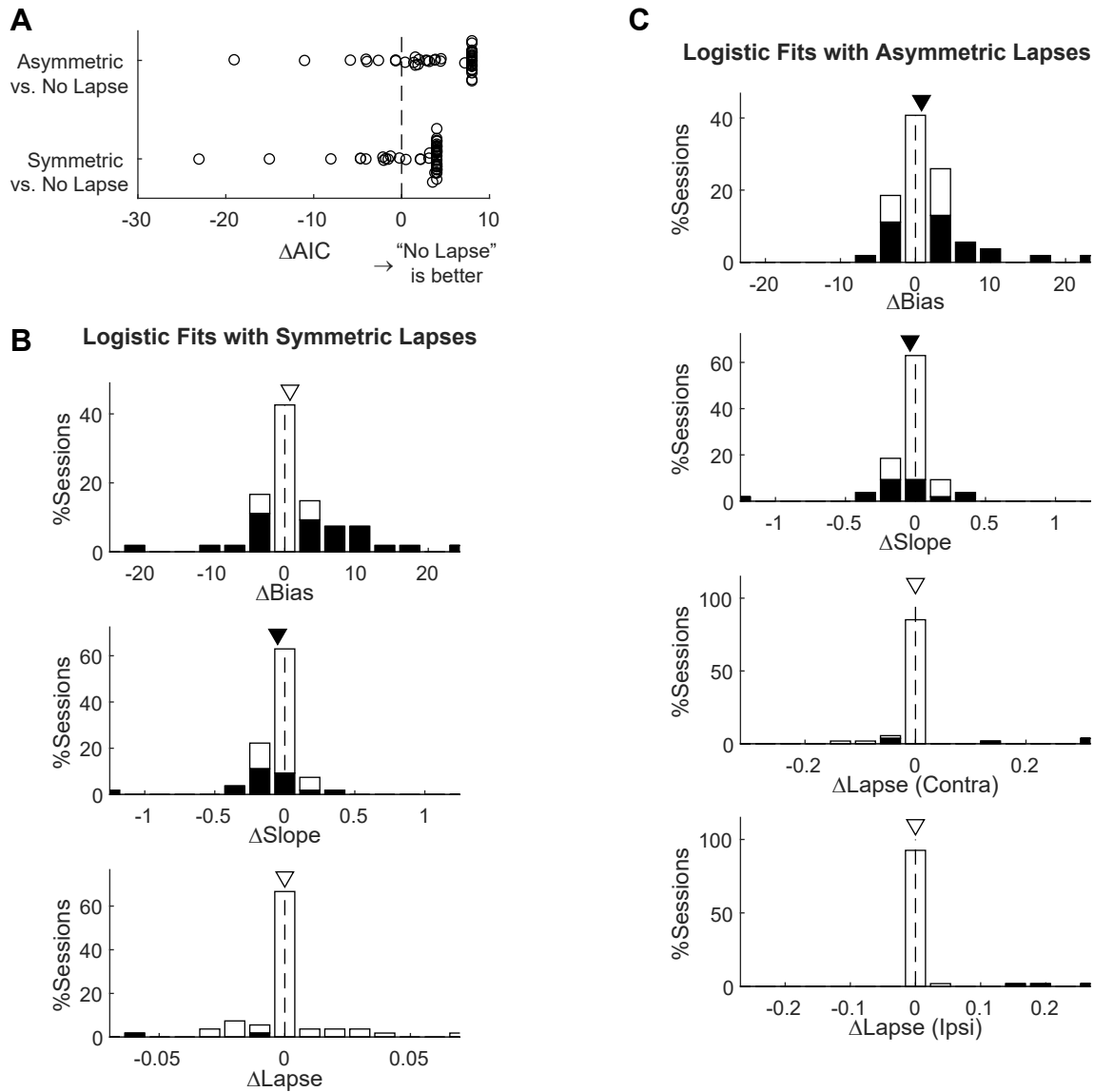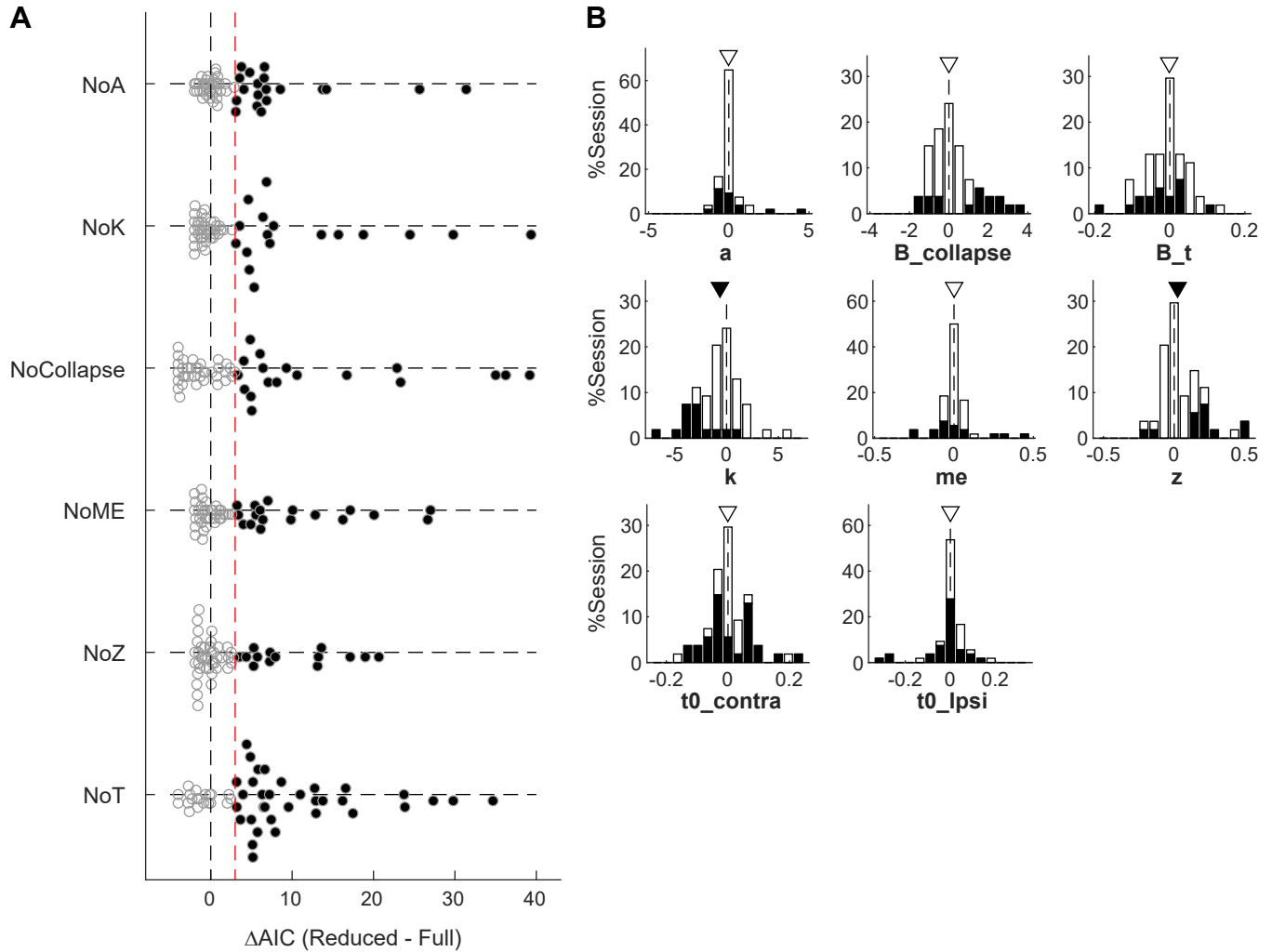
Suppl. Figure 2. Clustering parameters. A, Silhouette plots for clustering results using different combinations of settings. Silhouette scores for neurons are grouped by clusters and sorted. Red lines indicate the mean scores. Yellow shaded box indicates the chosen setting for results in Figure 3. B, Average Rand indices for different clustering settings. For each setting, the k-means algorithm was run 50 times, each time picking the best clusters out of 100 repetitions. Higher Rand index indicates greater cluster stability across different runs. C, Mean silhouette scores and the number of negative scores as a function of number of clusters, using the firing rate vectors and correlation distance. Higher mean score and fewer negative scores indicate better clustering.

Suppl. Figure 3. Clustering results using alternative numbers of clusters, visualized in tSNE space. Same format as Figure 3E.

Suppl Figure 4. Comparison of different logistic models. A, The No Lapse model was associated with the lowest AIC for most sessions. The Symmetric Lapse model was associated with lower AICs for 12 sessions. The Asymmetric Lapse model was associated with lower AICs for 8 sessions. B, Histograms of microstimulation effects on bias, slope, and lapse terms in the Symmetric Lapse model. C, Histograms of microstimulation effects on bias, slope, and two lapse (for each choice) terms in the Asymmetric Lapse model. Same format as the histograms in Figure 6A.

Suppl Figure 5. A, Differences in AIC between reduced and Full models. Filled circles indicate sessions for which $AIC_{Reduced} - AIC_{Full} > 3$ (red line). Note that for three sessions, the Full model outperformed the None model but not any of the reduced models. B, Histograms of difference in DDM parameters between trials with and without microstimulation. Filled bars represent sessions considered to show significant microstimulation effects on the given parameter, based on AIC comparisons. Triangles indicate median values. Filled triangles: Wilcoxon sign-rank test, $p < 0.05$.