

COGNITION

Striatal stimulation enhances cognitive control and evidence processing in rodents and humans

Adriano E. Reimer, Evan M. Dastin-van Rijn, Jaejoong Kim, Megan E. Mensinger, Elizabeth M. Sachse, Aaron Wald, Eric Hoskins, Kartikeya Singh, Abigail Alpers, Dawson Cooper, Meng-Chen Lo, Amanda Ribeiro de Oliveira†, Gregory Simandl, Nathaniel Stephenson, Alik S. Widge*

Copyright © 2024 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

Brain disorders, in particular mental disorders, might be effectively treated by direct electrical brain stimulation, but clinical progress requires understanding of therapeutic mechanisms. Animal models have not helped, because there are no direct animal models of mental illness. Here, we propose a potential path past this roadblock, by leveraging a common ingredient of most mental disorders: impaired cognitive control. We previously showed that deep brain stimulation (DBS) improves cognitive control in humans. We now reverse translate that result using a set-shifting task in rats. DBS-like stimulation of the midstriatum improved reaction times without affecting accuracy, mirroring our human findings. Impulsivity, motivation, locomotor, and learning effects were ruled out through companion tasks and model-based analyses. To identify the specific cognitive processes affected, we applied reinforcement learning drift-diffusion modeling. This approach revealed that DBS-like stimulation enhanced evidence accumulation rates and lowered decision thresholds, improving domain-general cognitive control. Reanalysis of prior human data showed that the same mechanism applies in humans. This reverse/forward translational model could have near-term implications for clinical DBS practice and future trial design.

INTRODUCTION

Deep brain stimulation (DBS) uses surgically implanted electrodes to modulate target brain circuits. DBS is highly successful in movement disorders (1) and is being actively investigated for treatment of psychiatric disorders, including treatment-refractory obsessive-compulsive disorder (OCD) (2) and major depressive disorder (MDD) (3). In all of these disorders, DBS' mechanism of action remains unclear (1, 4). The lack of mechanistic understanding leads to difficulties in programming stimulation to individual patients' needs, which contributes to clinical trial failures (4). Human neurophysiology studies offer hope for better stimulator programming (4–6) but are similarly limited by the internal heterogeneity of psychiatric diagnoses (4, 7). In movement disorders, mechanistic understanding arose from animal models (8), and those models are testbeds for new treatment paradigms (9). In contrast, psychiatric syndromes are difficult to model in animals, limiting therapy development (10, 11).

Diagnostic heterogeneity might be overcome, and reliable animal models created, by modeling constructs that cut across and form the basic ingredients of psychiatric illnesses (7, 11, 12). Cognitive control is a particularly promising construct, because (i) it is highly relevant to multiple illnesses and (ii) it can be modified by DBS. Cognitive control is the ability to adjust and regulate thoughts and behavior in service of a specific goal (13). It involves the ability to suppress or override prepotent responses in favor of more adaptive responses. Cognitive control depends on a distributed brain network connecting prefrontal regions with the basal ganglia, striatum, and thalamus (14, 15). Impaired cognitive control is linked to

a wide range of psychiatric disorders (16), including MDD (17, 18), OCD (19, 20), and addictions (21). Conversely, improved cognitive control has been suggested to predict clinical response, mainly in MDD (17, 22).

Multiple studies have enhanced cognitive control with DBS-like stimulation of the human striatum and internal capsule, with correlation between that enhancement and improvement in mood/anxiety symptoms (23–26). Others have shown similar effects with DBS of the subthalamic nucleus (STN) (27), which may connect through the internal capsule to the prefrontal cortex (PFC) (28). Those studies were limited by windows of opportunity in clinical volunteers. They did not verify that the effects were repeatable or demonstrate mechanisms. The last point is critical, because cognitive control is often measured as response times (RTs) on a psychophysical task. An RT decrease might reflect improved cognitive control, but it might also represent a deleterious effect, such as impulsivity. If the task was sufficiently easy, participants might have been able to increase their speed without increased caution. Conversely, cognitive control has multiple subcomponents (13, 29). RT shortening might reflect improvement in motivation, attentional control to focus on task demands, precommitment to high-control behaviors (proactive allocation), adjustment to changing requirements (reactive allocation), or conflict processing (domain-general control). All are consistent with the subjective effects of striatum/capsule stimulation, including increased motivation (30), ability to refocus attention away from distress (23), ability to effortfully resist symptoms (31), and drive toward pleasure and social interaction (32, 33). Recipients also report less concern for negative outcomes (33, 34). Animal models could adjudicate between those mechanisms. Rodent models would be particularly useful, because rodents have a range of toolkits for manipulating specific neuronal subpopulations and circuits (35). Further, there is rodent-human homology in circuits believed to underpin cognitive

Department of Psychiatry & Behavioral Sciences, University of Minnesota, Minnesota, MN 55454, USA.

*Corresponding author. Email: awidge@umn.edu

†Present address: Department of Psychology, Federal University of São Carlos (UFSCar), São Carlos, Brazil.

control (36, 37). Related functions such as extinction learning are improved by stimulation of these homologous structures (38).

Here, we leveraged that homology to develop a reverse translational, mechanistic rodent model of psychiatric DBS. First, we show that human RT enhancements can be replicated in rodents by stimulating the midstriatum and embedded corticofugal fibers. This enhancement replicates a dorsoventral anatomic gradient observed in humans. We then applied this model to dissect DBS' cognitive effects. We explore the potential mechanisms above, and through multiple behavioral paradigms and computational modeling, show that DBS improves task performance by improving domain-general cognitive control. Physiologically, we link this improvement to the simultaneous modulation of multiple prefrontal regions, possibly through their corticothalamic projections. Closing the translational circle, we demonstrate that the same cognitive mechanism is present in the original human data, showing that the animal model has explanatory power.

RESULTS

Replicating DBS' effects on cognitive control in rodents

We probed cognitive control in rats using a set-shift task where rats shift between cue (light) and spatial (side) rules (Fig. 1, A to D; figs. S1 and S2, A to B; and table S1) (39, 40). Human cognitive control enhancement depends on stimulation of specific sites in the mid/dorsal striatum and internal capsule (23). We mapped the effects of DBS-like stimulation at homologous subregions of rat striatum ($n = 35$; Fig. 1E). This model replicated the human findings. Midstriatum stimulation significantly improved RTs ($\beta = -34$ ms and $P < 2.7 \times 10^{-17}$); other sites did not ($P > 0.05$; Fig. 1F and table S2). As in humans, there was no significant effect on response accuracy or any other outcome ($P > 0.05$; Fig. 1G; fig. S2, C and D; and tables S3 to S5). The RT improvement was consistently present across animals and days (fig. S2, E and F). Raw behavioral results are summarized in fig. S3 (A to D). RT improvement was not explained by changes in general activity ($P > 0.05$; fig. S4, A to C and E, and tables S6 to S10), motivation ($P > 0.05$; fig. S4, D and E, and tables S9 and S10) or lesion effects ($P > 0.05$ and table S2). It required the higher frequencies generally used for DBS. Twenty-hertz stimulation worsened the total errors ($\beta = 0.23$ and $P = 5.50 \times 10^{-4}$; fig. S5 and table S11) and RTs ($\beta = 25$ ms and $P = 2.96 \times 10^{-4}$; fig. S5 and table S12) in an independent cohort that also replicated the results of the initial 130-Hz study ($\beta = -48$ ms and $P = 1.54 \times 10^{-34}$; fig. S4 and table S12). Twenty-hertz stimulation had no effect on the probability of correct responses (fig. S5 and table S13) but did increase overall session duration ($\beta = 188.395$ and $P = 0.004$; fig. S5 and table S14). The midstriatal RT improvement was also present across sexes ($\beta = -17$ ms and $P = 0.001$), with a significantly larger effect size in females ($\beta = -80$ ms and $P = 5.03 \times 10^{-19}$) but no baseline sex differences ($P > 0.05$; fig. S6, A and B, and table S15). There were also no sex- or stimulation-related differences in errors ($P > 0.05$; fig. S6C and table S16).

RT improvements caused by midstriatal stimulation are not related to impulsivity

We tested for increased impulsivity or motivation using the five-choice serial reaction time task (5-CSRTT; $n = 6$; Fig. 1H, fig. S7, and tables S17 and S18) (41). There was no change in premature responses, the most common impulsivity metric ($P = 0.35$; Fig. 1I

and table S19). 5-CSRTT RTs increased ($\beta = 49$ ms and $P = 0.005$; Fig. 1I and table S20), which may reflect improved cognitive control in a situation where response withholding is essential. Stimulation also increased omissions ($\beta = 0.45$ and $P = 1.06 \times 10^{-8}$; Fig. 1I and table S21), arguing against an increase in general motivation. There was no change in accuracy ($P = 1.000$; Fig. 1I and table S22). Raw behavioral results for the 5-CSRTT are summarized in fig. S8 (A to D).

Midstriatal stimulation nonspecifically increased proactive control

Cognitive control involves both planned, proactive adaptive responses and reactive cancellation of maladaptive responses (13, 29). We examined whether midstriatal stimulation during set-shift affected either process. Rats frequently nose-poked during the intertrial interval (ITI), and we reasoned that these behaviors might reflect rehearsal or planning (proactive control). Supporting that concept, during the ITI, rats showed both early nose-pokes that were more common after incorrect responses, then a ramping pattern of late-ITI nose-pokes that were more common after correct responses (Fig. 2, A and B). After incorrect responses, ITI nose-pokes were commonly directed to the port that would have been correct ($\beta = 0.53$ and $P = 7.09 \times 10^{-18}$; Fig. 2A, left inset, and tables S23 and S24). After correct responses, rats initially showed reduced early-ITI poking ($\beta = -0.24$ and $P < 0.001$; table S23) but increased their activity in the late period ($\beta = 0.12$ and $P < 0.001$; table S23). Similarly, after correct responses, nose-pokes were more common in the port that was just rewarded both in the early ITI ($\beta = 8.27$ and $P < 0.001$; table S25) and late ITI ($\beta = 1.89$ and $P = 7.98 \times 10^{-302}$; Fig. 2B, right inset, and table S25). This pattern could reflect proactive control, because on side blocks, it is possible to plan the correct response before a trial begins. Thus, ITI pokes might reflect rehearsal to keep a desired response in working memory. Accordingly, rats were more likely to poke a side they had rehearsed during the immediately preceding ITI, but only if they ignored the cue light and chose the nonilluminated port ($\beta = 1.84$ and $P = 4.50 \times 10^{-129}$; Fig. 2C and table S26). Further, late-ITI nose-pokes in the middle (trial initiation) port predicted faster RTs on the subsequent trial ($\beta = -36$ ms and $P = 1.66 \times 10^{-15}$; Fig. 2D and table S27). Also consistent with a proactive control effect, RT improvement was augmented by rehearsal. If a rat poked both the middle and the port it would choose on the upcoming trial (a "choice" poke), the RT improvement was greater ($\beta = -12$ ms and $P = 0.021$; Fig. 2D and table S27), whereas nose-pokes into a port discordant with the upcoming trial behavior ("nonchoice" pokes) had no effect beyond that of the accompanying middle-port poke ($P = 0.445$; Fig. 2D and table S27). Proactive rehearsal did not, however, explain the RT effect of stimulation, because stimulation did not specifically increase it. Rather, midstriatal stimulation increased all forms of ITI pokes (all $P < 6.42 \times 10^{-3}$; Fig. 2E and table S28), including nonchoice pokes unrelated to cognitive control.

Stimulation improves cognitive control by improving decisional efficiency

The RT decrease could also reflect reactive adaptation or a general improvement in conflict resolution. To explore these, we fit a reinforcement learning–drift diffusion model (RLDDM) (42) to the behavior

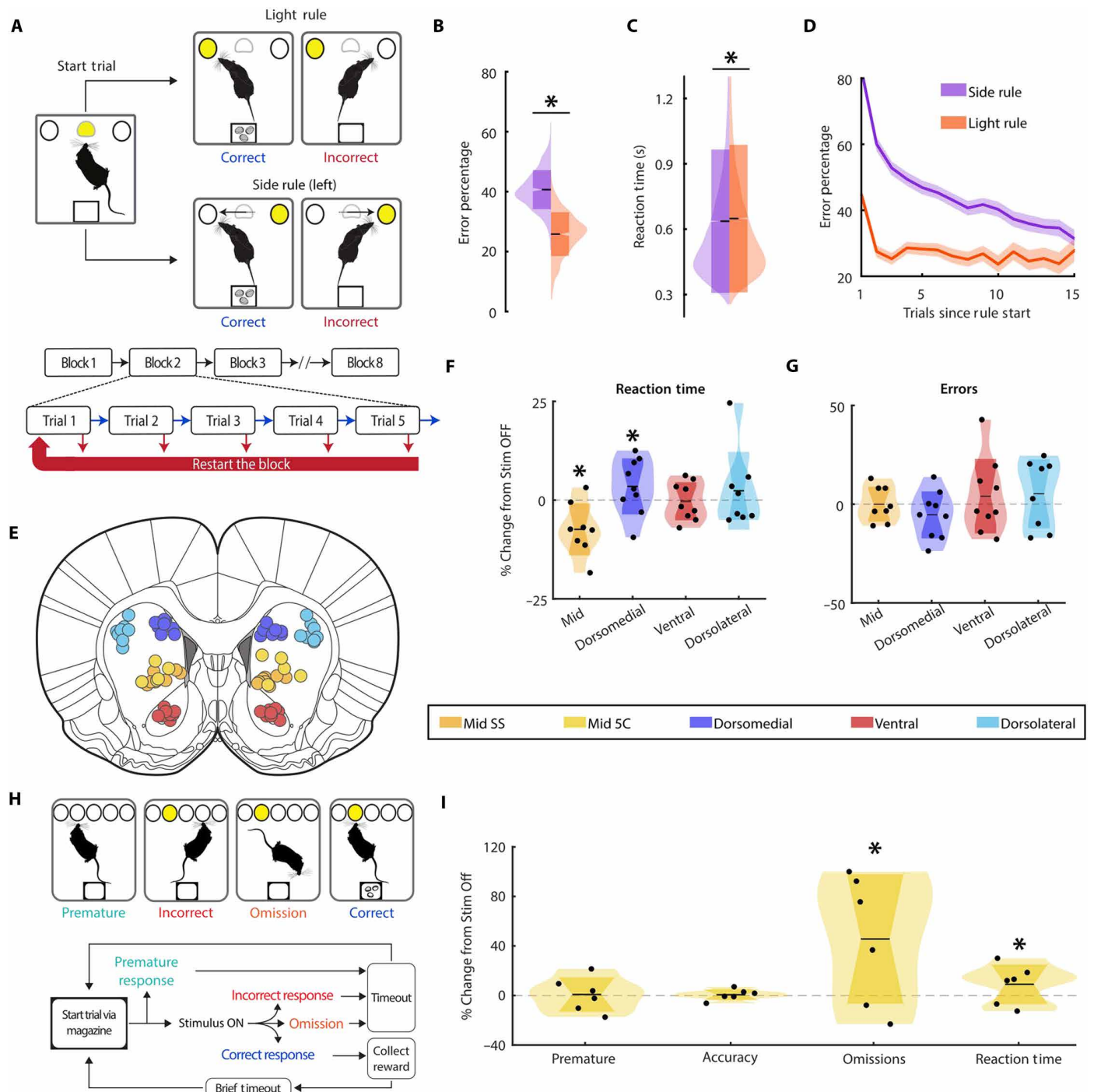


Fig. 1. Behavioral testing paradigms and stimulation outcomes. (A to D) Set-shift task. (A) Illustration of the set-shift task design. Rats must nosepoke to either an illuminated hole (light rule) or ignore the light and nosepoke on a specific side of the chamber (side rule). The rules are not cued and shift after the rat has completed five correct trials of the current rule. Errors reset to the beginning of the current five-trial block. (B) Probability of error plotted as a function of the current rule. (C) Distribution of reaction time (RT) separated by rule type. (D) Learning curves showing error probability across trials following each rule shift. (E to G) Stimulation experiment. (E) Histologically confirmed sites of implant ($n = 35$ rats total; midstriatum $n = 8$, ventral striatum $n = 9$, dorsomedial striatum $n = 9$, and dorsolateral striatum $n = 9$). The midstriatum target was the only target tested during the 5-CSRTT (see below, $n = 6$), whereas all four targets were tested for the primary set-shift experiment (SS). (F) Distributions of RT as a function of stimulation type, expressed as a percent change from stimulation-off. Distributions were computed over rats; each dot shows the median for one rat. (G) Distributions of percent change in error count during whole-session stimulation. (H and I) 5-CSRTT. (H) Task schematic. A trial began when a rat entered the food magazine. After a 5-s interval, a brief light stimulus was presented in one of five ports. If the rat poked the lit aperture, it received a food pellet (“correct response”). If the rat responded before the light was presented (“premature response”), in an incorrect aperture (“incorrect response”), or did not respond within the defined time period (“omission”), it received no reward. (I) Outcomes, following the plotting conventions of (F) and (G). $*P < 0.05$; all P values represent Wald Z tests of model parameters from GLMMs. See Materials and Methods for formulae and tables S1 to S3 (set-shift) and S19 to S22 (5-CSRTT) for statistical details.

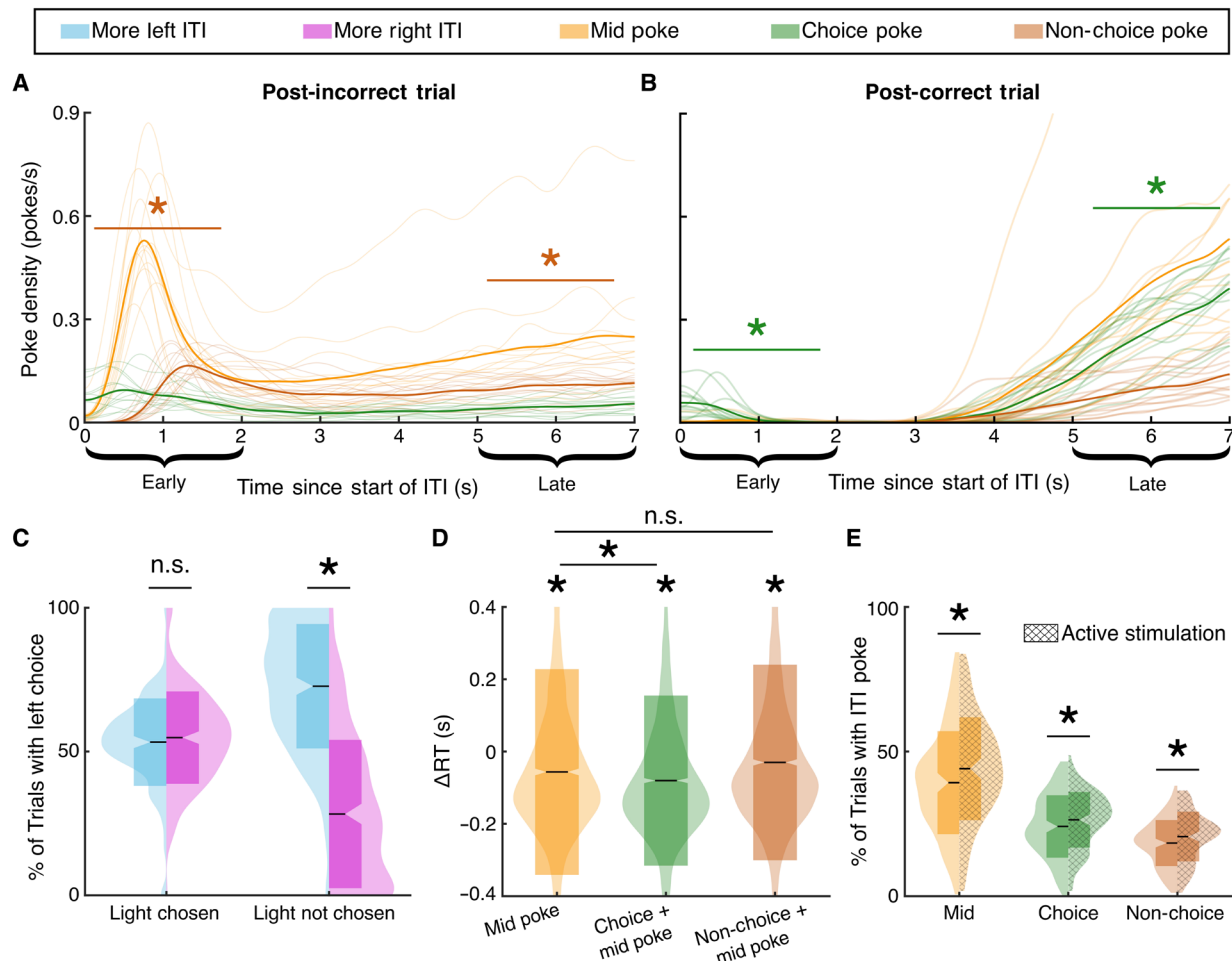
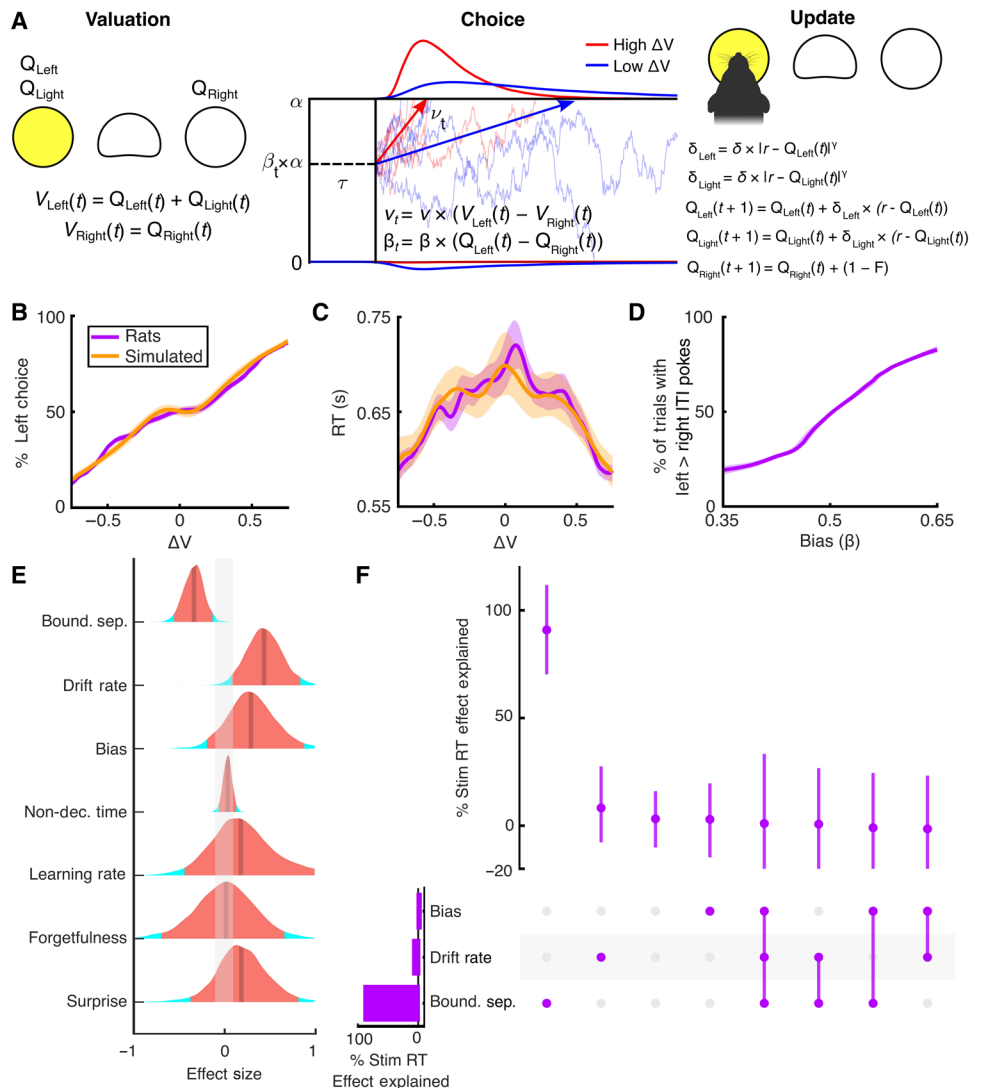


Fig. 2. Precommitment behaviors during ITIs of the set-shift task. (A and B) Probability of poking the mid (trial initiation) and side ports after incorrect (A) and correct responses (B). Thick lines show the mean over all animals and sessions, whereas light shaded lines show individual animals. Each plot is a kernel density estimate of the distribution, using a Gaussian kernel with 0.2-s SD. “Choice” and “nonchoice” pokes are pokes in the ports that were and were not chosen in the preceding trial, respectively. We emphasize “early” (first 2 s) and “late” (last 2 s) behavior. (C and D) Effects of ITI behavior on the subsequent trial. In these panels, choice and nonchoice now refer to the upcoming trial rather than the preceding trial as in (A) and (B). (C) Effect of ITI pokes on subsequent trial behavior (precommitment) during trials when the rat ignored the cue light (for example, during a side rule). Distributions show the probability of choosing the left or right port based on which port was most often poked during the Late portion of the preceding ITI. (D) Effect of ITI pokes in the late portion of the ITI on subsequent decision times. Plots show the distribution of RTs of the subsequent trial, conditioned on different types of ITI behavior. Asterisks above individual distributions indicate a mean different from the distribution when no ITI pokes occurred. (E) Distributions of late ITI poking into the mid and side ports with (hatched) and without midstriatal stimulation. Asterisk above each distribution reflects a significant difference between stimulation on and off. All *P* values represent Wald *Z* tests of model parameters from GLMMs; see Materials and Methods for formulae.

(Fig. 3A and fig. S9), incorporating both reactive reinforcement learning [RL; (29, 43)] and diffusion-based evidence accumulation [DDM; a model of conflict resolution (44)]. Like the traditional DDM, the RLDDM has trial-independent boundary separation and nondesicion time parameters that reflect the evidence required for a decision and the speed of sensorimotor processing, respectively. However, unlike the traditional DDM, the drift rate and bias terms in the RLDDM vary trial to trial as a function of the difference in expected rewards from the RL component. These two terms can capture efficacy of evidence accumulation and the ability to precommit to a high-probability option, respectively (Fig. 3A). The RL component included an adaptive learning rate and forgetting function based on a model selection analysis (fig. S10A). Simulated responses from the RLDDM strongly matched empirical behavior and psychometrics (Fig. 3, B to D, and fig. S10, B to E).

The ITI behaviors in Fig. 2 were captured in the RLDDM’s bias term, which incorporates value information specifically related to the side rule (Fig. 3D). Midstriatal stimulation did not alter any learning components of the model, arguing against a reactive control mechanism (Fig. 3E). Rather, it altered evidence accumulation components in a specific fashion that promotes efficient conflict resolution. Stimulation lowered the boundary separation term [probability of direction (*pd*) = 99.85%, median = -0.334 , and 1.05% in region of practical equivalence (ROPE)], which reflects the amount of evidence needed to reach a choice. This on its own would decrease RTs but also increase errors, because narrower boundaries are more easily reached by noise. However, stimulation also increased the drift rate (*pd* = 99.25%, median = 0.441, and 2.925% in ROPE), which drives the model toward the correct choice. When combined with tighter bounds, this leads to

Fig. 3. Midstriatal stimulation improves evidence processing and conflict resolution. (A) Illustration of the RLDDM. The RLDDM first assigns values V to the ports on the basis of its underlying estimates Q of the current value of the sides (left and right) and the light. The model then makes choices through a drift-diffusion process controlled by three key parameters: the decision threshold (boundary separation, α) which determines how much evidence is needed for a choice; the speed of evidence accumulation (drift rate, ν); and the time for nondecision processes (τ). The total value difference between options determines the drift rate, whereas the difference between sides determines the bias β . Trials with high value differences lead to shorter reaction times and more consistent choices (red versus blue curves above). After the port choice, values Q are updated on the basis of a RL process, with γ scaling the magnitude of errors and F controlling decay of unchosen option values. The model was fit hierarchically to the data from all rats, allowing individual-specific estimates of model parameters and their change with stimulation. (B to D) Posterior predictive simulations of behavioral outcomes for models (orange) and rat behavior (purple) over 4000 hierarchical posterior draws; see Materials and Methods. Shading shows the 95% highest density interval across animals and trials. (B) Choice as a function of the difference in value between the left and right sides (ΔV), where higher x -axis values represent a higher value for the left port. (C) RT as a function of the difference in value between the left and right sides, with the same conventions as (B). (D) Probability of an ITI poke as a function of the model's bias term. Higher x -axis values correspond to greater bias β toward the left port, and the y axis shows the percentage of left versus right pokes during the ITI. (E) Distributions of stimulation effect on model parameters over 4000 posterior draws. The median of the distribution is indicated by the solid line, and the 95% highest density interval is in orange. The gray shaded area around zero represents the ROPE for a null effect (effect size less than 0.1). (F) UpSet plot illustrating the percentage of the stimulation-related decrease in RT explained by stimulation effects on RLDDM model parameters. Each row corresponds to a parameter, with the bar indicating the total percentage of the stimulation-RT effect explained by that parameter. Each column corresponds to an intersection between parameters, with the filled-in dots indicating which parameters are part of the intersection. The upper panel indicates the percentage of the stimulation-RT effect that is unique to that intersection. The median and 95% highest density interval across posterior draws from the model are shown by the point and the bar, respectively.



more efficient conflict resolution, where the system needs less time to reliably reach the correct choice. There was a much smaller stimulation effect on the bias term ($pd = 88.525\%$, median = 0.296, and 15.425% in ROPE), which connects the RL side-specific valuation to the DDM. These three factors explained nearly 100% of the stimulation effect (Fig. 3F). This modeling implicates general conflict resolution, rather than specific strategies for control allocation, as the primary mechanism of cognitive control enhancement.

The RLDDM also explained the nonspecific increase in choice and nonchoice pokes described in Fig. 2E. Midstriatal stimulation only increased ITI poke behaviors at low bias levels, when rats had relatively little information about the correct answer to rehearse (fig. S11). This suggests that stimulation might reduce the gating threshold for converting internal preparation (bias) to an overt behavior (ITI poke).

Stimulation effects on cognitive control correlate with modulation of the PFC

Capsule/striatum stimulation is believed to involve retrograde modulation of the PFC (4, 28, 45), with mixed excitatory-inhibitory effects (24, 46). Consistent with this model, striatal stimulation increased *c-fos* expression throughout the PFC (fig. S12, A and B). The improvements modeled in Fig. 3 appeared to arise from prelimbic (PL) and infralimbic (IL) cortices, given that in these regions *c-fos* expression correlated moderately ($r > 0.24$) with RLDDM parameters. Boundary and drift rate changes were separable, with the former loaded more onto PL and the latter more onto IL cortices (fig. S12C).

Decisional efficacy changes translate across species

We finally tested whether these rodent findings explained the original human control improvements (23, 24). In those studies,

participants performed a multisource interference task (MSIT; Fig. 4A), where trials switched rapidly and unpredictably between two types. We thus fit DDMs without the RL component (Fig. 4B) to the participants from (23) who received the mid/dorsal striatal stimulation that most improved RT (fig. S13, A to D). Stimulation again increased the drift rate (Fig. 4C) ($pd = 0.978$, median = 1.31, and 1.47% in ROPE). This effect was present in four of five participants (Fig. 4D) and directly correlated with the RT improvement (Fig. 4E). Model selection argued against a change in the boundary separation (fig. S9B), as would be expected given the differing structure of the rodent and human tasks (see Materials and Methods and Discussion). We replicated these findings in an independent dataset from (24) (fig. S14, A to C; $pd = 0.996$, median = 0.62, and 1.03% in ROPE).

DISCUSSION

We developed a reverse translational model of psychiatric DBS, with homology to and explanatory power for effects previously shown in humans. This model highlights the power of a domain-focused approach. There are no strong animal models of psychiatric illness, because the core features of those illnesses are subjective self-reports (15, 16). Experts have repeatedly proposed to instead model cognitive and emotional deficits that serve as core “ingredients” of mental dysfunction (7, 47). We show the first working example of such a model. Our example uses cognitive control, but other domains of function such as risk-reward evaluation also change with DBS (48, 49), and the same approach likely applies.

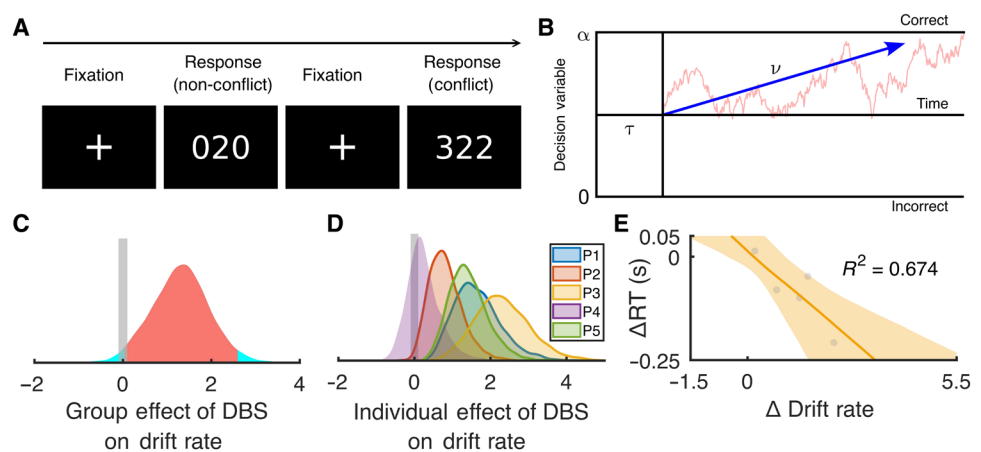
Our model could have clinical implications. Impulsivity has been proposed as a mechanism of DBS’ therapeutic effect, for example, by making a patient with OCD less sensitive to the “risk” of foregoing a ritual (34). Our findings suggest that impulsivity is dissociable from cognitive and likely clinical improvement. Similarly, our findings explain recent anatomic papers. Striatal DBS’ efficacy correlates with engagement of corticofugal fibers linking the PFC

to the STN (50, 51). That tract aligns with our midstriatal target. In primates, this part of internal capsule carries corticofugal fibers from the dorsal anterior cingulate (dACC), a structure associated with cognitive control (13, 14, 37). We showed increased *c-fos* in PL, a homolog of dACC that projects through the midstriatum (36). Primate cognitive control and DBS’ effects also load onto the lateral PFC (13, 14, 24), which has no known rodent homolog. Nevertheless, the alignment across other regions demonstrates the relevance of a cross-species model.

We used drift-diffusion modeling (DDM) to elucidate mechanisms of cognitive change, in part because DDM parameters have broad evidence for utility in describing psychiatric disorders (52). Elevated boundary separation (53, 54) and reduced drift rate (55, 56) are consistently seen separately and together (57, 58) for patients compared with controls. Consistent with a transdiagnostic construct, those findings span multiple disorders. Across disorders, DDM parameter changes predict elevated reaction times on decision-making tasks. This has been interpreted as “sluggish” but not less-accurate evidence accumulation/conflict resolution. Unexpectedly, very few studies have investigated how psychiatric treatments affect DDM parameters. Noninvasive stimulation of the PFC can reduce drift rate (59) or increase a bias toward effortful decisions (60) but these effects have only been investigated in healthy participants. Outside of cognitive control, STN DBS for Parkinson’s has been associated with impulsivity, reflected in a decrease in boundary separation (61, 62). Our finding of a stimulation-dependent, concurrent decrease in boundary separation and increase in drift rate has not previously been reported but represents a direct inversion of the typically observed deficits in these parameters for psychiatric patients (57, 58). Although a decrease in boundary separation alone would likely indicate impulsivity and result in more errors, the increased drift rate has a compensatory effect, leading to a net improvement in evidence processing (63, 64).

This and similar animal models could improve therapy development through new stimulation paradigms by enabling more rapid

Fig. 4. Changes in evidence processing identified in rats also explain previously reported stimulation-dependent changes in human cognitive control. (A) Schematic of the MSIT. In each trial, participants are shown three numbers and must identify the number that differs from the other two by pressing a corresponding button (1, 2, or 3). In nonconflict trials (e.g., “020”), the target number appears in its matching position (position 2 contains “2”). In conflict trials (e.g., “322”), the target number appears out of position (position 1 contains “2”) and is flanked by other numbers that could be valid targets, creating response interference. In (23, 24), we reported that striatal/internal capsule stimulation improves RT on this task, similar to the set-shift RT improvement reported in previous figures. (B) Model schematic.



We fit a standard DDM, exploring the potential effect of striatal stimulation on boundary separation α and drift rate ν . On the basis of Fig. 3’s results, we did not model potential effects on nondecision time τ . (C) Distributions of stimulation effect on drift rate over 4000 posterior draws. The median of the distribution is indicated by the solid line, and the 95% highest density interval is in orange. The gray shaded area around zero represents the ROPE for a null effect (effect size less than 0.1). (D) Individual distributions of drift rate changes for five participants, over 4000 posterior draws, on the same scale as (C). (E) Correlation between drift rate change and stimulation-induced RT change. Each dot represents one participant, specifically the RT change plotted against the point estimate of the drift rate change [median of the distributions shown in (D)]. The solid line shows a line of best fit. The shaded region represents the 95% highest density interval of this line. We calculated this by running a separate correlation for each posterior draw [distributions in (D)] against the change in median RT.

screening. For instance, our RT-based measures of cognitive control can be tracked in real time to identify parameters that optimally improve them (65), which may enable rapid screening of novel stimulation paradigms. Similar screening in humans would require weeks of expensive hospitalization (5, 6). Robust animal models could also improve patient selection. Because mental illnesses are internally heterogeneous (7, 11), neurostimulation of any given target will not work for most patients (4). Methods are emerging to identify patients whose symptoms arise from specific dysfunctions, including cognitive control (17). Pairing that identification with cognition-optimized neurostimulation could improve outcomes. Our mapping of DBS' effects to specific circuits could also yield physiologic biomarkers for closed-loop DBS. We (4) and others (66) have argued that those biomarkers will be difficult to find for ill-defined constructs such as "depression" but might well exist for robustly quantifiable metrics such as DDM parameters. Last, our results enable causal neuroscience. Nonhuman primates have strong internal capsule and prefrontal homology to humans (36). Striatal stimulation could manipulate subprocesses of cognitive control and uncover their neural substrates.

Striatal stimulation altered drift rate across species but altered boundary separation only in rats. This is due to the task structure (see Materials and Methods). Rats require multiple trials to learn each set-shift rule (Fig. 1). The human MSIT switched between high and low conflict (analogous to the set-shift rules) every one or two trials. Rats can precommit to a strategy (following the light or going to a specific side) before they initiate a trial and can learn this strategy during a block. Figures 1 to 3 show that both processes occur. Strategic anticipation decreases boundary separation, because less sensory evidence is needed to decide. Humans in MSIT can only process the sensory evidence once it is available, which loads onto the drift rate. Similarly, whereas rats frequently make errors, human performance on MSIT is almost always above 95% accuracy (23, 24). Without a high error rate, it is impossible to detect boundary separation changes, because that DDM term captures speed-accuracy tradeoffs. In a different task design, with long runs of high- or low-conflict trials, humans do show within-task learning (67), which enables better separation of the boundary separation and drift rate terms. Our rodent results predict that DBS effects on both terms should be dissociable by modulating internal capsule fibers descending from Brodmann area 24 (anterior cingulate, PL homolog, and boundary separation) versus Brodmann area 25 (subgenual cingulate, IL homolog, and drift rate) (36). This is not trivial, because the only way to identify such circuit-specific neurostimulation parameters would be through diffusion tractography, which is challenging in the internal capsule (68).

Neurobehavioral correlations highlighted the PL and IL cortices, but we did not see strong differences in c-fos evoked by stimulation at different striatal sites (fig. S8B), despite their different behavioral effects. Anatomy suggests that ventral stimulation should preferentially evoke c-fos expression in the mOFC (45), but mOFC c-fos was numerically higher for midstriatal and dorsolateral stimulation. This discordance is explained by our stimulation paradigm. First, our current (300 μ A) likely spreads well beyond the electrode tip, creating overlap between different targets' electric fields. Second, we stimulated for an hour before euthanizing for c-fos. Corticofugal fibers and striatal neurons form recurrent interconnected loops (69). Stimulation might have propagated through these connections to affect the cortex broadly. Third, 130-Hz stimulation has complex

excitatory/inhibitory effects (24, 46). C-fos only reflects excitation and thus may not fully reflect cortical engagement.

This study had specific limitations in the model species, task, randomization approach, and target. We used outbred rats, not a disease model. This could limit the translatability of results. However, as noted above, it is not clear that any putative "model of" psychiatric illness reflects human pathophysiology (10, 11). Rat models of transdiagnostic impairments, such as compulsivity, exhibit motor alterations that could confound our results (70, 71). Given that we (24) and others (26) have shown cognitive control improvements in psychiatrically diagnosed patients, we believe that the results would generalize to rodents or humans with cognitive impairments. Regarding the task, we used set-shifting to measure the broader construct of cognitive control. This is the most reported measure of cognitive control in rodents. Other assays of cognitive flexibility, such as reversal learning, have less need to suppress a prepotent response (72). This suppression of competing rules is critical; there was no RT effect of DBS in tasks without interference (24). There was similarly no RT effect of interference in a rodent flanker paradigm (73), hence why we did not use that task to model the human MSIT. Human internal capsule stimulation with a different cognitive control task, however, produced nearly identical RT improvement (26). Further, stimulation of comparable corticofugal fibers at a different target, the anteromedial STN, improved performance on a human extradimensional set-shift task that closely resembles our rat task (74). An important limitation is that we did not compare different stimulation sites within the same animal; we implanted separate cohorts with electrodes implanted in our four targets. Hence, each stimulation effect was measured against a slightly different baseline. The random intercepts in our models control for this, but slightly less robustly than our human work (23) that permitted direct within-participant comparison of different stimulation sites.

Another potential limitation is that our study differs from another recent attempt to model DBS using stimulation of the more posterior internal capsule in mice (46). In rats, the homologous fibers to the human anterior limb of the internal capsule are diffusely interpenetrated throughout the striatal gray matter rather than forming a cohesive tract. As a result, electrical stimulation of our optimal mid striatal target would have activated both passing corticofugal fibers and local striatal cell bodies. On the other hand, this is also true of DBS in humans; the relatively high currents used in most studies create tissue activation volumes that capture multiple aspects of striatum (75). We chose a more anterior target because the corticofugal fibers and synaptic terminals from the PFC have two useful properties when the implant is >1.2 mm anterior to bregma. First, because sensory-motor components synapse more posteriorly and PFC components synapse more anteriorly in the striatum (76), targeting an anterior plane more directly modulates the PFC, which is classically implicated in cognitive control. Second, because the striatum is overall larger in these more anterior planes, corticofugal fibers from different PFC subregions are more widely separated and can be separately activated (77). In contrast, in the rat posterior striatum or internal capsule, PFC projections are intermixed with motor/sensory fibers and more tightly packed together. This makes it much more difficult to map out the effects of different cortical sub-bundles. Thus, although there may be other ways to model DBS' effects on cognitive control, the overall results and prior literature make this anterior placement a particularly useful model. At the same time, we do note that we modulated both striatal neurons and corticofugal fibers; a more posterior placement in the capsule would

have permitted pure white matter stimulation, and comparing these might be valuable.

Whereas we describe above the translational implications of our findings/model, there are also important cautions. First, although cognitive control is impaired across many/most mental disorders (16, 21), control deficits may arise from multiple mechanisms. In that case, stimulation to improve cognitive control through the mechanism shown here may not treat all patients with cognitive control deficits. Newer work links control deficits to specific functional neuroimaging signatures (17, 18), suggesting that a relatively homogeneous population may be detected and treated, but this remains to be proven in clinical trials. Second, even in the presence of cognitive control impairments, human psychiatric disorders encompass a broad array of factors that cannot be captured in any one construct. Our model can likely identify biomarkers and inform targeted therapies for cognitive control impairments, and we have detailed how those biomarkers could be used to guide treatment (4, 65). It remains to be proven that modifying cognitive control leads to improved subjective well-being. There is evidence for this; our original studies reported subjective improvements in anxiety/distress (23), and neuroimaging studies suggest that cognitive control biomarkers mediate improvements in depression (17, 22). Prospective trials are still needed.

In summary, we show that a circuit-directed intervention and its cognitive mechanisms translate across species. These findings may enable a new set of psychiatric interventions. Knowing cognitive mechanisms should allow identification of patients who may benefit (specific deficits in cognitive control) and personalized intervention matched to individual biology (using cognitive change as a “read-out” of target engagement). In parallel, animal studies using this model could develop more reliable interventions, e.g., stimulation approaches that more reliably produce the desired cognitive change. Our findings similarly can inform the basic science of cognitive control and executive function by demonstrating that causal manipulations can affect specific subcomponents of those processes.

MATERIALS AND METHODS

Study design

We sought to develop an animal model of the core effect in our prior clinical work (23, 24): that 130-Hz stimulation of striatum and its passing corticofugal fibers leads to faster RTs in a task that requires cognitive control and that this improvement occurs on all trial types. We further sought to replicate the finding of (23) that the effect varies with the striatal region being stimulated, with the largest effects from stimulation slightly dorsal to the ventral striatum. We thus tested a set-shift task that, similar to human tasks, requires inhibition of a prepotent response (see below). Electrical stimulation parameters were chosen as homologs of those used in prior work (23–25), and statistical testing of the primary outcome similarly followed those papers. Partitioning of the striatum into stimulation zones (see below) was based on theoretical suggestions of a dorsal/ventral and lateral/medial functional segregation (15, 78), anatomic studies showing a specific topography of corticofugal fibers passing through different striatal subregions (77), and prior work suggesting cognitive benefits of stimulation in the mid-striatum (38). We performed secondary (unplanned) analyses through drift-diffusion and related models because those models are well suited to dissecting decision-making under response conflict (44, 52). Because the initial work used only male rats, we ran an additional mixed male/female cohort to check for sex differences, which was not

part of the initial study plan. All experiments were approved by the University of Minnesota Institutional Animal Care and Use Committee (protocols 1806-35990A and 2104-39021A) and complied with National Institutes of Health guidelines.

The primary set-shift experiment design and analysis were preplanned, with the four stimulation sites selected on the basis of prior theory/anatomic subdivisions of the rodent striatum as above and the analysis method (see below) selected to match the generalized linear models (GLMs) used in prior human studies (23, 24). Animal counts and number of sessions were predetermined by a power analysis. Because we report four behavioral outcomes from the same task (see below), we targeted 80% power at a Bonferroni-corrected $\alpha = 0.0125$. Because a GLM is functionally equivalent to a repeated-measures analysis of variance (ANOVA), we used that power estimator within G*Power 3.1. Assuming eight measurements per animal and four groups, the canonical medium effect size ($f = 0.25$) required 24 animals (seven per group). To provide a safety factor, we empirically increased this to eight animals per group and 16 measurements per animal. We then allocated and implanted at least 10 animals per primary group (fig. S1A) to ensure that we had sufficient power after animal exclusions for technical failures. Animals were assigned to one of the four implant targets without a predetermined order or formal randomization procedure. There was no blinding of experimenters or analysts, although analyses were pre-specified; see below.

We ran additional cohorts of the set-shift experiment, targeting only midstriatal stimulation on the basis of the primary findings. This aimed to validate the initial result and to test further controls, namely, different stimulation frequencies and sex differences. Because this study involved fewer comparisons, the effect direction was known, and because we had observed in the primary experiments that effect sizes were larger than anticipated, we empirically decreased this to planned samples of five rats per group. As described in the main text, these samples fully replicated the result of the main experiment.

The 5-CSRTT sample size was preplanned to be identical in size and protocol to the primary set-shift sample. We similarly planned to do this experiment only if the set-shift experiment identified a behaviorally effective stimulation site. The analysis strategy was preplanned to use GLMs with the same fixed/random effects as set-shift. Because of animals that did not learn the task or had off-target electrodes, we eventually implanted a total of 27 animals rather than the 10 initially planned (fig. S5).

The c-fos regions for analysis and neurobehavioral correlation strategy against set-shift and 5-CSRTT behavior changes were similarly preplanned. No power analysis was conducted for this component, because our emphasis was on adequate power for the behavioral effect, and we did not plan to perform inferential statistics. With the large number of planned assessment regions, behavioral outcomes, and stimulation sites, a well-powered study would require hundreds of animals.

Set-shifting task

The operant set-shift task (Fig. 1A) was modified from (39, 79) and described in detail in a previous study (80). In brief, animals learned and then shifted their responses between two distinct perceptual discrimination rules or dimensions: a cue-driven “light” rule and a spatial “side” rule. Rats were required to poke the illuminated middle nosepoke hole to initiate a trial. In all trials, one of the two peripheral

nosepoke holes was then illuminated. The light discrimination rule required the rats to poke the illuminated nosepoke hole, regardless of its spatial location. The side rule required that the animals poke at a designated spatial location across trials (left or right), regardless of which one was illuminated. The light rule is easier because it requires a simple response to a visually salient stimulus and thus creates a prepotent response tendency that must be overcome on side trials (see Fig. 1, B to D). Rats were reinforced with a single reward pellet for each correct response. After five consecutive correct choices, the rule was switched to the other dimension, requiring rats to shift their behavior to continue receiving rewards. The sequential trials on a single rule, regardless of correct/incorrect choice, were grouped into a block (Fig. 1A). No explicit cue was provided, besides the absence of expected reward, to signal rule changes. The task required the rats to reach the performance criterion eight times, resulting in seven shifts per test session.

Five-choice serial reaction time task

The 5-CSRTT (41) required rats to initiate a trial by nosepoking the illuminated food trough at the back of the chamber, then wait 5 s before a brief flash of light appeared in one of the five nosepokes on the opposite wall of the chamber. The flash of light appeared for 0.5 s (stimulus duration), and rats had an additional 5 s (limited hold) to nosepoke the aperture that lit up. Rats were reinforced with a single reward pellet for each correct response. A 5-s timeout before the next trial could be initiated occurred regardless of response type. If rats nosepoked too early, nosepoked the wrong aperture, or failed to respond within 5 s, the house light extinguished briefly to signify failure, rats were not rewarded, and there was a timeout as above.

Surgical approach

Prior work in humans suggests that striatum/internal capsule stimulation's effects on cognitive flexibility are spatially specific, with the greatest effects just dorsal to the ventral striatum/capsule (23). In rats, the homologous white matter bundles show a topographic distribution, such that different subregions of the rat striatum contain projections and corticofugal fibers from different PFC components (36, 77). Studies of extinction learning found that electrical stimulation enhanced extinction when delivered to a region just dorsal to the ventral striatum (38). Considering these prior results, we planned to test stimulation at four subtargets within the rat striatum (Fig. 1E): midstriatum (1.4 mm AP, ± 2.0 mm ML, and -6.0 mm DV), ventral striatum (1.3 mm AP, ± 2.0 mm ML, and -6.7 mm DV), dorsomedial striatum (1.4 mm AP, ± 1.8 mm ML, and -4.5 mm DV), or dorsolateral striatum (1.4 mm AP, ± 3.4 mm ML, and -4.5 mm DV). Each of these sites contains both cell bodies of the accumbens/caudate/putamen and corticofugal axons that project to both the thalamus and brainstem (37, 77). Rats were randomly assigned to each implantation group (fig. S1A).

DBS-like electrical stimulation

Electrical stimulation designed to model DBS was performed either using a PC running a custom-made LabVIEW program (National Instruments, Austin, TX) connected to a NI USB-6343 BNC analog/digital interface unit (National Instruments), connected in turn to an analog stimulus isolator (model 2200, A-M Systems, Sequim, WA), or using a PC running an Arduino script connected to a StimJim dual-channel electrical stimulator [(81); assembled by Labmaker, Berlin, Germany]. We verified that both setups produced equivalent

output current and waveforms. Parameters for both devices were set at 0.3 mA, 130 Hz, bipolar, biphasic, and charge-balanced square wave pulses with a pulse width of 50 μ s per phase, totaling 100 μ s per pulse. One hundred thirty hertz was the frequency used in our prior human experiments and is a typical human DBS frequency (2, 23, 24).

Statistical analysis

For analysis of primary behavioral and video data, we used generalized linear mixed models (GLMMs) to account for the repeated-measures design and non-Gaussian distribution of the data. Link functions and error distributions were chosen to be appropriate to each specific variable (for example, identity link and gamma distribution for RT). In cases where multiple behavioral measures were tested for the same purpose, critical *P* values were adjusted using a Bonferroni correction. A detailed accounting of all the variables and interactions used in each regression is included in Supplementary Methods.

RLDDMs (42) were fit to both rat and previously collected human behavioral data (23, 24) using Markov-Chain Monte Carlo with the HDDM Python package (82). Four independent chains were run for each model, and convergence was confirmed by testing whether the Gelman-Rubin statistic (83) was <1.1 . The predictive accuracy of the models was compared using deviance information criterion (DIC). The best model according to DIC was further evaluated using posterior predictive checks to assess its ability to account for key features of the data. To determine the presence of a stimulation effect on model parameters, we assessed the group level distributions for the stimulation-specific parameters through two Bayesian statistics: the pd and ROPE (84). pd refers to the proportion of the parameter distribution greater than 0 with a value above 0.5 (indicating a general increase in the parameter) and vice versa for values below 0.5. ROPE refers to the fraction of the density of the parameter distribution that lies within a region that is practically equivalent to a null effect. To define our ROPE, we normalized our parameter estimates by dividing the group estimate by the pooled SD parameters for the baseline parameter and stimulation-specific parameter. Using this normalized representation, we set our ROPE to be the region between ± 0.1 . pd is used as an indication for the existence/direction of an effect, whereas ROPE establishes significance. For the set-shift data, we additionally assessed the degree to which three model parameters (boundary separation, drift rate scaling, and bias scaling) could fully explain the observed decrease in reaction time with stimulation either uniquely or in common through a commonality analysis (85). Extensive details on the background for these models and their analysis are included in Supplementary Methods.

Supplementary Materials

The PDF file includes:

Supplementary Methods
Figs. S1 to S14
Tables S1 to S28
References (86–106)

Other Supplementary Material for this manuscript includes the following:

Data file S1
MDAR Reproducibility Checklist

REFERENCES AND NOTES

1. J. K. Krauss, N. Lipsman, T. Aziz, A. Boutet, P. Brown, J. W. Chang, B. Davidson, W. M. Grill, M. I. Hariz, A. Horn, M. Schulder, A. Mammis, P. A. Tass, J. Volkmann, A. M. Lozano,

- Technology of deep brain stimulation: Current status and future directions. *Nat. Rev. Neurol.* **17**, 75–87 (2021).
2. N. C. R. McLaughlin, D. D. Dougherty, E. Eskandar, H. Ward, K. D. Foote, D. A. Malone, A. Machado, W. Wong, M. Sedrak, W. Goodman, B. H. Kopell, F. Issa, D. C. Shields, O. A. Abulseoud, K. Lee, M. A. Frye, A. S. Widge, T. Deckersbach, M. S. Okun, D. Bowers, R. M. Bauer, D. Mason, C. S. Kubu, I. Bernstein, K. Lapidus, D. L. Rosenthal, R. L. Jenkins, C. Read, P. F. Malloy, S. Salloway, D. R. Strong, R. N. Jones, S. A. Rasmussen, B. D. Greenberg, Double blind randomized controlled trial of deep brain stimulation for obsessive-compulsive disorder: Clinical trial design. *Contemp. Clin. Trials Commun.* **22**, 100785 (2021).
 3. S. K. Conroy, P. E. Holtzheimer, Neuromodulation strategies for the treatment of depression. *Am. J. Psychiatry* **178**, 1082–1088 (2021).
 4. A. S. Widge, Closing the loop in psychiatric deep brain stimulation: Physiology, psychometrics, and plasticity. *Neuropsychopharmacology* **49**, 138–149 (2024).
 5. K. A. Scangos, A. N. Khambhati, P. M. Daly, G. S. Makhou, L. P. Sugrue, H. Zamanian, T. X. Liu, V. R. Rao, K. K. Sellers, H. E. Dawes, P. A. Starr, A. D. Krystal, E. F. Chang, Closed-loop neuromodulation in an individual with treatment-resistant depression. *Nat. Med.* **27**, 1696–1700 (2021).
 6. S. A. Sheth, K. R. Bijanki, B. Metzger, A. Allawala, V. Pirtle, J. A. Adkinson, J. Myers, R. K. Mathura, D. Oswalt, E. Tzolaki, J. Xiao, A. Noecker, A. M. Strutt, J. F. Cohn, C. C. McIntyre, S. J. Mathew, D. Borton, W. Goodman, N. Pouratian, Deep brain stimulation for depression informed by intracranial recordings. *Biol. Psychiatry* **92**, 246–251 (2022).
 7. B. N. Cuthbert, T. R. Insel, Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Med.* **11**, 126 (2013).
 8. J. L. Vitek, L. A. Johnson, Understanding Parkinson's disease and deep brain stimulation: Role of monkey models. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 26259–26265 (2019).
 9. B. Rosin, M. Slovik, R. Mitelman, M. Rivlin-Etzion, S. N. Haber, Z. Israel, E. Vaadia, H. Bergman, Closed-loop deep brain stimulation is superior in ameliorating parkinsonism. *Neuron* **72**, 370–384 (2011).
 10. L. M. Monteggia, H. Heimer, E. J. Nestler, Meeting report: Can we make animal models of human mental illness? *Biol. Psychiatry* **84**, 542–545 (2018).
 11. A. D. Redish, A. Kepecs, L. M. Anderson, O. L. Calvin, N. M. Grissom, A. F. Haynos, S. R. Heilbronner, A. B. Herman, S. Jacob, S. Ma, I. Vilares, S. Vinogradov, C. J. Walters, A. S. Widge, J. L. Zick, A. Zilverstand, Computational validity: Using computation to translate behaviours across species. *Philos. Trans. R. Soc. B* **377**, 20200525 (2022).
 12. Q. J. M. Huys, T. V. Maia, M. J. Frank, Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* **19**, 404–413 (2016).
 13. T. Egner, Principles of cognitive control over task focus and task switching. *Nat. Rev. Psychol.* **2**, 702–714 (2023).
 14. V. Menon, M. D'Esposito, The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology* **47**, 90–103 (2022).
 15. M. J. Sharpe, T. Stalnaker, N. W. Schuck, S. Killcross, G. Schoenbaum, Y. Niv, An integrated model of action selection: Distinct modes of cortical control of striatal decision making. *Ann. Rev. Psychol.* **70**, 53–76 (2019).
 16. L. M. McTeague, M. S. Goodkind, A. Etkin, Transdiagnostic impairment of cognitive control in mental illness. *J. Psychiatr. Res.* **83**, 37–46 (2016).
 17. L. M. Hack, L. Tozzi, S. Zenteno, A. M. Olmsted, R. Hilton, J. Jubeir, M. S. Korgaonkar, A. F. Schatzberg, J. A. Yesavage, R. O'Hara, L. M. Williams, A cognitive biotype of depression linking symptoms, behavior measures, neural circuits, and differential treatment outcomes: A prespecified secondary analysis of a randomized clinical trial. *JAMA Netw. Open* **6**, e2318411 (2023).
 18. L. Tozzi, X. Zhang, A. Pines, A. M. Olmsted, E. S. Zhai, E. T. Anene, M. Chesnut, B. Holt-Gosselin, S. Chang, P. C. Stetz, C. A. Ramirez, L. M. Hack, M. S. Korgaonkar, M. Wintermark, I. H. Gotlib, J. Ma, L. M. Williams, Personalized brain circuit scores identify clinically distinct biotypes in depression and anxiety. *Nat. Med.* **30**, 2076–2087 (2024).
 19. P. Gruner, C. Pittenger, Cognitive inflexibility in obsessive-compulsive disorder. *Neuroscience* **345**, 243–255 (2017).
 20. T. W. Robbins, M. M. Vaghi, P. Banca, Obsessive-compulsive disorder: Puzzles and prospects. *Neuron* **102**, 27–47 (2019).
 21. A. S. Widge, S. R. Heilbronner, B. Y. Hayden, Prefrontal cortex and cognitive control: New insights from human electrophysiology. *F1000Res* **8**, 1696 (2019).
 22. N. A. Crane, L. M. Jenkins, R. Bhaumik, C. Dion, J. R. Gowins, B. J. Mickey, J.-K. Zubieta, S. A. Langenecker, Multidimensional prediction of treatment response to antidepressants with cognitive control and functional MRI. *Brain* **140**, 472–486 (2017).
 23. I. Basu, A. Yousefi, B. Crocker, R. Zelmann, A. C. Paulk, N. Peled, K. K. Ellard, D. S. Weisholtz, G. R. Cosgrove, T. Deckersbach, U. T. Eden, E. N. Eskandar, D. D. Dougherty, S. S. Cash, A. S. Widge, Closed-loop enhancement and neural decoding of cognitive control in humans. *Nat. Biomed. Eng.* **7**, 576–588 (2023).
 24. A. S. Widge, S. Zorowitz, I. Basu, A. C. Paulk, S. S. Cash, E. N. Eskandar, T. Deckersbach, E. K. Miller, D. D. Dougherty, Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function. *Nat. Commun.* **10**, 1536 (2019).
 25. S. T. Olsen, I. Basu, M. T. Bilge, A. Kanabar, M. J. Boggess, A. P. Rockhill, A. K. Gosai, E. Hahn, N. Peled, M. Ennis, I. Shiff, K. Fairbank-Haynes, J. D. Salvi, C. Cusin, T. Deckersbach, Z. Williams, J. T. Baker, D. D. Dougherty, A. S. Widge, Case report of dual-site neurostimulation and chronic recording of cortico-striatal circuitry in a patient with treatment refractory obsessive compulsive disorder. *Front. Hum. Neurosci.* **14**, 569973 (2020).
 26. A. Allawala, S. Vartany, R. Mathura, H. Ritz, J. Adkinson, R. Gadot, D. Oswalt, A. Shenhav, W. Goodman, N. Pouratian, K. R. Bijanki, D. Borton, S. A. Sheth, DBS-induced improvement in cognitive control is mediated by theta oscillations in human intracranial recordings. *Neurosurgery* **69**, 36 (2023).
 27. R. Kelley, O. Flouty, E. B. Emmons, Y. Kim, J. Kingyon, J. R. Wessel, H. Oya, J. D. Greenlee, N. S. Narayanan, A human prefrontal-subthalamic circuit for cognitive control. *Brain* **141**, 205–216 (2018).
 28. J. C. Baldermann, C. Melzer, A. Zapf, S. Kohl, L. Timmermann, M. Tittgemeyer, D. Huys, V. Visser-Vandewalle, A. A. Kühn, A. Horn, J. Kuhn, Connectivity profile predictive of effective deep brain stimulation in obsessive-compulsive disorder. *Biol. Psychiatry* **85**, 735–743 (2019).
 29. T. S. Braver, The variable nature of cognitive control: A dual mechanisms framework. *Trends Cogn. Sci.* **16**, 106–113 (2012).
 30. L. Chen, N. Li, S. Ge, A. M. Lozano, D. J. Lee, C. Yang, L. Li, Q. Bai, H. Lu, J. Wang, X. Wang, J. Li, J. Jing, M. Su, L. Wei, X. Wang, G. Gao, Long-term results after deep brain stimulation of nucleus accumbens and the anterior limb of the internal capsule for preventing heroin relapse: An open-label pilot study. *Brain Stimul.* **12**, 175–183 (2019).
 31. I. Graat, S. Franken, G. van Rooijen, P. de Koning, N. Vulink, M. de Kroo, D. Denys, R. Mocking, Cognitive behavioral therapy in patients with deep brain stimulation for obsessive-compulsive disorder: A matched controlled study. *Psychol. Med.* **53**, 5861–5867 (2023).
 32. M. S. Okun, G. Mann, K. D. Foote, N. A. Shapira, D. Bowers, U. Springer, W. Knight, P. Martin, W. K. Goodman, Deep brain stimulation in the internal capsule and nucleus accumbens region: Responses observed during active and sham programming. *J. Neurol. Neurosurg. Psychiatry* **78**, 310–314 (2007).
 33. A. S. Widge, E. Licon, S. Zorowitz, A. Corse, A. R. Arulagasam, J. A. Camprodon, C. Cusin, E. N. Eskandar, T. Deckersbach, D. D. Dougherty, Predictors of hypomania during ventral capsule/ventral striatum deep brain stimulation. *J. Neuropsychiatry Clin. Neurosci.* **28**, 38–44 (2016).
 34. S. de Haan, E. Rietveld, M. Stokhof, D. Denys, Effects of deep brain stimulation on the lived experience of obsessive-compulsive disorder patients: In-depth interviews with 18 patients. *PLOS ONE* **10**, e0135524 (2015).
 35. J. Muir, J. Lopez, R. C. Bagot, Wiring the depressed brain: Optogenetic and chemogenetic circuit interrogation in animal models of depression. *Neuropsychopharmacology* **44**, 1013–1026 (2019).
 36. S. R. Heilbronner, J. Rodriguez-Romaguera, G. J. Quirk, H. J. Groenewegen, S. N. Haber, Circuit-based corticostriatal homologies between rat and primate. *Biol. Psychiatry* **80**, 509–521 (2016).
 37. D. N. Bullock, E. A. Hayday, M. D. Grier, W. Tang, F. Pestilli, S. R. Heilbronner, A taxonomy of the brain's white matter: Twenty-one major tracts for the 21st century. *Cereb. Cortex* **32**, 4524–4548 (2022).
 38. J. Rodriguez-Romaguera, F. H. M. Do Monte, G. J. Quirk, Deep brain stimulation of the ventral striatum enhances extinction of conditioned fear. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8764–8769 (2012).
 39. J. Park, J. Wood, C. Bondi, A. Del Arco, B. Moghaddam, Anxiety evokes hypofrontality and disrupts rule-relevant encoding by dorsomedial prefrontal cortex neurons. *J. Neurosci.* **36**, 3322–3335 (2016).
 40. D. A. Hamilton, J. L. Brigman, Behavioral flexibility in rats and mice: Contributions of distinct frontocortical regions. *Genes Brain Behav.* **14**, 4–21 (2015).
 41. A. Bari, J. W. Dalley, T. W. Robbins, The application of the 5-choice serial reaction time task for the assessment of visual attentional processes and impulse control in rats. *Nat. Protoc.* **3**, 759–767 (2008).
 42. M. L. Pedersen, M. J. Frank, G. Biele, The drift diffusion model as the choice rule in reinforcement learning. *Psychon. Bull. Rev.* **24**, 1234–1251 (2017).
 43. R. Rescorla, A. Wagner, A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement in *Classical Conditioning II: Current Research and Theory*, vol. 2 (Appleton-Century-Crofts, 1972).
 44. R. Ratcliff, G. McKoon, The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput.* **20**, 873–922 (2008).
 45. S. N. Haber, A. Yendiki, S. Jbadi, Four deep brain stimulation targets for obsessive-compulsive disorder: Are they different? *Biol. Psychiatry* **90**, 667–677 (2021).
 46. B. J. G. Van Den Boom, A. Elhazaz-Fernandez, P. A. Rasmussen, E. H. Van Beest, A. Parthasarathy, D. Denys, I. Willuhn, Unraveling the mechanisms of deep-brain stimulation of the internal capsule in a mouse model. *Nat. Commun.* **14**, 5385 (2023).
 47. J. A. Gordon, On being a circuit psychiatrist. *Nat. Neurosci.* **19**, 1385–1386 (2016).

48. M. Figeo, M. Vink, F. de Geus, N. Vulink, D. J. Veltman, H. Westenberg, D. Denys, Dysfunctional reward circuitry in obsessive-compulsive disorder. *Biol. Psychiatry* **69**, 867–874 (2011).
49. S. R. Patel, T. M. Herrington, S. A. Sheth, M. Mian, S. K. Bick, J. C. Yang, A. W. Flaherty, M. J. Frank, A. S. Widge, D. Dougherty, E. N. Eskandar, Intermittent subthalamic nucleus deep brain stimulation induces risk-averse behavior in human subjects. *eLife* **7**, e36460 (2018).
50. A. S. Widge, F. Zhang, A. Gosai, G. Papadimitrou, P. Wilson-Braun, M. Tsintou, S. Palanivelu, A. M. Noecker, C. C. McIntyre, L. O'Donnell, N. C. R. McLaughlin, B. D. Greenberg, N. Makris, D. D. Dougherty, Y. Rathi, Patient-specific connectomic models correlate with, but do not reliably predict, outcomes in deep brain stimulation for obsessive-compulsive disorder. *Neuropsychopharmacology* **47**, 965–972 (2022).
51. J. C. Baldermann, T. Schüller, S. Kohl, V. Voon, N. Li, B. Hollunder, M. Figeo, S. N. Haber, S. A. Sheth, P. E. Mosley, D. Huys, K. A. Johnson, C. Butson, L. Ackermans, T. B. van der Vlis, A. F. G. Leentjens, M. Barbe, V. Visser-Vandewalle, J. Kuhn, A. Horn, Connectomic deep brain stimulation for obsessive-compulsive disorder. *Biol. Psychiatry* **90**, 678–688 (2021).
52. C. N. White, R. Ratcliff, M. W. Vasey, G. McKoon, Using diffusion models to understand clinical disorders. *J. Math. Psychol.* **54**, 39–52 (2010).
53. M. Rouault, T. Seow, C. M. Gillan, S. M. Fleming, Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biol. Psychiatry* **84**, 443–451 (2018).
54. I. M. Berwian, J. G. Wenzel, A. G. E. Collins, E. Seifritz, K. E. Stephan, H. Walter, Q. J. M. Huys, Computational mechanisms of effort and reward decisions in patients with depression and their association with relapse after antidepressant discontinuation. *JAMA Psychiatry* **77**, 513–522 (2020).
55. D. G. Dillon, T. Wiecki, P. Pechtel, C. Webb, F. Goer, L. Murray, M. Trivedi, M. Fava, P. J. McGrath, M. Weissman, R. Parsey, B. Kurian, P. Adams, T. Carmody, S. Weyandt, K. Shores-Wilson, M. Touns, M. McInnis, M. A. Oquendo, C. Cusin, P. Deldin, G. Bruder, D. A. Pizzagalli, A computational analysis of flanker interference in depression. *Psychol. Med.* **45**, 2333–2344 (2015).
56. C. Shen, O. L. Calvin, E. Rawls, A. D. Redish, S. R. Sponheim, Clarifying cognitive control deficits in psychosis via drift diffusion modeling and attractor dynamics. *Schizophr. Bull.* **50**, 1357–1370 (2024).
57. V. M. Lawlor, C. A. Webb, T. V. Wiecki, M. J. Frank, M. Trivedi, D. A. Pizzagalli, D. G. Dillon, Dissecting the impact of depression on decision-making. *Psychol. Med.* **50**, 1613–1622 (2020).
58. A. Mandali, K. Weidacker, S.-G. Kim, V. Voon, The ease and sureness of a decision: Evidence accumulation of conflict and uncertainty. *Brain* **142**, 1471–1482 (2019).
59. M. G. Philiastides, R. Aukstulewicz, H. R. Heekeren, F. Blankenburg, Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Curr. Biol.* **21**, 980–983 (2011).
60. A. Soutschek, L. Nadporozhskaia, P. Christian, Brain stimulation over dorsomedial prefrontal cortex modulates effort-based decision making. *Cogn. Affect. Behav. Neurosci.* **22**, 1264–1274 (2022).
61. J. F. Cavanagh, T. V. Wiecki, M. X. Cohen, C. M. Figueroa, J. Samanta, S. J. Sherman, M. J. Frank, Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat. Neurosci.* **14**, 1462–1467 (2011).
62. G. J. Pagnier, W. F. Asaad, M. J. Frank, Double dissociation of dopamine and subthalamic nucleus stimulation on effortful cost/benefit decision making. *Curr. Biol.* **34**, 655–660.e3 (2024).
63. C. C. Liu, T. Watanabe, Accounting for speed–accuracy tradeoff in perceptual learning. *Vision Res.* **61**, 107–114 (2012).
64. G. Dutilh, J. Vandekerckhove, F. Tuerlinckx, E.-J. Wagenmakers, A diffusion model decomposition of the practice effect. *Psychon. Bull. Rev.* **16**, 1026–1036 (2009).
65. S. S. Nagrle, A. Yousefi, T. I. Netoff, A. S. Widge, In silico development and validation of Bayesian methods for optimizing deep brain stimulation to enhance cognitive control. *J. Neural Eng.* **20**, 036015 (2023).
66. L. L. Oganessian, M. M. Shanchei, Brain–computer interfaces for neuropsychiatric disorders. *Nat. Rev. Bioeng.* **2**, 653–670 (2024).
67. Z. Fu, D. Beam, J. M. Chung, C. M. Reed, A. N. Mamelak, R. Adolphs, U. Rutishauser, The geometry of domain-general performance monitoring in the human medial frontal cortex. *Science* **376**, eabm9922 (2022).
68. K. Sretavan, H. Braun, Z. Liu, D. Bullock, T. Palnitkar, R. Patriat, J. Chandrasekaran, S. Brenny, M. D. Johnson, A. S. Widge, N. Harel, S. R. Heilbronner, A reproducible pipeline for parcellation of the anterior limb of the internal capsule. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging*, S2451–9022(24)00196-4 (2024).
69. J. Cox, I. B. Witten, Striatal circuits for reward learning and decision-making. *Nat. Rev. Neurosci.* **20**, 482–494 (2019).
70. A. E. Reimer, A. R. De Oliveira, J. B. Diniz, M. Q. Hoexter, E. C. Miguel, M. R. Milad, M. L. Brandão, Fear extinction in an obsessive-compulsive disorder animal model: Influence of sex and estrous cycle. *Neuropharmacology* **131**, 104–115 (2018).
71. A. H. Alkhatib, A. Dvorkin-Gheva, H. Szechtman, Quinpirole and 8-OH-DPAT induce compulsive checking behavior in male rats by acting on different functional parts of an OCD neurocircuit. *Behav. Pharmacol.* **24**, 65–73 (2013).
72. L. Q. Uddin, Cognitive and behavioural flexibility: Neural mechanisms and clinical considerations. *Nat. Rev. Neurosci.* **22**, 167–179 (2021).
73. M. A. Robble, H. S. Schroder, B. D. Kangas, S. Nickels, M. Breiger, A. M. Iturra-Mena, S. Perlo, E. Cardenas, A. Der-Avakian, S. A. Barnes, S. Leutgeb, V. B. Risbrough, G. Vitaliano, J. Bergman, W. A. Carlezon, D. A. Pizzagalli, Concordant neurophysiological signatures of cognitive control in humans and rats. *Neuropsychopharmacology* **46**, 1252–1262 (2021).
74. H. Tyagi, A. M. Apergis-Schoute, H. Akram, T. Foltynie, P. Limousin, L. M. Drummond, N. A. Fineberg, K. Matthews, M. Jahanshahi, T. W. Robbins, B. J. Sahakian, L. Zrinzo, M. Hariz, E. M. Joyce, A randomized trial directly comparing ventral capsule and anteromedial subthalamic nucleus stimulation in obsessive-compulsive disorder: Clinical and imaging evidence for dissociable effects. *Biol. Psychiatry* **85**, 726–734 (2019).
75. D. Denys, I. Graat, R. Mocking, P. De Koning, N. Vulink, M. Figeo, P. Ooms, M. Mantione, P. Van Den Munckhof, R. Schuurman, Efficacy of deep brain stimulation of the ventral anterior limb of the internal capsule for refractory obsessive-compulsive disorder: A clinical cohort of 70 patients. *AJP* **177**, 265–271 (2020).
76. B. J. Hunnicutt, B. C. Jongbloets, W. T. Birdsong, K. J. Gertz, H. Zhong, T. Mao, A comprehensive excitatory input map of the striatum reveals novel functional organization. *eLife* **5**, e19103 (2016).
77. V. Coizet, S. R. Heilbronner, C. Carcenac, P. Mailly, J. F. Lehman, M. Savasta, O. David, J.-M. Deniau, H. J. Groenewegen, S. N. Haber, Organization of the internal limb of the internal capsule in the rat. *J. Neurosci.* **37**, 2539–2554 (2017).
78. B. J. Everitt, T. W. Robbins, From the ventral to the dorsal striatum: Devolving views of their roles in drug addiction. *Neurosci. Biobehav. Rev.* **37**, 1946–1954 (2013).
79. J. M. Darrah, M. R. Stefani, B. Moghaddam, Interaction of N-methyl-D-aspartate and group 5 metabotropic glutamate receptors on behavioral flexibility using a novel operant set-shift paradigm. *Behav. Pharmacol.* **19**, 225–234 (2008).
80. A. R. De Oliveira, A. E. Reimer, G. J. Simandl, S. S. Nagrle, A. S. Widge, Lost in translation: No effect of repeated optogenetic cortico-striatal stimulation on compulsivity in rats. *Transl. Psychiatry* **11**, 315 (2021).
81. N. Cermak, M. A. Wilson, J. Schiller, J. P. Newman, Stimjim: Open source hardware for precise electrical stimulation. bioRxiv [Preprint] (2019); <https://doi.org/10.1101/757716>.
82. T. Wiecki, I. Sofer, M. Frank, HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in python. *Front. Neuroinform.* **7**, 14 (2013).
83. A. Gelman, D. B. Rubin, Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–472 (1992).
84. D. Makowski, M. S. Ben-Shachar, S. H. A. Chen, D. Lüdtke, Indices of effect existence and significance in the Bayesian framework. *Front. Psychol.* **10**, 2767 (2019).
85. D. R. Seibold, R. D. McPhee, Commonality analysis: A method for decomposing explained variance in multiple regression analyses. *Human Commun. Res.* **5**, 355–365 (1979).
86. E. M. Dastin-van Rijn, E. Sachse, F. Iacobucci, M. Mensinger, A. S. Widge, OSCAR: An open-source controller for animal research. bioRxiv [Preprint] (2023); <https://doi.org/10.1101/2023.02.03.527033>.
87. E. M. Dastin-van Rijn, J. Nielsen, E. M. Sachse, C. Li, M. E. Mensinger, S. G. Simpson, M. C. Bucci, D. J. Titus, A. S. Widge, Pybehavior: A hardware agnostic, Python-based framework for controlling behavioral neuroscience experiments. *J. Open Source Softw.* **9**, 6515 (2024).
88. T. Sesia, V. Bulthuis, S. Tan, L. W. Lim, R. Vlamings, A. Blokland, H. W. M. Steinbusch, T. Sharp, V. Visser-Vandewalle, Y. Temel, Deep brain stimulation of the nucleus accumbens shell increases impulsive behavior and tissue levels of dopamine and serotonin. *Exp. Neurol.* **225**, 302–309 (2010).
89. G. Grassi, M. Figeo, P. Ooms, L. Righi, T. Nakamae, S. Pallanti, R. Schuurman, D. Denys, Impulsivity and decision-making in obsessive-compulsive disorder after effective deep brain stimulation or treatment as usual. *CNS Spectr.* **23**, 333–339 (2018).
90. T. M. Herrington, J. J. Cheng, E. N. Eskandar, Mechanisms of deep brain stimulation. *J. Neurophysiol.* **115**, 19–38 (2016).
91. L. di Biase, A. Fasano, Low-frequency deep brain stimulation for Parkinson's disease: Great expectation or false hope? *Mov. Disord.* **31**, 962–967 (2016).
92. T. Nath, A. Mathis, A. C. Chen, A. Patel, M. Bethge, M. W. Mathis, Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat. Protoc.* **14**, 2152–2176 (2019).
93. O. Sturman, L. Von Ziegler, C. Schläpfi, F. Akyol, M. Privitera, D. Slominski, C. Grimm, L. Thieren, V. Zerbi, B. Grewe, J. Bohacek, Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology* **45**, 1942–1952 (2020).
94. J. J. DiCarlo, J. W. Lane, S. S. Hsiao, K. O. Johnson, Marking microelectrode penetrations with fluorescent dyes. *J. Neurosci. Methods* **64**, 75–81 (1996).
95. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Elceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

96. I. Arganda-Carreras, V. Kaynig, C. Rueden, K. W. Eliceiri, J. Schindelin, A. Cardona, H. Sebastian Seung, Trainable Weka segmentation: A machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017).
97. G. Paxinos, C. Watson, *The Rat Brain in Stereotaxic Coordinates* (Elsevier, 6th ed., 2007).
98. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, 2021); <https://R-project.org/>.
99. RStudio Team, RStudio: Integrated development environment for R (RStudio, 2020); <http://rstudio.com/>.
100. M. L. Delignette-Muller, C. Dutang, fitdistrplus: An R package for fitting distributions. *J. Stat. Soft.* **64**, 1–34 (2015).
101. W. N. Venables, B. D. Ripley, Modern applied statistics with S, in *Statistics and Computing* (Springer, 4th ed., 2002).
102. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* **67**, 1–48 (2015).
103. Y. Niv, R. Daniel, A. Geana, S. J. Gershman, Y. C. Leong, A. Radulescu, R. C. Wilson, Reinforcement learning in multidimensional environments relies on attention mechanisms. *J. Neurosci.* **35**, 8145–8157 (2015).
104. T. Womelsdorf, M. R. Watson, P. Tiesinga, Learning at variable attentional load requires cooperation of working memory, meta-learning, and attention-augmented reinforcement learning. *J. Cogn. Neurosci.* **34**, 79–107 (2021).
105. A. Lex, N. Gehlenborg, H. Strobel, R. Vuilleumot, H. Pfister, UpSet: Visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).
106. G. Bush, L. M. Shin, The multi-source interference task: An fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network. *Nat. Protoc.* **1**, 308–313 (2006).

Acknowledgments

Funding: This work was supported by the OneMind Institute, the Minnesota Medical Discovery Team on Addictions, the MnDRIVE Brain Conditions initiative, the Tourette Association of America, and the US National Institutes of Health (R01NS120851 to A.S.W.

and R01MH124687 to A.S.W.). A.R.d.O. was supported by the São Paulo Research Foundation (FAPESP, Brazil, 2017/22473-9). EDvR was supported by a National Science Foundation Graduate Research Fellowship under award number 2237827. The opinions herein are fully those of the authors and do not represent the positions of any funding body. **Author contributions:** A.S.W. and A.E.R. provided the conceptual framework for the study. A.S.W., A.E.R., E.M.D.-v.R., M.E.M., E.M.S., and A.W. designed the experiments. A.E.R., A.W., E.M.S., D.C., M.-C.L., A.R.d.O., G.S., and N.S. performed the set-shifting experiments. M.E.M., A.W., and A.E.R. performed the 5-CSSRT experiment. E.M.D.-v.R., A.E.R., A.A., K.S., and E.H. analyzed the rat data. J.K. reanalyzed data from the prior human studies. A.E.R., E.M.D.-v.R., J.K., M.E.M., and A.S.W. wrote the manuscript. A.S.W. supervised the project and acquired funding. **Competing interests:** A.S.W. has consulted with Abbott on DBS and anonymously to investors interested in psychiatric indications through expert networks that prohibit him from revealing specific clients. None of those clients involves any ongoing relationship, financial, or otherwise. A.S.W. has received nonfinancial research support from Medtronic and Boston Scientific, companies that manufacture deep brain stimulators. This work is indirectly related to patent US11241188B2, “System and methods for monitoring and improving cognitive flexibility” and patent application US20240017069A1, “Systems and methods for measuring and altering brain activity related to flexible behavior,” both of which name A.S.W. as an inventor. **Data and materials availability:** All data are available in the manuscript or supplemental materials. Data files and analysis code have been deposited in Zenodo at <https://doi.org/10.5281/zenodo.13942551>. Data files and analysis code are also available through our laboratory GitHub at <https://github.com/tne-lab/SetShiftManuscript>.

Submitted 13 March 2024

Resubmitted 5 July 2024

Accepted 22 November 2024

Published 18 December 2024

10.1126/scitranslmed.adp1723