Article

Neuron

The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment

Highlights

- The BG do not contribute to deciding which movement target choice is selected
- BG activity reflects an urgency signal and its adjustment between SAT policies
- The BG are involved in confirming the commitment to a cortically determined choice

Authors

David Thura, Paul Cisek

Correspondence

david.thura@umontreal.ca

In Brief

Thura and Cisek show that the basal ganglia do not influence the choice of the target for action, but rather provide a context-dependent urgency signal that invigorates the deliberation process before confirming the choice determined in the cerebral cortex.

Thura & Cisek, 2017, Neuron *95*, 1160–1170
August 30, 2017
http://dx.doi.org/10.1016/j.neuron.2017.07.039

CellPress

CellPress

Neuron
Article

The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment

David Thura[1,2,*] and Paul Cisek[1]
[1]Department of Neuroscience, Université de Montréal, Montréal, QC H3T 1J4, Canada
[2]Lead Contact
*Correspondence: david.thura@umontreal.ca
http://dx.doi.org/10.1016/j.neuron.2017.07.039

SUMMARY

Prominent theories of decision making suggest that the basal ganglia (BG) play a causal role in deliberation between action choices. An alternative hypothesis is that deliberation occurs in cortical regions, while the BG control the speed-accuracy trade-off (SAT) between committing to a choice versus continuing to deliberate. Here, we test these hypotheses by recording activity in the internal and external segments of the globus pallidus (GPi/GPe) while monkeys perform a task dissociating the process of deliberation, the moment of commitment, and adjustment of the SAT. Our data suggest that unlike premotor and motor cortical regions, pallidal output does not contribute to the process of deliberation but instead provides a time-varying signal that controls the SAT and reflects the growing urgency to commit to a choice. Once a target is selected by cortical regions, GP activity confirms commitment to the decision and invigorates the subsequent movement.

INTRODUCTION

The role of the basal ganglia (BG) in voluntary behavior has been a subject of debate for decades (Dudman and Krakauer, 2016; Turner and Desmurget, 2010). Influential models suggest a role in motor decision making (Frank, 2011; Mink, 1996; Redgrave et al., 1999), whereby desired actions are selected and competing ones suppressed through the direct and indirect dopamine-dependent pathways (Cox et al., 2015; DeLong, 1990; Leblois et al., 2006). This is supported by decision-related modulation of neural activity in the striatum (Ding and Gold, 2010; Samejima et al., 2005) and pallidum (Arimura et al., 2013; Pasquereau et al., 2007), and by reinforcement signals in dopamine neurons (Dayan and Daw, 2008; Schultz, 1997). However, activity in the BG output nucleus, the globus pallidus internus (GPi), reflects choices significantly later than cortical regions (Arimura et al., 2013; Seo et al., 2012) and in simple reaction time tasks is largely simultaneous with muscle contraction (Anderson and Horak, 1985; Mink and Thach, 1991a; Turner and Anderson, 1997). Furthermore, GPi inactivation does not produce deficits in reach target selection (Desmurget and Turner, 2008; Mink and Thach, 1991b) but instead reduces the velocity and extent of movements (Desmurget and Turner, 2010; Horak and Anderson, 1984; Mink and Thach, 1991b), consistent with regulation of motivation and response vigor (Turner and Desmurget, 2010; Yttri and Dudman, 2016). It therefore remains unclear whether the BG play a causal role in decision making and, if so, precisely what that role may be.

Here, we investigate the hypothesis that, at least in well-practiced tasks, the BG do not contribute to the choice between potential movements (which is determined by the cerebral cortex), but instead provide a time-dependent signal that influences both the urgency of commitment to that choice and the vigor of the selected movement (Braunlich and Seger, 2016; Thura and Cisek, 2014, 2016; van Maanen et al., 2016). This predicts that during the process of deliberation, activity in the GPi will be insensitive to the evolving evidence on which the decision is based and will reflect choices only after commitment is made in cortical regions. It also predicts that GPi activity will be time dependent and will reflect an animal's speed-accuracy trade-off (SAT) (Bogacz et al., 2010; Forstmann et al., 2010; Thura et al., 2014).

To test these predictions, we recorded the activity of 107 task-related pallidal neurons (51 GPe; 56 GPi; see Figure S1) in two monkeys performing a reach selection task that allows us to dissociate different aspects of decision making, including the process of deliberation, the moment of commitment, and the adjustment of speed versus accuracy trade-offs. In the "tokens task" (Cisek et al., 2009), monkeys had to guess which of two potential reach targets would receive the majority of 15 tokens that jumped, every 200 ms, from a central circle to one of the two targets (Figure 1A). The monkeys were free to decide at any time, and when the chosen target was reached the remaining token jumps accelerated, either to every 150 ms in "slow" blocks or every 50 ms in "fast" blocks (Figure 1B), which alternated every 75–125 trials. This presents a speed versus accuracy trade-off: the monkeys could either wait until the decision can be made with confidence, or guess ahead of time, which is not as reliable but could improve the overall reward rate. The SAT can be adjusted between the two blocks: hasty decisions are more advantageous in fast blocks than slow blocks because more time is saved by guessing quickly. Although each token jump was completely random, we classified individual trials post hoc based on the pattern of jumps, defining categories of easy, ambiguous, or misleading trials (Figure 1C). For comparison with previous studies, we also recorded these cells in a simple

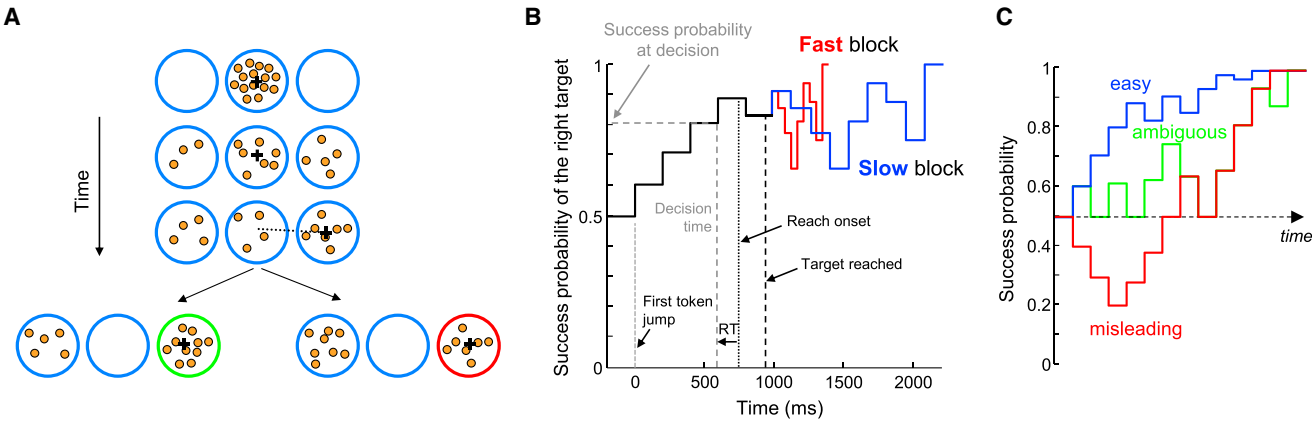

Figure 1. Behavioral Task

(A) The tokens task. Each row illustrates steps in an example trial.

(B) Success probability (SP) profile of an example trial, in which the right target was chosen. Black trace indicates SP before the chosen target is reached, while tokens jump every 200 ms. After the target is reached (vertical black dashed line), remaining tokens jump either every 150 ms (in slow blocks, blue trace) or every 50 ms (fast block, red trace). We subtract from movement onset time (vertical black dotted line) the mean RT calculated daily using a delayed reach task to yield an estimate of the decision time (DT, vertical gray dashed line), and the success probability at which the monkey committed to his choice (horizontal gray dashed line).

(C) Example success probability profiles of "easy" (blue), "ambiguous" (green), or "misleading" (red) trials (see criteria in STAR Methods). See also Figure S2.

delayed reaching task (DR task) in which only a single target is presented and the timing of movement is externally instructed by a GO signal.

Our previous studies of behavior in such dynamic decision-making tasks have led to the "urgency-gating model" (UGM), which suggests that during deliberation, the brain combines sensory evidence in favor of a given target choice with a growing "urgency" signal that is not specific to either target, and commits to a choice when their product reaches a threshold (Cisek et al., 2009; Thura et al., 2012). Consistent with this, we previously showed that neurons in dorsal premotor (PMd) and primary motor (M1) cortex continuously reflect the changing choice-specific evidence conveyed by the tokens and combine it with a non-specific urgency signal that grows over time in each trial (Thura and Cisek, 2014). Approximately 280 ms prior to movement onset, a peak of activity is reached in PMd, signaling commitment to the choice, followed 140 ms later by a peak in M1. We proposed that to maximize reward rate (Thura et al., 2012; Thura et al., 2014), the brain adjusts the SAT by modifying the baseline and slope of the urgency signal. Consistent with this, monkeys made hastier decisions in the fast than the slow blocks (Thura et al., 2014) and deliberation-related PMd/M1 activity was stronger in the fast than the slow block, while the commitment-related peaks were unchanged (Thura and Cisek, 2016). We also found that the vigor of selected movements was correlated with our estimate of urgency (Thura et al., 2014), leading us to hypothesize that a unified urgency/vigor signal comes from the basal ganglia via its projections from the GPi through the thalamus to PMd/M1. This predicts that activity in the GPi will reflect properties of the urgency signal—that is, it will be insensitive to the changing evidence during the process of deliberation but will instead depend on time (showing ramp-like patterns of activity) and reflect volitional adjustment of the SAT (showing block-dependent activation).

RESULTS

Behavioral Data

In the present paper, we report behavioral data collected during sessions in which globus pallidus neurons were recorded (Figure S2). Data gathered during PMd or M1 recording sessions have been published and extensively described in previous publications (Thura and Cisek, 2014; Thura et al., 2014). The present behavioral results replicate our previous observations. First, monkeys' decision duration and success probability at decision time strongly relied upon evidence provided by the distribution of tokens. Animals were faster to decide in easy trials compared to ambiguous or misleading trials. Their success probability at the time of choice was higher in the easy trials. Second, the manipulation of the post-decision interval between token jumps (150 ms in slow blocks versus 50 ms in fast blocks) strongly affected monkeys' strategy: in slow blocks, monkeys were more accurate and more conservative compared to fast blocks. This is because time saving is limited in the former condition, dissuading animals from choosing quickly. This also shows that as expected, monkeys traded accuracy for speed differently in the two conditions. Moreover, except for very fast guesses, animals committed to a choice with less evidence as time passed in each trial (see Figure S2 and Thura et al., 2014). This was true in both conditions, but the evidence needed for commitment was also higher in the slow condition compared to the fast condition. This suggests that a growing sense of urgency pushes animals to commit as time is elapsing. Finally, we previously showed that the influence of urgency extends to the vigor of the subsequent movement. Here, we replicate these findings by showing that decision duration as well as block condition strongly modulate the parameters of animals' reaching movements. Notably, both monkeys shortened their movement durations in fast blocks compared to slow blocks and during long decisions compared to short ones,

Table 1. Motor Discharges in the Globus Pallidus: Count and Percentage of Correlated or Modulated Cells

	GPe		GPi	
Movement-related[a]				
Modulated	26/44	59%	28/51	55%
Increase	15/26	58%	18/28	64%
Decrease	13/26	50%	13/28	46%
Spatial tuning[a]				
Tuned	16/44	36%	20/51	39%
Reaching kinematics[b]				
Reach peak velocity				
Slow block	15/51	29%	10/56	18%
Fast block	10/42	24%	13/52	25%
Reach amplitude				
Slow block	7/51	14%	13/56	23%
Fast block	11/42	26%	7/52	13%
Reach duration				
Slow block	9/51	18%	14/56	25%
Fast block	10/42	24%	20/52	38%
Reach vigor				
Slow block	13/51	25%	14/56	25%
Fast block	10/42	24%	17/52	33%

See also Table S1.

[a]Measured in the delayed reach task (95 cells tested).

[b]Measured in the tokens task (107 cells tested in the slow block, 94 in the fast block).

a pattern that resembles the urgency signal estimated on the basis of their decisions (Thura et al., 2014) in the same conditions. Finally, for one monkey, parameters related to the vigor of movements (reach peak velocity and amplitude) closely followed the adjustment of the context-dependent growing urgency signal derived from their choice behavior (see Thura et al., 2014 for an extensive description of these phenomena).

Neural Activity in the Globus Pallidus

Of the 95 GP cells tested in the DR task, we found that more than half (54/95) showed significant relations to arm movements (Table 1, see also Table S1 for data specific to each monkey), either increasing (33/95) or decreasing (26/95) their firing rate around movement onset (300 ms before until 200 ms after movement onset, Turner and Anderson, 1997) compared to the delay period (a 200 ms epoch preceding the GO signal, Wilcoxon rank-sum test, $p < 0.05$). As previously reported (Anderson and Turner, 1991; Mink and Thach, 1991a; Turner and Anderson, 1997, 2005), peri-movement activity of cells in GPe (36%) and GPi (39%) was directionally tuned in the DR task and often related to kinematic variables such as reach vigor (25% and 24% of GPe cells in the slow and fast blocks of the tokens task, respectively; 25% and 33% of GPi cells), peak velocity (29% and 24% of GPe cells, 18% and 25% of GPi cells) and reach amplitude (14% and 26% of GPe cells, 23% and 13% of GPi cells), implicating them in the "motor loop" of the basal ganglia (Alexander et al., 1990).

To assess how GP neurons contribute to deliberation and to compare this contribution with cortical cells, we first used a multiple linear regression model (see STAR Methods) to investigate how cortical and pallidal activity during deliberation is influenced by the direction of each token jump, the target that is ultimately chosen, the movement vigor, and whether the token jumps into the currently favored (expected) target.

Effect of the Direction of a Token Jump on Neuronal Activity

We found that during the first six token jumps, the multiple linear model significantly explained the neural response of 35%–47% (mean: 42%) of PMd cells, 20%–34% (mean: 28%) of M1 cells, 8%–24% (mean: 15%) of GPe cells, and 5%–21% (mean: 12%) of GPi cells (F values of the model calculated for each cell and after each token jump, $p < 0.05$). Among the four tested factors, the sensory evidence provided by single token jumps was the most influential factor (see Figure S3 for the effects of the other factors). Crucially, the effect of token jump direction was much stronger in cortical cells compared to pallidal cells (Figure 2A). During the first six jumps, the regression coefficient (β_1) significantly differed from 0 (null hypothesis of no effect of a token jump on the neural responses) in 21%–40% of PMd cells, 10%–38% M1 cells, 10%–20% GPe cells, and 5%–16% GPi cells (t-Stat values of the β_1 coefficient calculated for each cell and after each token jump, $p < 0.05$). It is also clear that β_1 distributions were wider in cortical cells compared to pallidal cells (Figure 2A), indicating that although a few GP cells clearly respond to a token jump (Figure 2B), they are overall only weakly modulated in comparison to cortical cells.

Influence of a Target Choice on Neural Activity

Next, to examine whether pallidal cells nevertheless contribute to target selection, we assessed their directional tuning in the tokens task using a receiver-operating characteristic (ROC) analysis on the epoch from 200 ms before commitment until movement onset. We found that 22 GPe (42%) and 27 GPi (48%) cells, although task related (see STAR Methods), did not exhibit any tuning at all and thus could not contribute to reach target selection. The remaining cells (29/51 in the GPe, 29/56 in the GPi) showed significant tuning but primarily around movement onset. To determine whether these tuned cells contribute to deliberation about the selected target, we analyzed their activity during trials classified as easy, ambiguous, and misleading, either toward or away from each cell's preferred target (PT). We found that in contrast with PMd and M1 (Figures 3A and 3B), pallidal activity only weakly reflects the evolving sensory evidence (Figures 3C and 3D). In particular, while there is a weak effect of sensory evidence on GPe activity, it is virtually absent in the GPi, the output nucleus of the BG. Note that this difference of influence of sensory evidence on cortical versus pallidal cells is not due to differences in cell count, as the time course of evidence can clearly be seen in PMd and M1 even at the level of individual neurons (Thura and Cisek, 2014), but this was never observed in GPe or GPi.

To further assess these regional differences, we calculated the correlation between cell activity and sensory evidence at different moments in time within trials. This analysis allows us to include all trial types and not only the special ones mentioned above. As

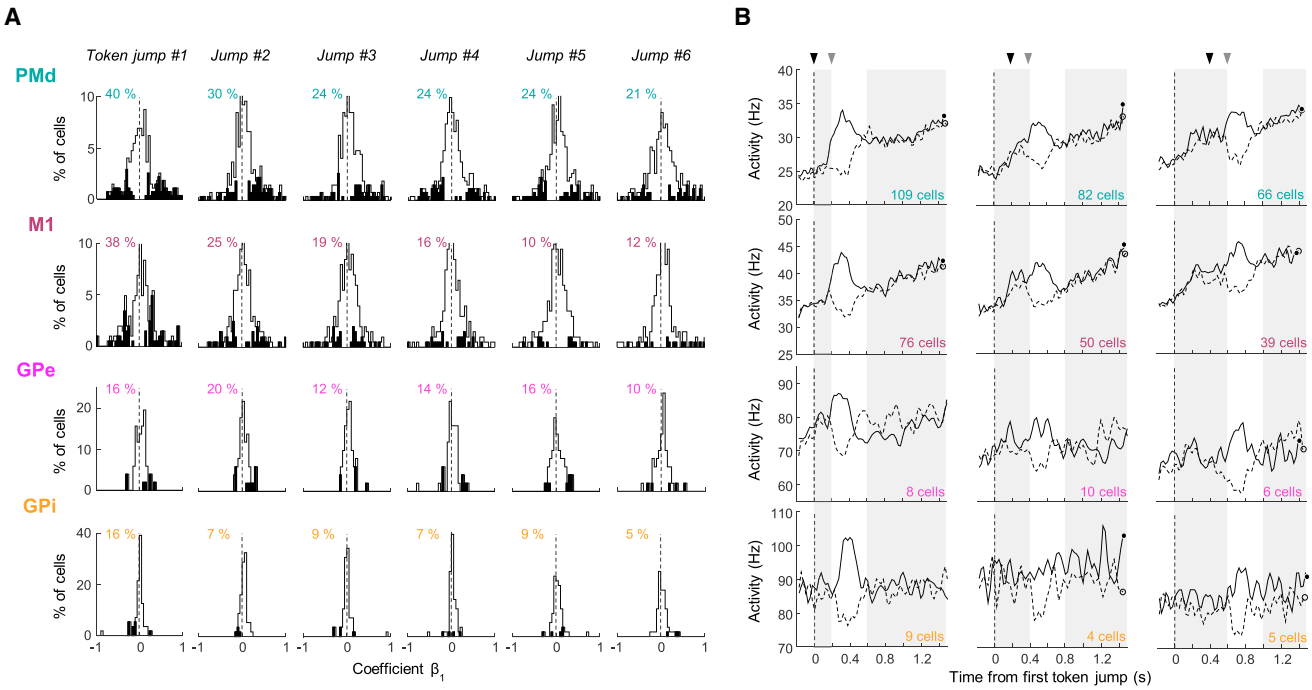

Figure 2. Effect of Token Jumps on Cortical and Pallidal Activity

(A) Distribution of regression coefficients reflecting the influence of each token jump on the neural response of PMd (top panels), M1 (second row panels), GPe (third row panels), and GPi (bottom panels) cells recorded in the slow condition. Coefficients are computed after each of the first six jumps (columns from left to right). Black histograms illustrate cells for which a token jump significantly modulated the neural response according to the multiple linear regression model (Equation 3). Percentages of significantly modulated cells are indicated above each panel.

(B) Left column: responses of PMd (top row), M1 (second row), GPe (third row), and GPi (bottom row) cells showing a significant effect of the first token jump direction (black histogram in A), averaged during trials in which the first token (black triangle) jumped into the cells' PT and the second token (gray triangle) jumped into the OT (solid curves) or trials in which the first token jumped into the OT and the second jumped into the PT (dashed curves). Middle column: same for cells that are significantly influenced by token jump #2, during trials in which the second token (black triangle) jumped into cells' PT and the third token (gray triangle) jumped into cells' OT (solid) or vice versa (dashed). Right column: same for cells that are significantly influenced by token jump #3, during trials in which the third token (black triangle) jumped into the PT and the fourth (gray triangle) jumped into the OT (solid) or vice versa (dashed). See also Figure S3.

reported previously (Thura and Cisek, 2014), we observed that PMd cells and (to a lesser extent) M1 cells are strongly and positively correlated with the available sensory evidence, from the beginning of deliberation until commitment (Spearman correlation coefficient, averaged across neurons and time, $\rho = 0.49$ in PMd, $\rho = 0.45$ in M1). This correlation is weaker in the GP ($\rho = 0.37$ in GPe; $\rho = 0.26$ in GPi). Most importantly, when normalized against background firing rate, the gain between GPe/GPi activity and sensory evidence is close to zero (Figure 3E). Moreover, we see a gradient of sensitivity: PMd shows the strongest gain that tends to increase with time; M1 cells exhibit a weaker and stable gain; while GPe and GPi activity is nearly insensitive to sensory evidence during the entire deliberation period (ANCOVA, main effect of recording area on slope, $F = 14.46$, $p < 0.0001$). In short, the evolving evidence on which the decision is based is clearly reflected in PMd and M1, but it is virtually absent across GPi cells. Thus, we conclude that BG output does not contribute to determining the target choice in our task.

It has been suggested that, in addition to firing rates, neural variability across trials can provide further insights about the neural computations underlying deliberation, choice commitment, and action execution (Churchland et al., 2006, 2010, 2011). We thus assessed the possibility that the contribution of BG to action selection takes part through changes in response variance rather than in changes in mean firing rates. Using the Fano factor as a measure of neural variability across trials (Churchland et al., 2006), we first replicated previous findings showing that premotor and motor neural variability decreases before the execution of reaching movements (Figure S4B) (Churchland et al., 2010). We also found such a decrease of variability in GP cells before movements, although this was weaker in GPi. Importantly, however, when aligned on the first token jump, we observed that across-trial variability in GP neurons was unaffected during the deliberation process (Figure S4A). By contrast, we noted a slight increase of variability in cortical neurons during decision formation, especially in PMd, as previously demonstrated in the oculomotor system (Churchland et al., 2011). It is also interesting to note that variability in the GPi was lower than cortical variability overall, with a Fano factor significantly lower than 1.

Effect of Speed-Accuracy Trade-Off on GP Activity during Deliberation

The results described above suggest that GP cells are only weakly involved in the process of selecting a reach target. But

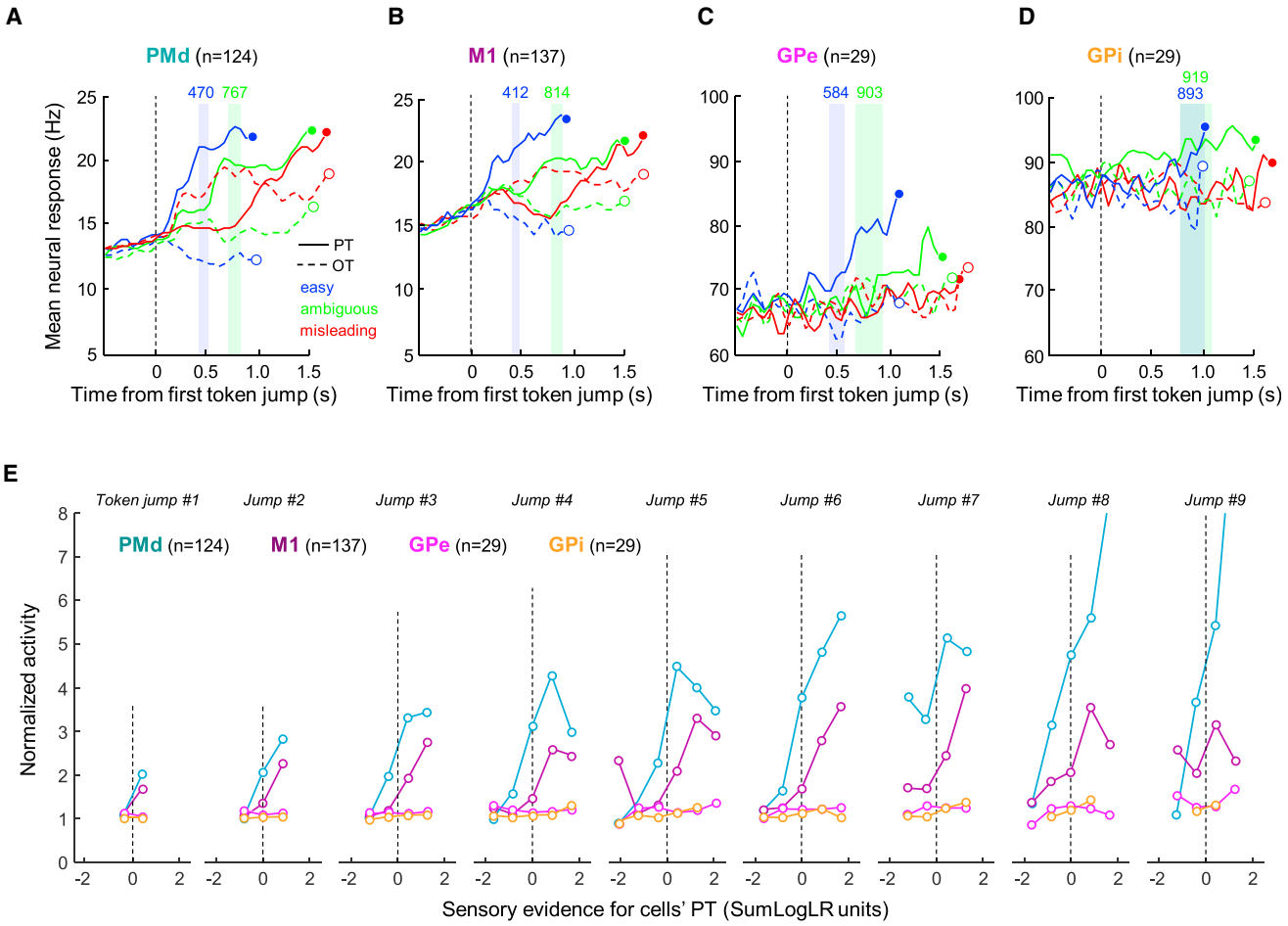

Figure 3. Comparison of Activity during Deliberation in PMd, M1, GPe, and GPi

(A) Average response, during the slow block, of 124 tuned PMd neurons during easy (blue), ambiguous (green), and misleading (red) trials in which the monkey correctly chose the PT (solid curves) or OT (dashed curves). Activity is aligned on the first token jump, and spikes recorded after decision commitment (circles) are discarded to avoid averaging artifacts. The widths of the colored horizontal bars illustrate the mean ± SE of cells' discrimination time in easy (blue) and ambiguous (green) trials, with numbers indicating the mean.

(B) Same as (A) for 137 tuned M1 cells.

(C) Same as (A) for 29 tuned GPe cells.

(D) Same as (A) for 29 tuned GPi cells.

(E) Each panel illustrates how activity varies as a function of evidence (calculated as SumLogLR with respect to cells' PT), 200 ms after each of the first nine token jumps, in the same PMd (blue-green), M1 (violet), GPe (magenta), and GPi (orange) cells shown in (A)–(D). Activity is normalized against baseline and spikes recorded after decision commitment are discarded. See also Figure S4.

does this mean that these cells do not contribute to any aspect of decision making? Several previous theoretical and experimental investigations suggested that the basal ganglia are crucial for the adjustment of speed-accuracy trade-offs during decision formation (Bogacz et al., 2010; Forstmann et al., 2010). Our previous results are consistent with the hypothesis that SAT occurs through adjustments of an urgency signal, possibly originating from the basal ganglia, to regulate decision-related activity in cortical structures (see Thura and Cisek, 2016; Thura et al., 2014).

To investigate whether GP activity reflects urgency, we examined individual cells for time-dependent changes of activity within each trial (see STAR Methods). We found that some cells (21/51 in GPe, 13/56 in GPi) gradually increased their activity during deliberation ("build-up" cells), and some cells showed the opposite pattern ("decreasing" cells, 8/51 GPe and 11/56 GPi). The remaining ~50% cells had "steady" activity during deliberation. Interestingly, build-up cells in the GPi tended to be more sensitive to sensory evidence compared to other neuron types, although still much less sensitive and highly lagged compared to cortical cells (Figure S5).

To test whether these time-dependent activities could be the neural correlates of the context-dependent urgency signal, we examined whether they changed between the slow and fast blocks (42/51 GPe and 52/56 GPi cells tested in both conditions) to reflect the predicted adjustment of the urgency signal with changes in the SAT (Thura et al., 2014). We found that most GP cells (70/94, 75%) were indeed significantly modulated by

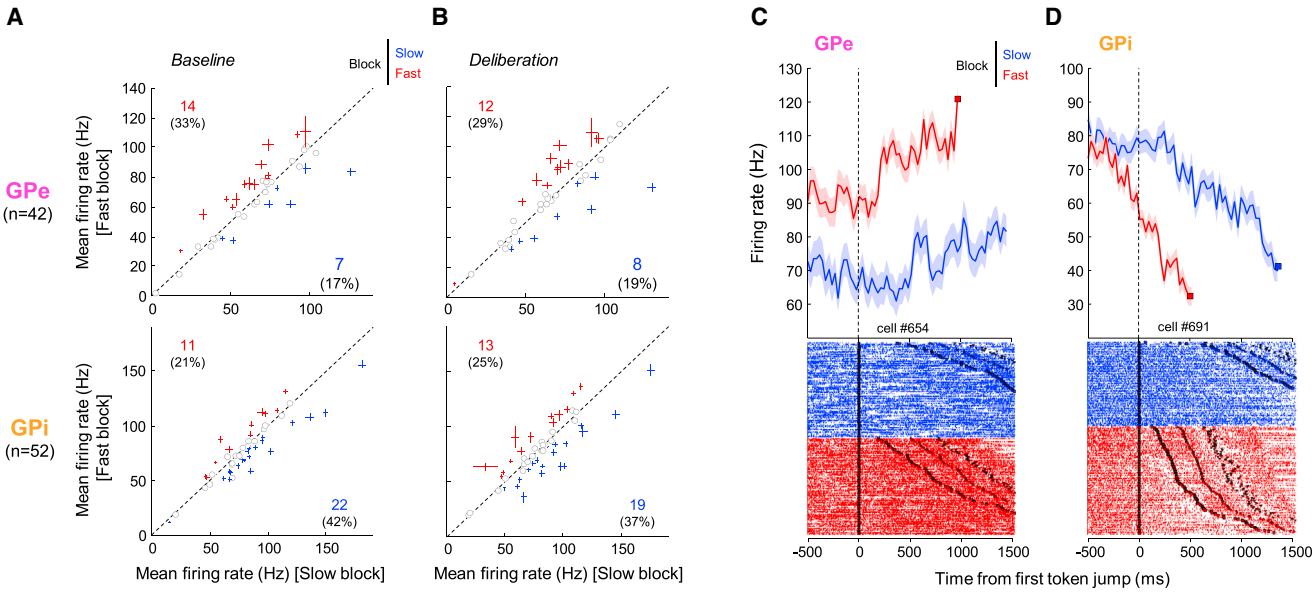

Figure 4. Effect of SAT Condition on GP Activity

(A) Comparison of mean neural activity, for both targets, of 42 GPe (top) and 52 GPi (bottom) neurons recorded during all trials in the slow block (abscissa) versus the fast block (ordinate) in a 200 ms period preceding the first token jump (baseline period). Colored crosses illustrate the mean (±SE) activity of cells significantly modulated by the block condition. Red represents stronger activity in the fast blocks. Blue represents stronger activity in the slow blocks. Circles indicate cells that were not significantly modulated between the blocks. Numbers indicate the number and percentage (in parenthesis) of significantly modulated cells.

(B) Same as (A), but here, activity is compared between blocks in a 200 ms epoch extending from 400 ms to 600 ms after the first token jump (deliberation period). For this analysis, trials in which decision duration was below 600 ms were excluded.

(C) Example of a GPe neuron's response in the slow (blue) and fast (red) blocks of trials in the tokens task. Activity is aligned on the first token jump and illustrated as spike density functions (mean ± SE, computed in bins of 30 ms, top) and rasters displays (bottom). Spikes after commitment time (280 ms before movement onset, squares) are discarded from the spike density functions.

(D) Example of a GPi neuron's response in the two speed conditions of the tokens task. Same conventions as in (C).

the block, either before the start of the trial (baseline period, GPe: 8/42, 19%; GPi: 9/52, 17%), during deliberation (GPe: 7/42, 17%; GPi: 8/52, 15%), or during both epochs (Figures 4A and 4B). Interestingly, while less than one-third (13/42) of the GPe cells are modulated by the SAT context during both baseline and deliberation, this proportion rises to reach almost 50% (24/52) in the GPi. Two examples of these modulated cells are shown in Figures 4C and 4D. The GPe cell, classified as a build-up cell, is more active in the fast condition compared to the slow condition. The GPi cell, classified as a decreasing cell, shows the opposite pattern. Crucially, we found that at the population level, the modulation between blocks was congruent with the temporal profile within each block (Figure 5): build-up cells tended to be *more* active during fast than slow blocks, especially in the GPe; while decreasing cells were usually *less* active in fast than slow blocks, especially in the GPi. Steady cells, at the population level, were less modulated. Most importantly, there was a significant correlation between within-trial profiles and between-block modulation in both GPe and GPi, and this was stronger in the latter (Pearson's correlation, GPe: $r = 0.31$, $p = 0.04$; GPi: $r = 0.42$, $p < 0.002$; Figures 5G and 5H). This means that the activity of cells with the highest positive (negative) slope was most strongly increased (decreased) in the fast block relative to the slow block. Additional control analyses suggest that these activity patterns are not simply related to the anticipation of reward, because the baseline responses are not reduced during trials in which the monkey is informed that no reward will be delivered (see Figure S6).

Neural Activity at the Time of Commitment

If the GPe and GPi do not contribute to target selection in our task, then why have previous studies (Arimura et al., 2013; Pasquereau et al., 2007) shown selection-related activity in these nuclei? As noted above, in our task deliberation between targets can be dissociated from commitment, but this was not the case in previous studies. This raises the possibility that the selection-related activity reported previously reflected the *commitment* to the selected target and not the process by which that target was selected.

To examine this possibility, we again compared activity in PMd, M1, GPe, and GPi but now aligned to movement onset (Figures 6A–6D). In each region, activity tuned to the preferred target (PT) reaches a peak before movement execution, regardless of trial type. That peak is earlier in PMd than in M1, GPe, or GPi (Figure 6E). A running ROC analysis performed on all cells with activity aligned to movement onset showed that choice-selective tuning emerges earlier and is much more common in PMd and M1 compared to GPe and GPi (Figure 6F). Notably, while some tuning begins to appear in the GP 200–300 ms before commitment, it remains rare (<10%) and then increases dramatically just around the moment when the PMd population exhibits its prominent peak (Figure 6F). Indeed, in both GPe and GPi,

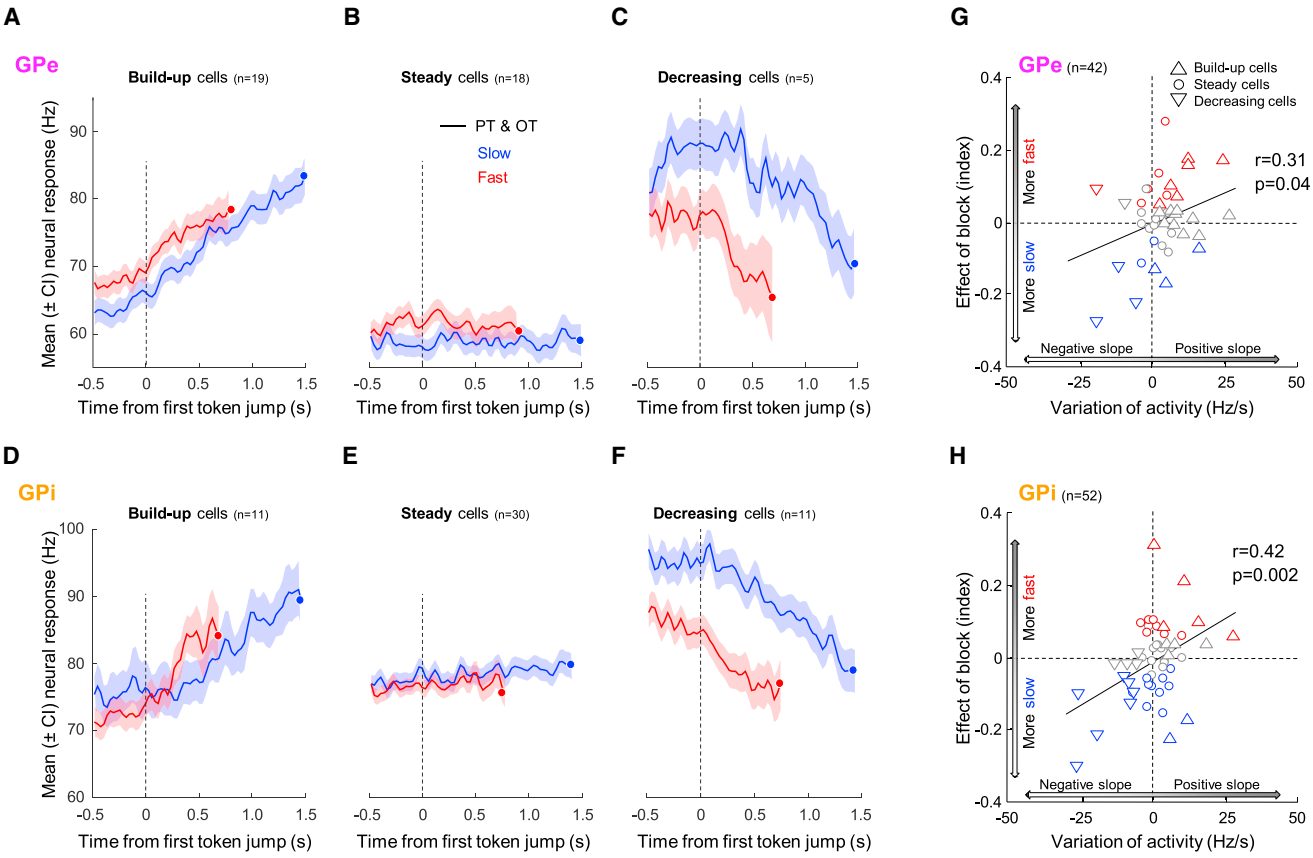

Figure 5. Time-Dependent Activity in the Globus Pallidus

(A) Average activity (with 95% confidence intervals) of 19 "build-up" GPe cells aligned on the first token jump during the fast (red) and slow blocks (blue) and truncated before decision commitment (circles). Both correct and errors trials are included.

(B) Same as (A) for a population of 18 "steady" GPe cells.

(C) Same as (A) for 5 "decreasing" GPe cells.

(D–F) Same as (A)–(C) for the same three categories of cells in the GPi.

(G) Correlation between the average variation of activity during deliberation of the 42 GPe neurons recorded in both blocks (calculated as a slope, including slow block trials from both targets, computed from the first token jump until decision commitment) and the modulation of the block condition on activity during deliberation (400 to 600 ms after the first token jump). The index of modulation by the block condition is calculated as $(FR_{fast} - FR_{slow})/(FR_{fast} + FR_{slow})$. The black line illustrates the result of a linear regression through the data. Cells whose activity was significantly higher (lower) in the fast block than the slow block are plotted in red (blue) and symbols indicate whether cells were classified as build-up, steady, or decreasing (see legend).

(H) Same as (G) for 52 GPi cells. See also Figures S5 and S6.

a sharp rise of activity tuned to the selected target conspicuously occurs around that "moment of commitment" (Figures 6C and 6D).

DISCUSSION

Although numerous theories have implicated the basal ganglia in decision making, along with fronto-parietal cortical regions (Bogacz et al., 2010; Cisek, 2007; Forstmann et al., 2010; Frank, 2011; Mazzoni et al., 2007; Mink, 1996; Redgrave et al., 1999; Thura et al., 2014; Turner and Desmurget, 2010), their precise contribution is still under debate (Dudman and Krakauer, 2016; Turner and Desmurget, 2010). In the present study, we aimed to determine their potential contribution to the component processes of deliberation between target choices, commitment to a single choice, and the adjustment of the SAT through neural recordings in the BG output nuclei. Our data show that during deliberation, information pertinent to selection of reaching movements is continuously influencing activity in reach-related regions of PMd and M1 but is much weaker and significantly delayed in the BG, particularly in its output via the GPi (Figures 2 and 3). This argues against the role of the BG in the deliberation process that determines which reach target is selected.

Instead, we found ramp-like patterns of activity, in both GPe and GPi, that were modulated by the SAT condition in terms of their baseline (Figures 5A, 5C, and 5F), and to a lesser extent their slope (Figure 5A). Our data are compatible with the hypothesis that the BG invigorate the decision-making process by providing an urgency signal that modulates the gain of how sensory evidence influences PMd and M1 activity (Thura and Cisek, 2014, 2016). The urgency signal grows over time and is adjusted to modulate the SAT such that higher urgency leads to hastier

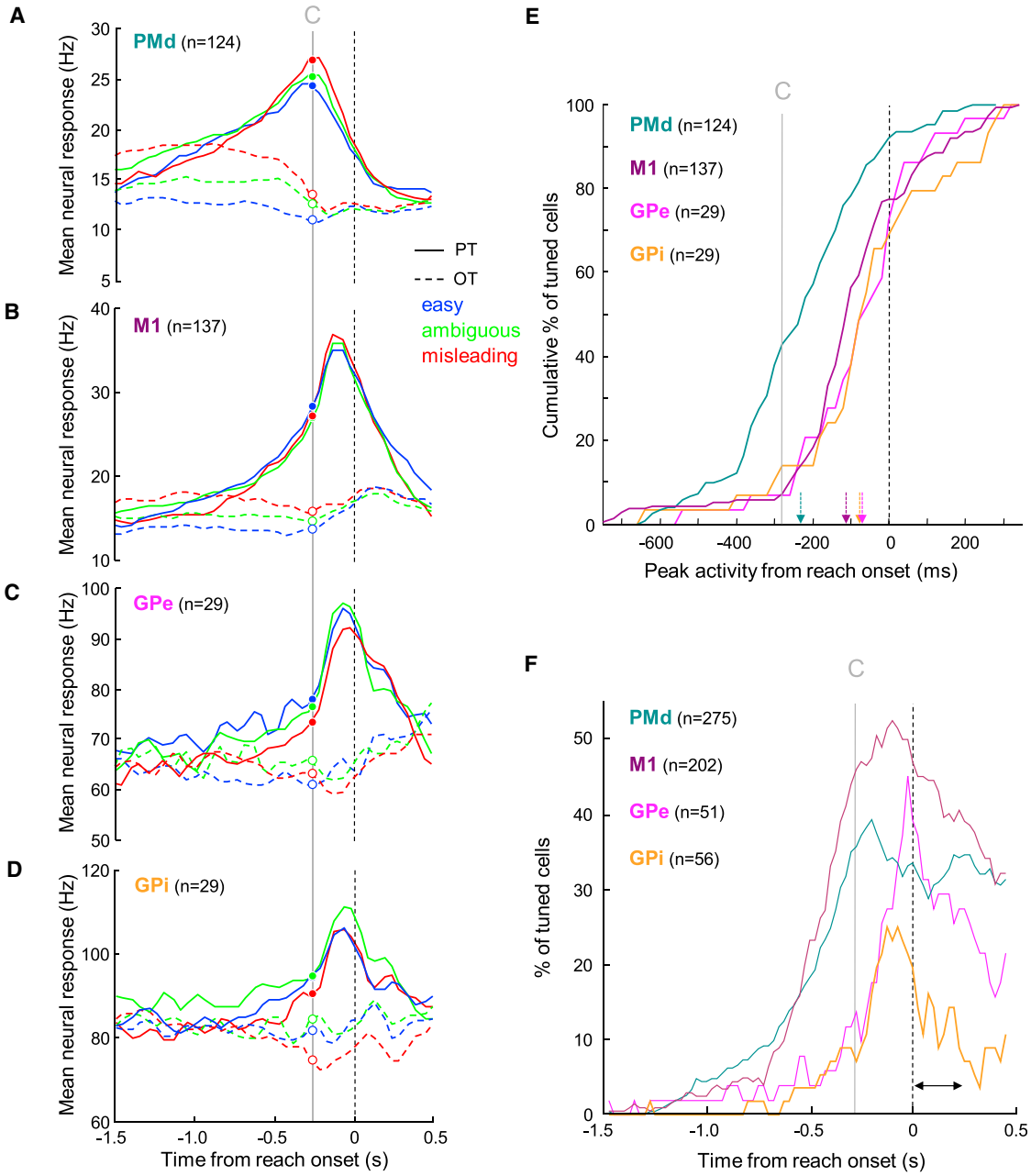

Figure 6. Comparison of Activity during Commitment in PMd, M1, GPe, and GPi

(A) Activity of the 124 tuned PMd neurons shown in Figure 3A but here aligned on movement onset. Circles and the vertical gray line indicate our estimate of the moment of commitment (Thura and Cisek, 2014).

(B) Same as (A) for 137 tuned M1 cells.

(C) Same as (A) for 29 tuned GPe cells.

(D) Same as (A) for 29 tuned GPi cells.

(E) Cumulative distributions of the timing of peak activity, relative to movement onset, of tuned cells recorded in PMd (blue-green), M1 (violet), GPe (magenta), and GPi (orange). The average timing of each cell's maximum firing rate is calculated for trials in which the monkey correctly chose the cell's preferred target, in bins of 20 ms from 750 ms before to 350 ms after movement initiation. Colored dotted and vertical arrows mark for each area the median of the distribution (PMd: −230 ms; M1: −110 ms; GPe and GPi: −70 ms). The distribution of activity peak timing is significantly earlier in PMd compared to the other areas (Wilcoxon rank-sum test, $p < 0.001$).

(F) Percentage of all PMd (blue-green), M1 (violet), GPe (magenta), and GPi (orange) cells showing target-selective activity (ROC score above 0.65 for at least 3 successive bins) as a function of time relative to movement onset in 30 ms time bins. The horizontal arrow marks the average movement period.

guesses in the fast block (Figures 4 and 5). Such a policy maximizes reward rate (Thura et al., 2012), consistent with the hypothesis that the BG motivate voluntary behavior (Pasquereau et al., 2007) and influence movement vigor (Dudman and Krakauer, 2016; Turner and Desmurget, 2010). However, several important questions remain unanswered. For example, while we observed that the majority of build-up neurons were in GPe and the majority of decreasing neurons were in GPi, these were not exclusive. In the context of current models of the basal ganglia, it is more difficult to interpret why there should exist decreasing cells in GPe and build-up cells in GPi, raising challenges for future work.

Although our data argue against models that ascribe a causal role of BG in selecting the target choice for goal-directed movement (Mink, 1996; Redgrave et al., 1999), they do implicate the BG in the final commitment to that choice. Our results are consistent with the proposal that cortical activity reflects a dynamic, biased competition between candidate actions (Cisek, 2007), which is gradually amplified by an urgency signal from the BG that effectively controls the amount of evidence needed before the animal commits to the currently favored reach choice. As the cortical bias grows in favor of one of the targets, it begins to influence activity in the GPe, producing the gradual emergence of tuning in the 200–300 ms before commitment (Figure 6). When that becomes strong enough to engage tuning in the GPi, the BG "gate" is opened, leading to a positive feedback that constitutes commitment to the action choice (Brown et al., 2004). The opening of that gate could involve both the suppression of GPi cells than inhibit the selected movement as well as increased activity of GPi cells that inhibit unselected movements, producing the conspicuous increase of tuned activity that we observed just after the cortical moment of commitment (Figure 6F).

It is important to note that our results are from a highly practiced task, which the monkeys have been performing for several years before we began recording in the GP. Strong evidence exists that the basal ganglia are implicated in choices during reinforcement learning (RL). For example, midbrain dopamine neurons encode a key component of RL models, the reward prediction error (e.g., Bayer and Glimcher, 2005), and disruption of information processing in the basal ganglia during learning strongly affects choice behavior (e.g., Knowlton et al., 1996; Piron et al., 2016). It has also been proposed that as a task is practiced, the associations between stimuli and actions initially encoded in the basal ganglia are consolidated in the neocortex (Ashby et al., 2007; Hadj-Bouziane et al., 2003). Thus, it is possible that in highly practiced tasks like ours, the basal ganglia are less critical for decision making than in tasks involving learning (Piron et al., 2016). Furthermore, while our data argue against the influence of the BG on selection of actions based on probabilistic information, they may be involved in selecting choices based on their value. Indeed, the BG receive critical information related to action value via the nigrostriatal dopamine signals (Kim et al., 2015; Morris et al., 2006) and future work should investigate whether this information influences value-based deliberation in the BG output nuclei.

Finally, while the BG urgency signal is not specific to the selection of the reach target in our task, it may be specific to the monkey's motivation to continue performing the task as opposed to doing something else. Indeed, our results are compatible with the hypothesis that while BG do not select potential actions within a given behavioral system (e.g., reach right versus reach left), they select which behavioral system (e.g., reaching versus locomotion versus feeding) is prioritized in a given situation (Grillner et al., 2013; Redgrave et al., 1999). More generally, the present results lead toward a role of the BG that strongly depends on motivational factors like reward rate maximization (Haith et al., 2012; Shadmehr, 2010). A better understanding of such motor operations in concert with non-motor factors should provide crucial information about BG-related neurological diseases affecting both the motor and cognitive aspects of behavior, as in Parkinson's disease or impulse control disorders (Mazzoni et al., 2007).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCES SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHODS DETAILS
  - Surgery
  - Apparatus
  - Behavioral tasks
  - Dataset
  - Neural recordings in PMd and M1
  - Neural recording in the globus pallidus
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral data
  - Neural data analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2017.07.039.

AUTHOR CONTRIBUTIONS

Conceptualization, P.C.; Methodology, P.C. and D.T.; Investigation, D.T.; Writing – Original Draft, D.T.; Writing – Review & Editing, P.C. and D.T.; Funding Acquisition, P.C. and D.T.; Resources, P.C.; Supervision, P.C.

ACKNOWLEDGMENTS

We are grateful to John Kalaska and Trevor Drew for comments on the manuscript, and Robert Turner for suggestions on techniques for recording activity in the GP. This work was supported by Canadian Institutes of Health Research Grant MOP-102662, the Canadian Foundation for Innovation, Fonds de Recherche en Santé du Québec, the EJLB Foundation to P.C., and fellowships from the FYSSEN Foundation and the Groupe de Recherche sur le Système Nerveux Central to D.T.

Received: April 12, 2017
Revised: July 7, 2017
Accepted: July 29, 2017
Published: August 17, 2017

REFERENCES

Alexander, G.E., Crutcher, M.D., and DeLong, M.R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Prog. Brain Res. *85*, 119–146.

Anderson, M.E., and Horak, F.B. (1985). Influence of the globus pallidus on arm movements in monkeys. III. Timing of movement-related information. J. Neurophysiol. *54*, 433–448.

Anderson, M.E., and Turner, R.S. (1991). A quantitative analysis of pallidal discharge during targeted reaching movement in the monkey. Exp. Brain Res. *86*, 623–632.

Arimura, N., Nakayama, Y., Yamagata, T., Tanji, J., and Hoshi, E. (2013). Involvement of the globus pallidus in behavioral goal determination and action specification. J. Neurosci. *33*, 13639–13653.

Ashby, F.G., Ennis, J.M., and Spiering, B.J. (2007). A neurobiological theory of automaticity in perceptual categorization. Psychol. Rev. *114*, 632–656.

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron *47*, 129–141.

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

Bogacz, R., Wagenmakers, E.J., Forstmann, B.U., and Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. Trends Neurosci. *33*, 10–16.

Bogacz, R., Martin Moraud, E., Abdi, A., Magill, P.J., and Baufreton, J. (2016). Properties of Neurons in External Globus Pallidus Can Support Optimal Action Selection. PLoS Comput. Biol. *12*, e1005004.

Braunlich, K., and Seger, C.A. (2016). Categorical evidence, confidence, and urgency during probabilistic categorization. Neuroimage *125*, 941–952.

Brown, J.W., Bullock, D., and Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. Neural Netw. *17*, 471–510.

Churchland, M.M., Yu, B.M., Ryu, S.I., Santhanam, G., and Shenoy, K.V. (2006). Neural variability in premotor cortex provides a signature of motor preparation. J. Neurosci. *26*, 3697–3712.

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. *13*, 369–378.

Churchland, A.K., Kiani, R., Chaudhuri, R., Wang, X.J., Pouget, A., and Shadlen, M.N. (2011). Variance as a signature of neural computations during decision making. Neuron *69*, 818–831.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. Philos. Trans. R. Soc. Lond. B Biol. Sci. *362*, 1585–1599.

Cisek, P., Crammond, D.J., and Kalaska, J.F. (2003). Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm. J. Neurophysiol. *89*, 922–942.

Cisek, P., Puskas, G.A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. J. Neurosci. *29*, 11560–11571.

Cox, S.M., Frank, M.J., Larcher, K., Fellows, L.K., Clark, C.A., Leyton, M., and Dagher, A. (2015). Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. Neuroimage *109*, 95–101.

Dayan, P., and Daw, N.D. (2008). Decision theory, reinforcement learning, and the brain. Cogn. Affect. Behav. Neurosci. *8*, 429–453.

DeLong, M.R. (1971). Activity of pallidal neurons during movement. J. Neurophysiol. *34*, 414–427.

DeLong, M.R. (1990). Primate models of movement disorders of basal ganglia origin. Trends Neurosci. *13*, 281–285.

Desmurget, M., and Turner, R.S. (2008). Testing basal ganglia motor functions through reversible inactivations in the posterior internal globus pallidus. J. Neurophysiol. *99*, 1057–1076.

Desmurget, M., and Turner, R.S. (2010). Motor sequences and the basal ganglia: kinematics, not habits. J. Neurosci. *30*, 7685–7690.

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

Ditterich, J. (2010). A Comparison between Mechanisms of Multi-Alternative Perceptual Decision Making: Ability to Explain Human Behavior, Predictions for Neurophysiology, and Relationship with Decision Theory. Front. Neurosci. *4*, 184.

Dudman, J.T., and Krakauer, J.W. (2016). The basal ganglia: from motor commands to the control of vigor. Curr. Opin. Neurobiol. *37*, 158–166.

Forstmann, B.U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.J., Bogacz, R., and Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. Proc. Natl. Acad. Sci. USA *107*, 15916–15920.

Frank, M.J. (2011). Computational models of motivated action selection in corticostriatal circuits. Curr. Opin. Neurobiol. *21*, 381–386.

Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (Wiley).

Grillner, S., Robertson, B., and Stephenson-Jones, M. (2013). The evolutionary origin of the vertebrate basal ganglia and its role in action selection. J. Physiol. *591*, 5425–5431.

Hadj-Bouziane, F., Meunier, M., and Boussaoud, D. (2003). Conditional visuomotor learning in primates: a key role for the basal ganglia. J. Physiol. Paris *97*, 567–579.

Haith, A.M., Reppert, T.R., and Shadmehr, R. (2012). Evidence for hyperbolic temporal discounting of reward in control of movements. J. Neurosci. *32*, 11727–11736.

Horak, F.B., and Anderson, M.E. (1984). Influence of globus pallidus on arm movements in monkeys. I. Effects of kainic acid-induced lesions. J. Neurophysiol. *52*, 290–304.

Kim, H.F., Ghazizadeh, A., and Hikosaka, O. (2015). Dopamine Neurons Encoding Long-Term Memory of Object Value for Habitual Behavior. Cell *163*, 1165–1175.

Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. Science *273*, 1399–1402.

Leblois, A., Boraud, T., Meissner, W., Bergman, H., and Hansel, D. (2006). Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. J. Neurosci. *26*, 3567–3583.

Mazzoni, P., Hristova, A., and Krakauer, J.W. (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. J. Neurosci. *27*, 7105–7116.

Middleton, F.A., and Strick, P.L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res. Brain Res. Rev. *31*, 236–250.

Mink, J.W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. Prog. Neurobiol. *50*, 381–425.

Mink, J.W., and Thach, W.T. (1991a). Basal ganglia motor control. II. Late pallidal timing relative to movement onset and inconsistent pallidal coding of movement parameters. J. Neurophysiol. *65*, 301–329.

Mink, J.W., and Thach, W.T. (1991b). Basal ganglia motor control. III. Pallidal ablation: normal reaction time, muscle cocontraction, and slow movement. J. Neurophysiol. *65*, 330–351.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., and Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. Nat. Neurosci. *9*, 1057–1063.

Pasquereau, B., Nadjar, A., Arkadir, D., Bezard, E., Goillandeau, M., Bioulac, B., Gross, C.E., and Boraud, T. (2007). Shaping of motor responses by incentive values through the basal ganglia. J. Neurosci. *27*, 1176–1183.

Piron, C., Kase, D., Topalidou, M., Goillandeau, M., Orignac, H., N'Guyen, T.H., Rougier, N., and Boraud, T. (2016). The globus pallidus pars interna in goal-oriented and routine behaviors: Resolving a long-standing paradox. Mov. Disord. *31*, 1146–1154.

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

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. Science *310*, 1337–1340.

Sato, T.R., and Schall, J.D. (2003). Effects of stimulus-response compatibility on neural selection in frontal eye field. Neuron *38*, 637–648.

Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. Curr. Opin. Neurobiol. *7*, 191–197.

Seo, M., Lee, E., and Averbeck, B.B. (2012). Action selection and action value in frontal-striatal circuits. Neuron *74*, 947–960.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. *16*, 1486–1510.

Shadmehr, R. (2010). Control of movements and temporal discounting of reward. Curr. Opin. Neurobiol. *20*, 726–730.

Thura, D., and Cisek, P. (2014). Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making. Neuron *81*, 1401–1416.

Thura, D., and Cisek, P. (2016). Modulation of premotor and primary motor cortical activity during volitional adjustments of speed-accuracy trade-offs. J. Neurosci. *36*, 938–956.

Thura, D., Beauregard-Racine, J., Fradet, C.W., and Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. J. Neurophysiol. *108*, 2912–2930.

Thura, D., Cos, I., Trung, J., and Cisek, P. (2014). Context-dependent urgency influences speed-accuracy trade-offs in decision-making and movement execution. J. Neurosci. *34*, 16442–16454.

Turner, R.S., and Anderson, M.E. (1997). Pallidal discharge related to the kinematics of reaching movements in two dimensions. J. Neurophysiol. *77*, 1051–1074.

Turner, R.S., and Anderson, M.E. (2005). Context-dependent modulation of movement-related discharge in the primate globus pallidus. J. Neurosci. *25*, 2965–2976.

Turner, R.S., and Desmurget, M. (2010). Basal ganglia contributions to motor control: a vigorous tutor. Curr. Opin. Neurobiol. *20*, 704–716.

van Maanen, L., Fontanesi, L., Hawkins, G.E., and Forstmann, B.U. (2016). Striatal activation reflects urgency in perceptual decision making. Neuroimage *139*, 294–303.

Yttri, E.A., and Dudman, J.T. (2016). Opponent and bidirectional control of movement velocity in the basal ganglia. Nature *533*, 402–406.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Rhesus macaque (*Macaca mulatta*)	N/A	N/A
Software and Algorithms		
MATLAB 2014b	MathWorks	https://www.mathworks.com/product/matlab.html
LabVIEW 8.2	National Instruments	http://www.ni.com/fr-fr/shop/labview.html
Brainsight	Rogue Research	https://www.rogue-research.com/veterinary/research/
Plexon Offline Sorter	Plexon	http://www.plexon.com/products/offline-sorter
Fano Factor computation algorithm	Mark Churchland's lab website	http://churchlandlab.neuroscience.columbia.edu/links.html
Other		
AlphaLab acquisition and recording system	Alpha-Omega	https://alphaomega-eng.com/alphalab-snr.html
Microelectrodes (tungsten, glass coated)	FHC	https://www.fh-co.com/category/microelectrodes
Microelectrodes (tungsten, glass coated)	Alpha-Omega	https://alphaomega-eng.com//electrodes-arrays/acute-electrode.html
Microdrive	NAN Instruments	http://naninstruments.com/
Eye tracker	ASL	N/A
Digitizing Tablet	GTCO CalComp	http://www.gtcocalcomp.com/large-format-digitizers

CONTACT FOR REAGENT AND RESOURCES SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, David Thura (david.thura@umontreal.ca).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Experiments were conducted with two male macaque monkeys (*Macaca mulatta*; S: 4-9 years old, 5-9 kg; Z: 4-6 years old, 4-7 kg). Both animals were pair-housed in a vivarium with a 12 hr light cycle (6am to 6pm). The local animal ethics committee approved surgery, testing procedure, and animal care.

METHODS DETAILS

Surgery

Both animals were first implanted, under anesthesia and aseptic conditions, with a titanium head fixation post for head stabilization. Then, after recovery and head-fixed training, animals were implanted with recording chambers allowing vertical penetration of microelectrodes into the brain via trans-dural guide tubes. A grid, positioned in the horizontal plane and composed of holes (1mm spacing), was used to ensure an accurate mapping of the successive penetrations performed in PMd, M1 and the basal ganglia.

Apparatus

Monkeys sat head-fixed in a custom primate chair and performed two planar reaching tasks using a vertically-oriented cordless stylus whose position was recorded by a digitizing tablet (*CalComp*, 125 Hz). Their non-acting hand was restrained on an arm rest with Velcro bands. In some sessions, unconstrained eye movements were recorded using an infrared camera (*ASL*, 120 Hz). Stimuli and continuous cursor feedback were projected onto a mirror suspended between the monkey's gaze and the tablet, creating the illusion that they are in the plane of the tablet. Neural activity was recorded from the hemisphere contralateral to the acting hand with 1-4 independently moveable (*NAN* microdrive) microelectrodes (*FHC, Alpha-Omega Eng.*) and data were acquired with the AlphaLab system (*Alpha-Omega Eng.*).

Behavioral tasks

Monkeys were trained to perform the "tokens" task (Figure 1A) in which they are presented with one central starting circle (1.75 cm radius) and two peripheral target circles (1.75 cm radius, arranged at 180° around a 5 cm radius circle). The monkey begins each trial by placing a handle in the central circle, in which 15 small tokens are randomly arranged. The tokens then begin to jump, one-by-one every 200 ms ("pre-decision interval"), from the center to one of the two peripheral targets. The monkey's task is to move the handle to the target that he believes will ultimately receive the majority of tokens. The monkey is allowed to make the decision as soon as he feels sufficiently confident, and has 500 ms to bring the cursor into a target after leaving the center. When the monkey reaches a target, the remaining tokens move more quickly to their final targets ("post-decision interval," which was either 150 ms in "slow" blocks or 50 ms in "fast" blocks. In a few sessions, the post-decision interval was reduced to 20 ms in fast blocks). Once all tokens have jumped, visual feedback is provided to the monkey (the chosen target turns green for correct choices or red for error trials) and a drop of water or fruit juice is delivered for choosing the correct target. A 1500 ms inter-trial interval precedes the following trial. We alternated between slow and fast blocks for about 75-125 trials each, typically several times in each recording session.

The monkeys were also trained to perform a delayed reach (DR) task (usually 30-48 trials per recording session). In this task, the monkey again begins by placing the cursor in the central circle containing the 15 tokens. Next, one of six peripheral targets is presented (1.75 cm radius, spaced at 60° intervals around a 5cm radius circle) and after a variable delay (500 ± 100 ms), the 15 tokens simultaneously jump into that target. This "GO signal" instructs the monkey to move the handle to the target to receive a drop of juice. This task is used to determine cell's task response and tuning as well as the animal's mean reaction time (RT), used as an estimate of the total delays attributable to sensory processing and response initiation (see Figure 1B).

In a subset of sessions, we trained monkeys to perform "never rewarded" trials, which were intermixed with regular trials in the tokens task. This manipulation was performed to control for the possibility that some of the effects described in the present report were due to the animals' expectation of reward. In the never rewarded trials, the targets were orange (as opposed to blue in regular trials) and monkeys never received the reward, even if they made the correct decision. Monkeys thus associated the orange circles with the "never rewarded" condition, so that they knew at trial start that they would not receive any reward regardless of their choice. Nevertheless, they still had to complete it to move on to the next, regular trial. Because monkeys typically did not react well to this condition, we had to keep the number of never rewarded trials very low (typically ~5% of trials) during sessions in which the manipulation was performed.

Dataset

The last stage of monkeys' training in the tokens task involved providing animals with alternating blocks of slow and fast trials of the tokens task. Based on behavioral data (Thura et al., 2014), we defined two periods during this last stage: first, when behavior was comparable between the two blocks; and second, when the monkeys began to behave differently in the two blocks, in terms of decision duration and success probability.

All neurophysiological data reported here were acquired from correct or error trials in which the monkeys completed the tokens task by choosing one of the two targets. Neurons were selected according to their anatomical location and physiological properties (see Figure S1).

In analyses aimed to explore the effect of sensory evidence as well as the time course of action selection in cortical areas and globus pallidus, we focused on trials from the slow block. The dataset for these analyses consists of 202 M1 neurons (monkey S, n = 78), 275 PMd (monkey S, n = 175), 51 GPe (monkey S, n = 19) and 56 GPi (monkey S, n = 22) neurons. 33 PMd cells were recorded during the first period of the last stage of Monkey S's training.

In analyses aimed to explore the neural correlates of monkeys' SAT adjustment between the blocks, we only considered globus pallidus neurons. The effect of block on PMd and M1 cells is the subject of a separate, recently published paper (Thura and Cisek, 2016). Here, all globus pallidus cells were recorded while monkeys behaved significantly differently in the two SAT conditions, and only cells recorded in both blocks were included in these analyses (GPe, n = 42, monkey S, n = 15; GPi, n = 52, monkey S, n = 20).

Neural recordings in PMd and M1

The standard procedures for single-unit recordings in the PMd and M1, signal processing and data management have been described previously (Thura and Cisek, 2014). During recording sessions, we focused on cells showing a change of activity in the tokens task, and monkeys were usually performing the task while we were searching for cells. When one or more task related cells were isolated, we ran a block of 30-48 trials of the DR task to determine spatial tuning and select a preferred target (PT) for each cell (i.e., the target associated with the highest firing rate during one or more task epochs). Next, we ran blocks of tokens task trials using the PT of an isolated cell and the 180° opposite target (OT). We sometimes simultaneously recorded several task-related cells showing different spatial preferences, and since we always selected a single pair of targets, the actual best direction for each of the recorded cells was not always among these two.

We usually started recording cells in the slow block because monkeys were more conservative in this condition. It was thus easier to assess cell properties online and more convenient to search for cells because fewer rewards were spent. When possible, cells were tested with multiple repetitions of slow and fast blocks to control for potential confounds related to evolving signals, elapsing time, and the monkey's fatigue or satiation (see Thura and Cisek, 2016 for control analyses on this question).

Neural recording in the globus pallidus

To record cells in the globus pallidus, we first estimated the position of the structure relative to the recording chambers based on MRI scans and 3D reconstructions using the Brainsight system (see Figure S1). Our goal was to reach GP regions where neurons show reach-related activity (Turner and Anderson, 1997) and belong to cortico-striato-pallidal loops that connect with M1 and PMd (i.e., middle of the GPi nucleus in the rostro-caudal axis; ventral, ventrolateral parts of the nucleus in the dorso-lateral axis (Middleton and Strick, 2000)). In both animals, penetrations began by crossing the ventral part of premotor cortex (PMv). Next, to guide our electrode to the globus pallidus, we followed the procedure described in DeLong's original experiment (DeLong, 1971). After passing beyond PMv, penetrations moved successively through about 3-8 mm of white matter (depending on the cortex thickness), the putamen (4-6 mm), the GPe (typically 2 mm) and finally the GPi (Figures S1B and S1C). The boundary between the putamen and GPe was straightforward to identify, as putamen cells usually exhibited low and sparse firing rates whereas GPe cells were usually spontaneously active with high firing rates and brief pauses characteristic of the GPe. As the electrode was lowered further, we usually observed an absence of activity for about 0.5 to 1.5 mm. Then, background activity increased again and high firing cells were observed, without pauses in discharge, indicating the beginning of GPi. On average, GPi cells were recorded about 2.5 mm deeper that GPe cells (Figures S1B and S1C), and as expected (DeLong, 1971), the average baseline firing rate was higher in GPi compared to GPe (81 Hz versus 64 Hz, Wilcoxon rank-sum test, $p = 0.009$).

When the globus pallidus (either GPe or GPi) was reached and a cell isolated, its activity in relation with the tokens task (increase or decrease of activity during any period of the task) was assessed online and only task-related cells were selected and further investigated. Once selected, the same procedure as the one described for PMd and M1 was applied to record spikes in the tokens task as well as in the DR task. Unlike during cortical recording sessions, only a single electrode was used at a time to record in GP.

QUANTIFICATION AND STATISTICAL ANALYSIS

All data and statistical analyses were performed using built-in and custom MATLAB (MathWorks, Natick, MA) functions and scripts. Details regarding the statistical tests used can be found in the following section, the main text and/or figure legends. Except when indicated, the significance level of all statistical tests was set at 0.05. Lower p values are reported when appropriate.

Behavioral data

Methods to analyze monkeys' behavior in the tokens task have been described previously (Cisek et al., 2009; Thura and Cisek, 2014). Briefly, the tokens task allows us to calculate, at each moment in time, the "success probability" associated with choosing each target. To characterize the success probability profile for each trial, we calculated this quantity (with respect to the target ultimately chosen by the monkey) for each token jump (Figure 1B). For example, with a total of 15 tokens, if at a particular moment in time the right target contains N_R tokens, the left contains N_L tokens, and N_C tokens remain in the center, then the probability that the target on the right will ultimately be the correct one (i.e., the success probability of guessing right) is:

$$p(R \mid N_R, N_L, N_C) = \frac{N_C!}{2^{N_C}} \sum_{k=0}^{\min(N_C, 7-N_L)} \frac{1}{k!(N_C - k)!} \quad \text{(Equation 1)}$$

Although each token jump and each trial was completely random, we could classify *a posteriori* some specific classes of trials embedded in the fully random sequence (e.g., "easy," "ambiguous," or "misleading" trials, Figure 1C). A trial is classified as "easy" if success probability (SP) exceeds 0.6 after two token jumps, 0.75 after five token jumps and 0.75 after eight token jumps. A trial is ambiguous if the SP is 0.5 after two jumps, between 0.4 and 0.65 after three token jumps, and then between 0.35 and 0.65 after five jumps. A trial is misleading if SP is below 0.4 after three token jumps.

Reaction time in the tokens task was calculated as the time of movement onset (based on kinematics) relative to the time of the first token jump. Decision time (DT) was estimated by subtracting from the reaction time the monkey's mean reaction time from the delayed reach task performed on the same day. We could then compute for each trial the duration of a decision as well as its success probability at the time of the decision (SP, Figure 1B). Wilcoxon rank-sum tests were used to compare reaction time, decision duration or success probability distributions between different conditions (Figure S2).

To quantify sensory evidence, we calculated a simple "first order" approximation of the success probability as the sum of log-likelihood ratios (SumLogLR) of individual token movements:

$$SumLogLR(n) = \sum_{k=1}^{n} \log \frac{p(e_k \mid S)}{p(e_k \mid U)} \quad \text{(Equation 2)}$$

where $p(e_k|S)$ is the likelihood of a token event e_k (a token jumping into either the selected or unselected target) during trials in which the selected target S is correct, and $p(e_k|U)$ is its likelihood during trials in which the unselected target U is correct. The SumLogLR is proportional to the difference in the number of tokens which have moved in each direction prior to the moment of decision (Cisek et al., 2009). Here, we thus defined "sensory evidence" as the information, pertinent to the correct choice, which is continuously present within the stimulus – i.e., the number of tokens in each target.

All arm and eye movement data were analyzed offline. Reaching characteristics (onset, offset, duration, amplitude, velocity peak) were assessed using monkeys' movement kinematics. Horizontal and vertical position data were first differentiated to obtain a velocity profile and then filtered using a 6th-order low-pass filter with a frequency cutoff of 15 Hz. Onset and offset of movements (used to calculate movement duration and amplitude) were determined using a 3 cm/s velocity threshold. Peak velocity was determined as the maximum value between these two events.

An analysis of covariance (ANCOVA) was used to assess the significance of the effects of block and decision duration (as well as any potential interactions) on sensory evidence at time of commitment as well as on movement parameters (Figure S2).

Neural data analysis

To be included in the analyses related to the time course of deliberation, neurons had to be recorded in the slow block and during at least 5 trials of each of the special trial types (easy, ambiguous, and misleading). To be included in the analyses related to the SAT adjustments between blocks, GP cells had to be recorded in at least 50 trials of the slow and the fast blocks (in one or more repetitions of each block).

We first defined a multiple linear regression model to assess the impact of a token jump as well as other factors on the neural response of all PMd, M1, GPe and GPi neurons recorded in the slow block of trials in the tokens task (Figure 2A). Normalized responses (relative to a 400 ms baseline period) aligned on the first token jump were measured during the deliberation epoch, in 200 ms bins after each of the 6 first token jumps and following a 200 ms delay allowing for sensory conductance (see Figure 2B). Spikes occurring after commitment time minus 200 ms are discarded to prevent any contamination of activity by action preparation processes. We used the following regression model to fit neural responses:

$$FR = I + \beta_1 * J + \beta_2 * V + \beta_3 * T + \beta_4 * E \quad \text{(Equation 3)}$$

where FR represents the discharge rate in a given epoch, I is the baseline activity, J is the direction of an individual token jump (either 0 or 1 depending on whether the token jumped into one target or the other), V is the vigor of the reaching movement, and T is the target chosen by the monkey (either 0 or 1 depending on whether the monkey chose one target or the other). Furthermore, inspired by a computational model of the basal ganglia proposed by Ditterich (2010) and based on a Bayesian framework for interpreting the role of the BG during decision-making (Bogacz and Gurney, 2007; Bogacz et al., 2016), we investigated whether activity in GPe and GPi is dependent on whether a given token jumps into the target that is currently expected (has the most tokens). To test this, we added an additional factor, E, defined as

$$E_{j+1} = N_j(T) - N_j(O) \quad \text{(Equation 4)}$$

where N(T) represents the number of tokens in the target into which token *j+1* jumped, and N(O) is the number of tokens in the other target.

In the model, the fitted value for the different coefficients provides an estimate of the change in response depending on the corresponding factor. Thus, this regression model provides a first-order test of the null hypothesis that a factor does not affect the collected responses. For instance, the null hypothesis is that jump direction does not affect a neuron activity (H_0: $\beta_1 = 0$). All fits were obtained using weighted least-squares and activity is normalized to account for differences of overall firing rates of neurons between cortical and sub-cortical areas.

For all cells recorded in PMd, M1, GPe and GPi, we next investigated those showing a significant spatial preference for one of the two potential targets before the execution of the reaching movement. To this aim, we calculated for each cell the mean activity related to each target in bins of 30 ms from 1500 ms before to 500 ms after reaching onset, and assessed the significance using a running receiver-operating characteristic (ROC) (Green and Swets, 1966; Shadlen et al., 1996) analysis with a criterion of 0.65. In a recent study (Thura and Cisek, 2014), we showed that some of the PMd cells reflect the deliberation process by tracking the evolution of sensory evidence and then signal the commitment to the choice about 280 ms before movement execution. Here, we classified a cell as tuned if the ROC scored above 0.65 during a period extending from 200 ms before this estimate of decision commitment until the movement onset (i.e., during a 480ms period prior to movement onset).

In this report, we also assessed the relationship of GP activity with movement execution during the delayed reach (DR) task (44/51 GPe and 51/56 GPi cells tested). We first compared the peri-movement activity (from 300 before to 200 ms after movement onset) (Turner and Anderson, 1997), for each cell and for each of the 6 potential targets, with activity during the late delay period (200 ms before the GO signal) in the same trial. Movement-related activity (either increasing, decreasing, or both increasing for particular targets and decreasing for others) was assessed with a Wilcoxon rank-sum test.

Next, we assessed the directional preference of GP cells during the delayed reach (DR) task (Table 1). We compared the peri-movement activity (from 300 before to 200 ms after movement onset) of each of these cells for the six movement directions and computed the significance of directional tuning with a non-parametric bootstrap procedure (1000 shuffles, $p < 0.05$, see Cisek et al., 2003).

The effect of movement vigor, movement amplitude, and peak velocity on GP cell activity was assessed with trial-by-trial Pearson correlation analyses in which movement vigor, amplitude or peak velocity and neural firing rate (from 300ms before movement onset to 200 ms after movement onset) were compared, separately for slow and fast trials in the tokens task (Table 1). If a cell was

directionally tuned during movement execution, only PT trials were included in the analysis. If not, trials for both directions were included. We defined vigor as the peak movement speed divided by the movement amplitude.

Cell discrimination times between targets (vertical bars in Figures 3A and 3D) were calculated as times when the difference in activity between compared conditions (PT- versus OT-related activity) exceeded 2 standard deviations in a sliding window (size, 10 ms; step, 2 ms) beginning at the first token jump (Sato and Schall, 2003).

For the tuned cells (ROC > 0.65), we investigated the relationship between sensory evidence (calculated as SumLogLR, see Equation 2) and normalized neural activity related to the cell's PT as a function of decision duration (in bins of 200 ms, corresponding to the token jumps) in the slow block of trials (Figure 3E). Normalization of each cell's activity was done by dividing the firing rate with its averaged baseline activity computed in a 400 ms period preceding the first token jump in slow block trials regardless of the final choice of the monkey. Spikes occurring after decision commitment were excluded and Spearman's correlation coefficient was calculated to assess the robustness of this relationship for each token jump for which at least 4 SumLogLR values were available (with a minimum of 5 samples for each value). From this analysis, we then focused on the evolution of the slope characterizing the relationship between sensory evidence and normalized neural activity. To compute the slope, we focused on data surrounding the 0 evidence point. For odd token jump numbers, the two closest data values located apart from the 0 evidence point were considered. For even jump numbers, three measures were considered (SumlogLR = 0 and the two closest data values apart from SumLogLR = 0). An analysis of covariance (ANCOVA) was used to assess the significance of the effect of neurons' recording location and decision duration on the steepness of the slopes.

The measure of the across-trial neural variability was performed using the Fano factor (Figure S4), defined as the spike-count variance divided by the spike-count mean, with counts computed in a 75ms sliding window, using the code employed in Churchland et al. (2010) and made available by the authors (http://churchlandlab.neuroscience.columbia.edu/links.html). In the present paper, we assessed the variability of the tuned PMd, M1, GPe and GPi cells during two epochs, either during early deliberation (from 300 ms before to 800 ms after the first token jump) or before and during movement execution (from 800 ms before to 300 ms after movement onset). For each population of cells, all trials in which decision duration was above 600 ms were included in the analysis.

The timing of peak activity in each area was calculated as cells' PT-related maximum firing rate relative to movement onset, computed in bins of 20 ms from 750 ms before movement onset to 350 ms after movement onset. Statistical significance of differences in peak activity timing across areas was assessed with a Wilcoxon rank-sum test (Figure 6E).

To assess the pattern of GP cells' activity variation during the deliberation period, we computed for each cell a trial-by-trial Pearson's correlation coefficient on activity binned every 5 ms from the first token jump until our estimate of decision commitment (280 ms before movement onset), regardless of the final choice of the monkey. If we observed more than 15% of trials with positive and significant correlation between activity and time compared to trials with negative and significant correlation between activity and time, we classified the cell as a ''build-up'' cell. If the cell showed the opposite result, it was classified as a ''decreasing'' cell. Cells that did not belong to either of the two above categories were classified as ''steady'' cells. This criterion delineated three categories of neurons in the globus pallidus, as shown in Figure 5.

The effects of block (slow versus fast) on GP cells were assessed via comparisons of averaged neural activity in two 200 ms epochs of the task: a baseline period, the 200 ms preceding the first token jump, and a deliberation period, from 400 to 600 ms after the first token jump. For the deliberation period, all decisions shorter than 600 ms were excluded from the analyses. In each individual cell, robustness of the effect of block was assessed with a Wilcoxon rank-sum test (Figures 4A and 4B).

We assessed the robustness of the effect of block on the GP population average response (Figures 5A–5F) with a bootstrap test, consisting in the resampling of the firing rates of each cell 1000 times in the two block conditions to produce distributions of means in each 30 ms bins from 500 ms before the first token jump until decision time. We then computed the 2.5%–97.5% percentiles of these distributions of resampled data to get the upper and lower bounds of the confidence interval (CI).

Instantaneous firing rate was assessed via a partial inter-spike interval method. When analyzing data with respect to the start of the trial (1st token jump), we always exclude all spikes occurring after our estimate of decision commitment (280ms before movement onset), i.e., any activity associated with movement initiation and/or execution. This is important in order to prevent averaging artifacts due to the very wide range of decision durations in the tokens task.

DATA AND SOFTWARE AVAILABILITY

Analysis-specific code and datasets are available by request to the Lead Contact: david.thura@umontreal.ca.