Title: Cross-species modeling and enhancement of cognitive control with striatal brain stimulation

Authors

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

Affiliations

All authors are affiliated with the Department of Psychiatry & Behavioral Sciences, University of Minnesota, Minnesota, MN, USA

* denotes equal contributions to the work

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

Correspondence: awidge@umn.edu

Keywords: executive function, neuromodulation, deep brain stimulation, cognitive conflict, cognitive control

Number of words in Abstract: 125

Number of Figures: 4

Number of Tables: 0

Number of References: 72 main text, 104 total

One Sentence Summary: Developing a reliable animal model of a human brain stimulation therapy reveals that this therapy works by enhancing the brain's ability to process conflicting pieces of evidence.

Abstract

Brain disorders, particularly 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. We show a 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, showing that DBS-like stimulation of the mid-striatum improves cognitive control in rats. Using this model, we identify a mechanism, improvement in domain-general cognitive control, and rule out competing hypotheses such as impulsivity. The rat findings explain prior human results and have immediate implications for clinical practice and future trial design.

Main Text:

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*), Tourette syndrome (*3*), and major depressive disorder (MDD) (*4*). In all these disorders, DBS' mechanism of action remains unclear (*1*, *5*, *6*). The lack of mechanistic understanding leads to difficulties in programming stimulation to individual patients' needs, which contributes to clinical trial failures (*5*, *6*). Human neurophysiology studies offer hope for better stimulator programming (*6–9*) but are similarly limited by the internal heterogeneity of psychiatric diagnoses (*6*, *10*, *11*). In movement disorders, mechanistic understanding arose from animal models (*12*), and those models are testbeds for new treatment paradigms (*13*, *14*). In contrast, psychiatric syndromes are difficult to model in animals, limiting therapy development (*15*, *16*).

Diagnostic heterogeneity might be overcome, and reliable animal models created, by modeling constructs that cut across and form the basic ingredients of psychiatric illnesses (*10*, *16*, *17*). Cognitive control is a particularly promising construct for modeling DBS, because (1) it is highly relevant to multiple illnesses and (2) it can be modified via DBS. Cognitive control is the ability to adjust and regulate thoughts and behavior in service of a specific goal (*18*, *19*). 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 (*20–22*). Impaired cognitive control is linked to a wide range of psychiatric disorders, including MDD (*23*), OCD (*24*, *25*), and addictions (*26*).

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 (*27–29*).

63 Those studies were limited by windows of opportunity in clinical volunteers. They did not verify that the effects
64 were repeatable or demonstrate mechanisms. The last point is critical, because the claimed cognitive control
65 enhancement was a decrease in response times (RTs) on a psychophysical task. An RT decrease might reflect
66 improved cognitive control, but might also represent a deleterious effect, such as impulsivity. If the task was
67 sufficiently easy, participants might have been able to increase their speed without increased caution.
68 Conversely, cognitive control has multiple subcomponents (*19*, *30*, *31*). The RT effect might reflect
69 improvement in motivation, attentional control to focus on task demands, pre-commitment to high-control
70 behaviors (proactive allocation), adjustment to changing requirements (reactive allocation), or conflict
71 processing (domain-general control). All are consistent with the subjective effects of striatum/capsule
72 stimulation, including increased: motivation (*32*, *33*), ability to refocus attention away from distress (*27*, *34*),
73 ability to effortfully resist symptoms (*35*), and drive towards pleasure and social interaction (*36*, *37*). Recipients
74 also report less concern for negative outcomes (*34*, *37*). Animal models could adjudicate between those
75 mechanisms. Rodent models would be particularly useful, as rodents have a wide range of toolkits for
76 manipulating specific neuronal subpopulations and circuits (*38*, *39*). Further, there is rodent-human homology
77 in circuits believed to underpin cognitive control (*40*, *41*). Related executive functions such as extinction
78 learning are improved by stimulation of these homologous structures (*42*, *43*).
79
80 We leverage that homology to develop a reverse translational rodent model of psychiatric DBS that reveals
81 underlying mechanisms. First, we show that human RT enhancements can be replicated in rodents by
82 stimulation of the mid-striatum and embedded corticofugal fibers. This enhancement replicates a dorso-ventral
83 anatomic gradient observed in humans. We then apply this model to dissect DBS' cognitive effects. We explore
84 the potential mechanisms above, and through multiple behavioral paradigms and computational modeling, show
85 that DBS improves task performance by increasing the efficiency of conflict adjudication (domain-general
86 cognitive control). Physiologically, we link this improvement to the simultaneous modulation of multiple
87 prefrontal regions, possibly via their cortico-thalamic projections. Closing the translational circle, we

demonstrate that the same cognitive mechanism is present in the original human data, i.e. the animal model has explanatory power for clinically relevant phenomena.

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 (Figure 1A-D, Table S1) (*44*, *45*). Human cognitive control enhancement depended on stimulating specific sites in the mid/dorsal striatum and internal capsule (*27*). We mapped the effects of DBS-like stimulation at homologous sub-regions of rat striatum (n=35, Figure 1E, Figure S1). This model replicated the human findings. Mid-striatum stimulation significantly improved RTs (β = -34 ms, $p < 2.7e-17$); other sites did not ($p > 0.05$, Figure 1F, Table S2). As in humans, there was no significant effect on response accuracy or any other outcome ($p > 0.05$, Figure 1G, Figure S2C-D, Tables S3-S5). The improvement was consistently present across animals and days (Figure S2E-F). It was not explained by changes in general activity ($p > 0.05$, Figure S3A-C, Table S6-S10), motivation ($p > 0.05$, Figure S3D-E, Tables S9-S10) or lesion effects ($p > 0.05$, Table S2). It required the higher frequencies generally used for DBS. 20 Hz stimulation worsened the total errors (β=0.23, p=5.50e-4, Figure S4, Table S11) and RTs (β = 25 ms, p=2.96e-4, Figure S4, Table S12), in an independent cohort that also replicated the results of the initial 130 Hz study (β = -48 ms, p=1.54e-34, Figure S4, Table S12).

RT improvements are not impulsivity

We tested for increased impulsivity or motivation using the 5-Choice Serial Reaction Time Task (5CSRTT, n=6, Figure 1E/H, Figure S5, Tables S15-S16) (*46*). There was no change in premature responses, the most common impulsivity metric (p=0.35, Figure 1I, Table S17). 5CSRTT RTs <u>increased</u> (β = 49 ms, p=0.005, Figure 1I, Table S18), which may reflect improved cognitive control in a situation where response withholding

is essential. Stimulation also increased omissions (β = 0.45, p=1.06e-8, Figure 1I, Table S19), arguing against an increase in general motivation.

Mid-striatal stimulation non-specifically increased proactive control

Cognitive control involves both planned, proactive adaptive responses and reactive cancellation of maladaptive responses (*19*, *31*). We examined whether mid-striatal stimulation affected either process. Rats frequently nose-poked during the inter-trial interval (ITI), and we reasoned that these might reflect rehearsal or planning (proactive control). Supporting that concept, during the ITI rats showed both early nosepokes that were more common after incorrect responses, then a ramping pattern of late-ITI nosepokes that were more common after correct responses (Figure 2A-B). After incorrect responses, ITI nosepokes suggested correction/regret – they were commonly directed to the port that would have been correct (β = 0.53, p = 7.09e-18, Figure 2A, left inset, Table S21-S22). Similarly, after correct responses, nosepokes were more common in the port that was just rewarded (β = 1.89, p = 7.98e-302, Figure 2B, right inset, Table S23). Both 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. Consistent with this thesis, 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 non-illuminated port (β = 1.84, p=4.50e-129, Figure 2C, Table S24). Further, late-ITI nosepokes in the middle (trial initiation) port predicted faster RTs on the subsequent trial (β = -36 ms, p=1.66e-15, Figure 2D, Table S25). Also consistent with a proactive control effect, this 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 (β = -12 ms, p=0.021, Figure 2D, Table S25), whereas nosepokes into a port discordant with the upcoming trial behavior ("Non-Choice" pokes) had no effect beyond that of the accompanying middle-port poke (p=0.4448, Figure 2D, Table S25). Proactive rehearsal did not, however, explain the RT effect of stimulation, because stimulation did not specifically increase it. Rather, mid-striatal

stimulation increased all forms of ITI pokes (all $p < 6.42e-3$, Figure 2E, Table S26), including Non-Choice 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/evidence. To explore these, we fit a hybrid model (*47*) to the behavior data (Figure 3A, Figure S6), incorporating both reactive reinforcement learning (RL, (*31*, *48*)) and diffusion-based evidence accumulation (DDM, a model of conflict resolution (*49*)). The RL component included an adaptive learning rate and forgetting function, based on a model selection analysis (Figure S7A). Simulated responses from the RLDDM strongly matched empirical behavior and psychometrics (Figure 3B-D, S7B-E). The ITI behaviors in Figure 2 were captured in the RLDDM's bias term, which incorporates value information specifically related to the Side rule (Figure 3D). Mid-striatal stimulation did not alter any learning components of the model, arguing against a reactive control mechanism (Figure 3E). Rather, it altered evidence accumulation components in a specific fashion that promotes efficient conflict resolution. Stimulation lowered the boundary separation term (pd=99.85%, Std.Median=-0.3335, 1.05% in 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%, Std.Median=-0.4406, 2.925% in ROPE), which drives the model towards the correct-choice boundary. When combined with tighter bounds, this leads to 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%, Std.Median=0.296, 15.425% in ROPE), which connects the RL component to the DDM. These three factors explained nearly 100% of the stimulation effect (Figure 3F). This modeling implicates the general process of conflict resolution, rather than specific strategies for cognitive control allocation, as the primary mechanism of cognitive control enhancement.

Stimulation effects on cognitive control correlate with modulation of prefrontal cortex.

Prominent theories argue that stimulation of the capsule/striatum works primarily by retrograde modulation of prefrontal cortex (PFC) (*50–52*), with mixed excitatory-inhibitory effects (*28*, *53*). Consistent with this model, striatal stimulation increased c-fos expression throughout PFC. (Figure S8A-B). The improvements modeled in Figure 3 appeared to arise from modulation of prelimbic (PL) and infralimbic (IL) cortices, as 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 modulation (Figure S8C).

Decisional efficacy changes translate across species

We finally tested whether these rodent findings explained the original human cognitive control improvements (*27*, *28*). In those studies, participants performed a Multi-Source Interference Task (MSIT, Figure 4A), where trials switched rapidly and unpredictably between two types. We thus fit DDMs without the RL component (Figure 4B) to the participants from (*27*) who received the mid/dorsal striatal stimulation that most improved RT (Figure S9). Stimulation again increased the drift rate (Figure 4C) (pd=0.978, Std.Median=1.31, 1.47% in ROPE), and this effect was present in 4/5 participants. Model selection argued against a change in the boundary separation (Figure S9B), as would be expected given the differing structure of the rodent and human tasks (see Methods, Discussion). We replicated these findings in an independent dataset from (*28*) (Figure S10; pd=0.996, Std.Median=0.62, 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 cross-diagnostic/domain-focused approach. There are no strong animal models of psychiatric illness, in part 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 (*10*, *54*, *55*). We show the first

working example of such a model, connecting a cognitive domain across species to a therapeutic intervention. Our example uses cognitive control, but other domains of function also change with DBS (*56–58*), and the same approach likely applies.

Our model has immediate clinical implications. For instance, impulsivity has been proposed as a mechanism of DBS' therapeutic effect (e.g., by making a patient with OCD less sensitive to the "risk" of foregoing a ritual) (*34*, *59*). We show that impulsivity is dissociable from cognitive improvement, i.e. is likely a side effect. Similarly, our findings explain recent anatomic papers. Striatal DBS' efficacy correlates with engagement of corticofugal fibers linking PFC to STN (*51*, *52*). That tract aligns with our mid-striatal target. In primates, this part of internal capsule carries corticofugal fibers from dorsal anterior cingulate (dACC), a structure heavily associated with cognitive control (*19*, *20*, *22*, *41*, *60*). We showed increased c-fos in PL, a potential homologue of dACC that projects through mid-striatum (*40*, *61*). Primate cognitive control and DBS' effects also load onto lateral PFC (*19*, *20*, *28*), which has no known rodent homologue. Nevertheless, the alignment across other regions demonstrates the relevance of a cross-species model.

This and similar animal models should improve therapy development through new stimulation paradigms, by enabling more rapid screening. For instance, our RT-based measures of cognitive control can be tracked in real time to identify parameters that optimally improve them (*62*), which may enable rapid screening of novel paradigms such as temporally irregular (*14*) or plasticity-directed (*63*) stimulation. Similar screening in humans would require weeks of expensive hospitalization (*7*, *8*). Robust animal models could also improve DBS patient selection. Because mental illnesses are internally heterogeneous (*10*, *10*, *16*, *55*), neurostimulation of any given target will not work for most patients (*6*, *11*). Methods are emerging to identify patients whose symptoms arise from specific dysfunctions, including cognitive control (*64*). Pairing that identification with cognition-optimized neurostimulation could dramatically improve outcomes. Finally, our results enable causal neuroscience. Non-human primates have strong internal capsule and prefrontal homology to humans (*40*, *60*).

Striatal stimulation could manipulate specific sub-processes of cognitive control and uncover their neural substrates.

Striatal stimulation altered drift rate across species, but boundary separation only in rats. This is due to the task structure and its effect on error rates (see Methods for elaboration). Rats require multiple trials to learn each Set-Shift rule (Figure 1). The human MSIT switched between high and low conflict (analogous to the Set Shift rules) every 1-2 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 1-3 show that both processes occur. Strategic anticipation decreases the boundary separation, since 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, while rats frequently make errors, human performance on MSIT is almost always above 95% accuracy (*27*, *28*, *65*). 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, e.g. with long runs of high or low conflict trials, humans would likely show learning and DBS would likely alter a boundary separation term.

Neuro-behavioral correlations highlighted PL and IL, but we did not see strong differences in c-fos evoked by stimulation at different striatal sites (Figure S8B), despite their different behavioral effects. Anatomy suggests that ventral stimulation should preferentially evoke c-fos expression in mOFC (*50*, *60*, *61*), but mOFC c-fos was numerically higher for mid-striatal and dorsolateral stimulation. This discordance is explained by our stimulation paradigm. First, our relatively high current (300 μA) likely spreads well beyond the electrode tip, creating overlap between different targets' electric fields. Second, we stimulated for an hour before sacrificing for c-fos. Corticofugal fibers and striatal neurons form recurrent interconnected loops (*60*, *66*). Stimulation might have propagated through these connections to impact cortex more broadly. Third, 130 Hz stimulation has complex excitatory/inhibitory effects (*9*, *28*, *53*). C-fos only reflects excitation, and thus may not fully reflect cortical engagement.

This study had specific limitations in the model species and task. We studied only male rats, obscuring any potential sex differences. However, given that no sex differences were observed in the corresponding human studies (*27*, *28*), we did not anticipate significant sex-related differences. We also used outbred rats, not a disease model. This could limit the translatability of results to humans with psychiatric illnesses. However, as noted above, it is not clear that any putative "model of" psychiatric illness reflects human pathophysiology (*15*, *16*). Rat models of transdiagnostic impairments, such as compulsivity, exhibit motor alterations that could confound our results (*67*, *68*).

Regarding the task, we used set shifting to measure of the broader construct of cognitive control. This is the most commonly reported measure of cognitive control in rodents. Other assays of cognitive flexibility, such as reversal learning, have less need to suppress a prepotent response (*69–71*). That suppression of competing rules/interference is critical –in the original paper, there was no RT effect of DBS in tasks without interference (*28*). Other tasks, such as recently proposed rodent flanker paradigms (*72*), may be useful for measuring the generalizability of our effects.

In summary, we provide the first complete example of mapping a circuit-directed intervention and its cognitive mechanisms across species, demonstrating the value of a transdiagnostic approach to animal models of psychiatric illnesses and interventions. Our findings open the door to clinical interventions guided both by rigorous phenotyping and objective markers of target engagement, paired with animal studies to develop novel therapeutic strategies. They similarly can inform the basic science of cognitive control and executive function, by demonstrating that causal manipulations can affect specific subcomponents of those processes. Collectively, these impacts bring us closer to robust and personalized therapies for severe psychiatric illnesses.

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. *Contemporary Clinical Trials Communications* **22**, 100785 (2021).

3. D. Martinez-Ramirez, J. Jimenez-Shahed, J. F. Leckman, M. Porta, D. Servello, F.-G. Meng, J. Kuhn, D. Huys, J. C. Baldermann, T. Foltynie, M. I. Hariz, E. M. Joyce, L. Zrinzo, Z. Kefalopoulou, P. Silburn, T. Coyne, A. Y. Mogilner, M. H. Pourfar, S. M. Khandhar, M. Auyeung, J. L. Ostrem, V. Visser-Vandewalle, M.-L. Welter, L. Mallet, C. Karachi, J. L. Houeto, B. T. Klassen, L. Ackermans, T. Kaido, Y. Temel, R. E. Gross, H. C. Walker, A. M. Lozano, B. L. Walter, Z. Mari, W. S. Anderson, B. K. Changizi, E. Moro, S. E. Zauber, L. E. Schrock, J.-G. Zhang, W. Hu, K. Rizer, E. H. Monari, K. D. Foote, I. A. Malaty, W. Deeb, A. Gunduz, M. S. Okun, Efficacy and safety of deep brain stimulation in Tourette syndrome: the international Tourette syndrome deep brain stimulation public database and registry. *JAMA Neurol* **75**, 353–359 (2018).

4. S. K. Conroy, P. E. Holtzheimer, Neuromodulation Strategies for the Treatment of Depression. *AJP* **178**, 1082–1088 (2021).

5. F. L. Hitti, A. S. Widge, P. Riva-Posse, D. A. Malone, M. S. Okun, M. M. Shanechi, K. D. Foote, S. H. Lisanby, E. Ankudowich, S. Chivukula, E. F. Chang, A. Gunduz, C. Hamani, A. Feinsinger, C. S. Kubu, W. Chiong, J. A. Chandler, R. Carbunaru, B. Cheeran, R. S. Raike, R. A. Davis, C. H. Halpern, N. Vanegas-Arroyave, D. Markovic, S. K. Bick, C. C. McIntyre, R. M. Richardson, D. D. Dougherty, B. H. Kopell, J. A. Sweet, W. K. Goodman, S. A. Sheth, N. Pouratian, Future directions in psychiatric neurosurgery: Proceedings of the 2022 American Society for Stereotactic and Functional Neurosurgery meeting on surgical neuromodulation for psychiatric disorders. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation* **16**, 867–878 (2023).

6. A. S. Widge, Closing the loop in psychiatric deep brain stimulation: physiology, psychometrics, and plasticity. *Neuropsychopharmacol.*, 1–12 (2023).

7. K. W. Scangos, A. N. Khambhati, P. M. Daly, G. S. Makhoul, 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).

8. S. A. Sheth, K. R. Bijanki, B. Metzger, A. Allawala, V. Pirtle, J. A. Adkinson, J. Myers, R. K. Mathura, D. Oswalt, E. Tsolaki, 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. *Biological Psychiatry* **92**, 246–251 (2022).

9. S. Alagapan, K. S. Choi, S. Heisig, P. Riva-Posse, A. Crowell, V. Tiruvadi, M. Obatusin, A. Veerakumar, A. C. Waters, R. E. Gross, S. Quinn, L. Denison, M. O'Shaughnessy, M. Connor, G. Canal, J. Cha, R.

Hershenberg, T. Nauvel, F. Isbaine, M. F. Afzal, M. Figee, B. H. Kopell, R. Butera, H. S. Mayberg, C. J. Rozell, Cingulate dynamics track depression recovery with deep brain stimulation. *Nature*, 1–9 (2023).

10. B. N. Cuthbert, T. R. Insel, Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med* **11**, 126 (2013).

11. C. R. P. Sullivan, S. Olsen, A. S. Widge, Deep brain stimulation for psychiatric disorders: From focal brain targets to cognitive networks. *NeuroImage* **225**, 117515 (2021).

12. J. L. Vitek, L. A. Johnson, Understanding Parkinson's disease and deep brain stimulation: Role of monkey models. *Proc Natl Acad Sci USA* **116**, 26259–26265 (2019).

13. 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).

14. J. Wang, S. Nebeck, A. Muralidharan, M. D. Johnson, J. L. Vitek, K. B. Baker, Coordinated reset deep brain stimulation of subthalamic nucleus produces long-lasting, dose-dependent motor improvements in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine non-human primate model of Parkinsonism. *Brain Stimulation* **9**, 609–617 (2016).

15. L. M. Monteggia, H. Heimer, E. J. Nestler, Meeting Report: Can We Make Animal Models of Human Mental Illness? *Biological Psychiatry* **84**, 542–545 (2018).

16. 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. *Phil. Trans. R. Soc. B* **377**, 20200525 (2022).

17. Q. J. M. Huys, T. V. Maia, M. J. Frank, Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience* **19**, nn.4238 (2016).

18. M. M. Botvinick, J. D. Cohen, The computational and neural basis of cognitive control: charted territory and new frontiers. *Cognitive Science* **38**, 1249–1285 (2014).

19. T. Egner, Principles of cognitive control over task focus and task switching. *Nat Rev Psychol* **2**, 702–714 (2023).

20. V. Menon, M. D'Esposito, The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacol.* **47**, 90–103 (2022).

21. 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. *Annual Review of Psychology* **70**, 1–24 (2018).

22. M. W. Cole, W. Schneider, The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage* **37**, 343–360 (2007).

23. I. Grahek, A. Shenhav, S. Musslick, R. M. Krebs, E. H. W. Koster, Motivation and cognitive control in depression. *Neuroscience & Biobehavioral Reviews* **102**, 371–381 (2019).

24. P. Gruner, C. Pittenger, Cognitive inflexibility in Obsessive-Compulsive Disorder. *Neuroscience* **345**, 243–255 (2017).

25. T. W. Robbins, M. M. Vaghi, P. Banca, Obsessive-compulsive disorder: puzzles and prospects. *Neuron* **102**, 27–47 (2019).

26. A. S. Widge, S. R. Heilbronner, B. Y. Hayden, Prefrontal cortex and cognitive control: new insights from human electrophysiology. *F1000Res* **8**, 1696 (2019).

27. 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).

28. 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).

29. S. 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. M. 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. *Frontiers in Human Neuroscience* **14**, 569973 (2020).

30. A. Shenhav, S. Musslick, F. Lieder, W. Kool, T. L. Griffiths, J. D. Cohen, M. M. Botvinick, Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience* **40**, 99–124 (2017).

31. T. S. Braver, The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences* **16**, 106–113 (2012).

32. C. J. Thomson, R. A. Segrave, P. B. Fitzgerald, K. E. Richardson, E. Racine, A. Carter, Personal and relational changes following deep brain stimulation for treatment-resistant depression: A prospective qualitative study with patients and caregivers. *PLoS ONE* **18**, e0284160 (2023).

33. 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 Stimulation* **12**, 175–183 (2019).

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. 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. *Psychological Medicine*, 1–7 (2022).

36. 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).

37. A. S. Widge, E. Licon, S. Zorowitz, A. Corse, A. R. Arulpragasam, J. A. Camprodon, C. Cusin, E. N. Eskandar, T. Deckersbach, D. D. Dougherty, Predictors of hypomania during ventral capsule/ventral striatum deep brain stimulation. *Journal of Neuropsychiatry and Clinical Neurosciences* **28**, 38–44 (2016).

38. J. Muir, J. Lopez, R. C. Bagot, Wiring the depressed brain: optogenetic and chemogenetic circuit interrogation in animal models of depression. *Neuropsychopharmacol.* **44**, 1013–1026 (2019).

39. A. R. Nectow, E. J. Nestler, Viral tools for neuroscience. *Nature Reviews Neuroscience* **21**, 669–681 (2020).

40. 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).

41. 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. *Cerebral Cortex* **32**, 4524–4548 (2022).

42. 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).

43. J. Rodriguez-Romaguera, F. H. Do-Monte, Y. Tanimura, G. J. Quirk, S. N. Haber, Enhancement of Fear Extinction with Deep Brain Stimulation: Evidence for Medial Orbitofrontal Involvement. *Neuropsychopharmacol* **40**, 1726–1733 (2015).

44. J. Park, J. Wood, C. Bondi, A. D. Arco, B. Moghaddam, Anxiety evokes hypofrontality and disrupts rule-relevant encoding by dorsomedial prefrontal cortex neurons. *J. Neurosci.* **36**, 3322–3335 (2016).

45. D. A. Hamilton, J. L. Brigman, Behavioral flexibility in rats and mice: contributions of distinct frontocortical regions. *Genes, Brain and Behavior* **14**, 4–21 (2015).

46. 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).

47. 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).

48. 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* (1972)vol. Vol. 2.

49. R. Ratcliff, G. McKoon, The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation* **20**, 873–922 (2008).

50. S. N. Haber, A. Yendiki, S. Jbabdi, Four Deep Brain Stimulation Targets for Obsessive-Compulsive Disorder: Are They Different? *Biological Psychiatry* **90**, 667–677 (2021).

51. J. C. Baldermann, T. Schüller, S. Kohl, V. Voon, N. Li, B. Hollunder, M. Figee, 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. *Biological Psychiatry* **90**, 678–688 (2021).

52. 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. *Neuropsychopharmacol.* **47**, 965–972 (2022).

53. 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).

54. J. A. Gordon, On being a circuit psychiatrist. *Nat Neurosci* **19**, 1385–1386 (2016).

55. A. S. Widge, K. K. Ellard, A. C. Paulk, I. Basu, A. Yousefi, S. Zorowitz, A. Gilmour, A. Afzal, T. Deckersbach, S. S. Cash, M. A. Kramer, U. T. Eden, D. D. Dougherty, E. N. Eskandar, Treating refractory mental illness with closed-loop brain stimulation: progress towards a patient-specific transdiagnostic approach. *Experimental Neurology* **287**, 361–472 (2017).

56. C. S. Inman, K. R. Bijanki, D. I. Bass, R. E. Gross, S. Hamann, J. T. Willie, Human amygdala stimulation effects on emotion physiology and emotional experience. *Neuropsychologia* **145**, 106722 (2020).

57. M. Figee, 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).

58. 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-aversive behavior in human subjects. *Elife* **7**, e36460 (2018).

59. A. S. Widge, D. D. Dougherty, "Managing patients with psychiatric disorders with deep brain stimulation" in *Deep Brain Stimulation Management*, W. J. Marks Jr., J. L. Ostrem, Eds. (Cambridge University Press, Cambridge : New York, ed. 3rd, 2022), pp. 198–214.

60. S. N. Haber, W. Tang, E. Y. Choi, A. Yendiki, H. Liu, S. Jbabdi, A. Versace, M. Phillips, Circuits, networks, and neuropsychiatric disease: transitioning from anatomy to imaging. *Biological Psychiatry* **87**, 318–327 (2020).

61. 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 anterior limb of the internal capsule in the rat. *The Journal of Neuroscience* **37**, 2539–2554 (2017).

62. S. S. Nagrale, 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).

63. M. Lo, R. Younk, A. S. Widge, Paired electrical pulse trains for controlling connectivity in emotion-related brain circuitry. *IEEE Trans. Neural Syst. Rehabil. Eng.* **20**, 2721–2730 (2020).

64. 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 Network Open* **6**, e2318411 (2023).

447 65. G. Bush, L. M. Shin, The Multi-Source Interference Task: an fMRI task that reliably activates the
448 cingulo-frontal-parietal cognitive/attention network. *Nat Protoc* **1**, 308–313 (2006).

449 66. J. Cox, I. B. Witten, Striatal circuits for reward learning and decision-making. *Nat Rev Neurosci* **20**, 482–
450 494 (2019).

451 67. A. E. Reimer, A. R. De Oliveira, J. B. Diniz, M. Q. Hoexter, E. C. Miguel, M. R. Milad, M. L. Brandão,
452 Fear extinction in an obsessive-compulsive disorder animal model: Influence of sex and estrous cycle.
453 *Neuropharmacology* **131**, 104–115 (2018).

454 68. A. H. Alkhatib, A. Dvorkin-Gheva, H. Szechtman, Quinpirole and 8-OH-DPAT induce compulsive
455 checking behavior in male rats by acting on different functional parts of an OCD neurocircuit.
456 *Behavioural Pharmacology* **24**, 65–73 (2013).

457 69. V. Boulougouris, J. W. Dalley, T. W. Robbins, Effects of orbitofrontal, infralimbic and prelimbic cortical
458 lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research* **179**, 219–228 (2007).

459 70. S. Ghods-Sharifi, D. M. Haluk, S. B. Floresco, Differential effects of inactivation of the orbitofrontal
460 cortex on strategy set-shifting and reversal learning. *Neurobiology of Learning and Memory* **89**, 567–573
461 (2008).

462 71. L. Q. Uddin, Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. *Nat*
463 *Rev Neurosci* **22**, 167–179 (2021).

464 72. M. A. Robble, H. S. Schroder, B. D. Kangas, S. Nickels, M. Breiger, A. M. Iturra-Mena, S. Perlo, E.
465 Cardenas, A. Der-Avakian, S. A. Barnes, S. Leutgeb, V. B. Risbrough, G. Vitaliano, J. Bergman, W. A.
466 Carlezon, D. A. Pizzagalli, Concordant neurophysiological signatures of cognitive control in humans and
467 rats. *Neuropsychopharmacol.* **46**, 1252–1262 (2021).

468 73. B. W. Balleine, The meaning of behavior: discriminating reflex and volition in the brain. *Neuron* **104**,
469 47–62 (2019).

470 74. B. J. Everitt, T. W. Robbins, From the ventral to the dorsal striatum: Devolving views of their roles in
471 drug addiction. *Neuroscience & Biobehavioral Reviews* **37**, 1946–1954 (2013).

472 75. C. N. White, M. Servant, G. D. Logan, Testing the validity of conflict drift-diffusion models for use in
473 estimating cognitive processes: A parameter-recovery study. *Psychon Bull Rev* **25**, 286–301 (2018).

474 76. J. M. Darrah, M. R. Stefani, B. Moghaddam, Interaction of N-methyl-D-aspartate and group 5
475 metabotropic glutamate receptors on behavioral flexibility using a novel operant set-shift paradigm.
476 *Behavioural Pharmacology* **19**, 225–234 (2008).

477 77. A. Ribeiro de Oliveira, A. E. Reimer, G. J. Simandl, S. S. Nagrale, A. S. Widge, Lost in translation: No
478 effect of repeated optogenetic cortico-striatal stimulation on compulsivity in rats. *Translational*
479 *Psychiatry* **11**, 315 (2021).

480 78. T. Sesia, V. Bulthuis, S. Tan, L. W. Lim, R. Vlamings, A. Blokland, H. W. M. Steinbusch, T. Sharp, V.
481 Visser-Vandewalle, Y. Temel, Deep brain stimulation of the nucleus accumbens shell increases
482 impulsive behavior and tissue levels of dopamine and serotonin. *Experimental Neurology* **225**, 302–309
483 (2010).

79. J. Luigjes, M. Mantione, W. Van Den Brink, P. R. Schuurman, P. Van Den Munckhof, D. Denys, Deep brain stimulation increases impulsivity in two patients with obsessive–compulsive disorder: *International Clinical Psychopharmacology*, 1 (2011).

80. G. Grassi, M. Figee, 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).

81. 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.

82. N. Cermak, M. A. Wilson, J. Schiller, J. P. Newman, "Stimjim: open source hardware for precise electrical stimulation" (preprint, Neuroscience, 2019); https://doi.org/10.1101/757716.

83. T. M. Herrington, J. J. Cheng, E. N. Eskandar, Mechanisms of deep brain stimulation. *J Neurophysiol* **115**, 19–38 (2016).

84. L. di Biase, A. Fasano, Low-frequency deep brain stimulation for Parkinson's disease: Great expectation or false hope? *Movement Disorders* **31**, 962–967 (2016).

85. 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).

86. O. Sturman, L. Von Ziegler, C. Schläppi, 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. *Neuropsychopharmacol.* **45**, 1942–1952 (2020).

87. J. J. DiCarlo, J. W. Lane, S. S. Hsiao, K. O. Johnson, Marking microelectrode penetrations with fluorescent dyes. *Journal of Neuroscience Methods* **64**, 75–81 (1996).

88. T. Naselaris, H. Merchant, B. Amirikian, A. P. Georgopoulos, Spatial Reconstruction of Trajectories of an Array of Recording Microelectrodes. *Journal of Neurophysiology* **93**, 2318–2330 (2005).

89. 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. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).

90. 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).

91. G. Paxinos, C. Watson, *The Rat Brain in Stereotaxic Coordinates* (Elsevier/Academic, 2009).

92. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2021; http://www.R-project.org/).

93. RStudio Team, *RStudio: Integrated Development Environment for R* (RStudio, PBC., Boston, MA, 2020; http://www.rstudio.com/).

94. M. L. Delignette-Muller, C. Dutang, fitdistrplus: An R Package for Fitting Distributions. *J. Stat. Soft.* **64** (2015).

95. L. Dubreuil-Vall, P. Chau, G. Ruffini, A. S. Widge, J. A. Camprodon, tDCS to the left DLPFC modulates cognitive and physiological correlates of executive function in a state-dependent manner. *Brain Stimulation* **12**, 1456–1463 (2019).

96. W. N. Venables, B. D. Ripley, W. N. Venables, *Modern Applied Statistics with S* (Springer, New York, 4th ed., 2002)*Statistics and computing*.

97. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using **lme4**. *J. Stat. Soft.* **67** (2015).

98. 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).

99. T. Womelsdorf, M. R. Watson, P. Tiesinga, Learning at Variable Attentional Load Requires Cooperation of Working Memory, Meta-learning, and Attention-augmented Reinforcement Learning. *Journal of Cognitive Neuroscience* **34**, 79–107 (2021).

100. T. V. Wiecki, I. Sofer, M. J. Frank, HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front. Neuroinform.* **7**, 14 (2013).

101. A. Gelman, D. B. Rubin, Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **7**, 457–472 (1992).

102. D. Makowski, M. S. Ben-Shachar, S. H. A. Chen, D. Lüdecke, Indices of Effect Existence and Significance in the Bayesian Framework. *Front Psychol* **10**, 2767 (2019).

103. D. R. Seibold, R. D. McPhee, Commonality Analysis: A Method for Decomposing Explained Variance in Multiple Regression Analyses. *Human Communication Research* **5**, 355–365 (1979).

104. A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, H. Pfister, UpSet: Visualization of Intersecting Sets. *IEEE Transactions on Visualization and Computer Graphics* **20**, 1983–1992 (2014).

547 **Acknowledgements**

548 **Funding**

549 This work was supported by the OneMind Institute, the Minnesota Medical Discovery Team on Addictions, the
550 MnDRIVE Brain Conditions initiative, the Tourette Association of America, and the US National Institutes of
551 Health (R01NS120851, R01MH124687). ARO was supported by the São Paulo Research Foundation
552 (FAPESP, Brazil—2017/22473-9. EDvR was supported by a National Science Foundation Graduate Research
553 Fellowship under award number 2237827. The opinions herein are fully those of the authors and do not
554 represent the positions of any funding body.

555 **Author Contributions**

556 Conceptualization: ASW, AER

557 Data curation: AER, EDvR, MEM, EMS, AW

558 Formal analysis: AER, EDvR, JK, AA

559 Funding acquisition: ASW

560 Investigation: AER, MEM, EMS, AW, EH, KS, AA, DC, M-C L, ARO, GS, NS

561 Methodology: ASW, AER, EDvR, MEM, EMS

562 Project administration, supervision: ASW

563 Writing - original drafts: AER, EDvR, JK, MEM

564 Writing - review and editing: AER, EDvR, MEM, ASW

565 **Competing Interests**

566 ASW is a named inventor on multiple patent applications related to neurostimulation and cognitive control,
567 none of which directly references the information herein and none of which is licensed to any entity.

Data and Code Availability

Datafiles and analysis code will be made available via our laboratory GitHub repositories at the time of publication.

List of Supplementary Materials

Methods

Figs. S1 to S10

Tables S1 to S26

Supplementary References

Figures

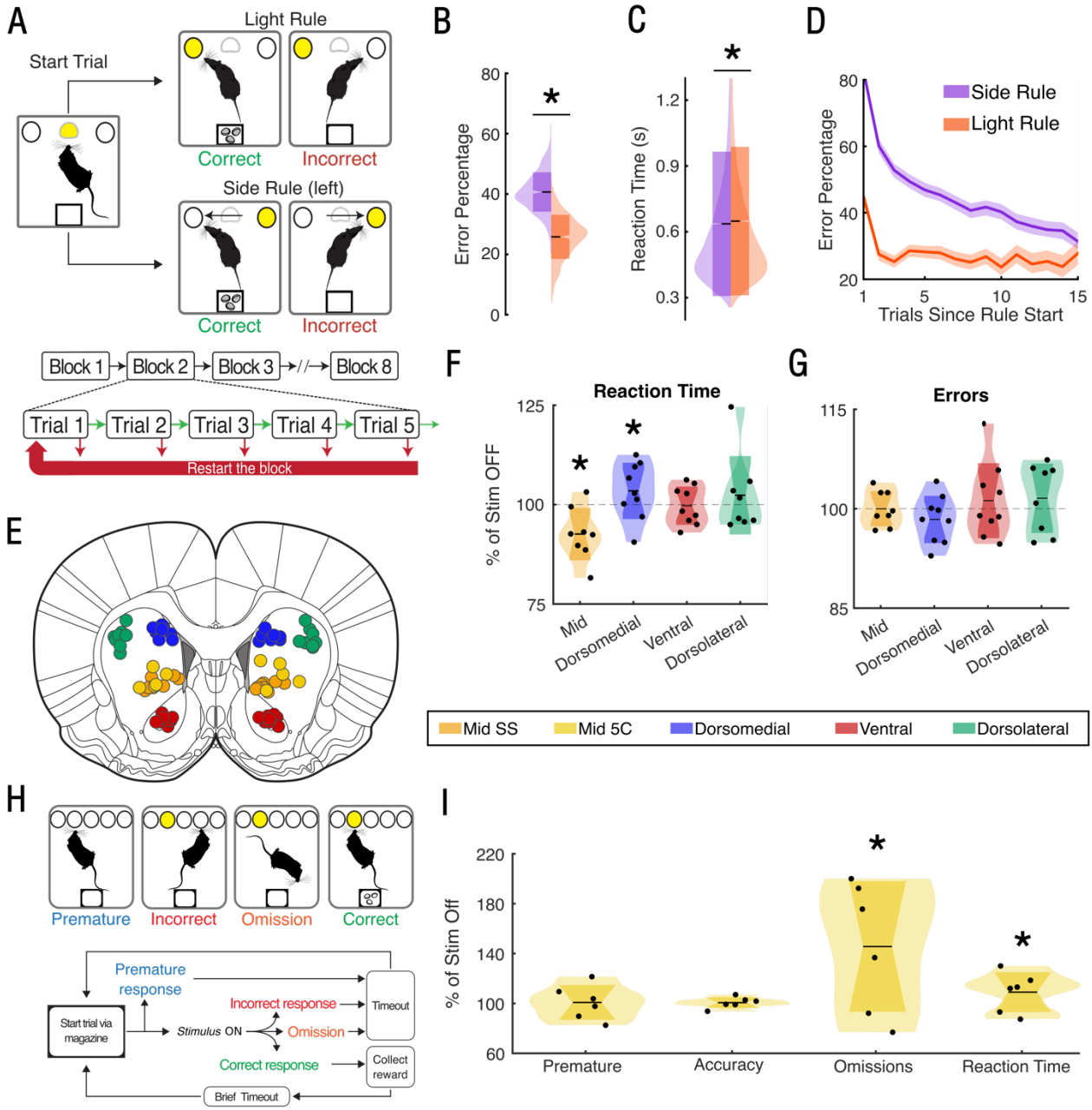

Figure 1. Behavioral testing paradigms and stimulation outcomes.

(A-D) Set-Shift task.

(A), task schematic. Rats must nosepoke to either an illuminated hole (Light rule) or ignore the light and always nosepoke on a specific side of the chamber (Side rule). The rules are not cued, and shift after the rat has completed 5 correct trials of the current rule. Errors reset to the beginning of the current 5-trial block.

585 (B), probability of error based on the current rule. The cue light acts as a pre-potent stimulus that is difficult to
586 ignore, leading to increased errors during Side blocks (β=0.64, p<0.001).
587
588 (C), RT as a function of rule. Because Side rules do not require sensory processing (the correct answer is known
589 even before the cue light is seen), they have slightly faster RTs (β=12 ms, p<0.001).
590
591 (D), probability of error after a rule shift. In both rules, rats learn the task and progress through blocks,
592 evidenced by decreasing error rate as a block proceeds. However, as in (B), Side rules are more difficult and the
593 error does not rapidly asymptote, whereas Light rules can be reliably learned in 2-3 trials.
594 (E-G), stimulation experiment.
595
596 (E), histologically confirmed sites of implant for sub-regions of the striatum. The Mid-striatal target was the
597 only target tested during the 5-choice experiment (see below), whereas all 4 targets were tested for the primary
598 Set-Shift experiment.
599
600 (F), distributions of RT as a function of stimulation type, expressed as a percentage of stimulation-off.
601 Distributions are computed over rats; each dot shows the median for one rat. Mid-striatal stimulation
602 specifically replicates the effect seen in prior human data stimulating at this target (see main text; β= -34 ms, p
603 < 0.001), whereas dorsomedial striatal stimulation worsens RT.
604
605 (G), distributions of error count (during the whole session) as a function of stimulation type, showing no
606 change, which also replicates prior human results.
607
608 (H-I), 5-choice serial reaction time task (5-CSRTT).

(H), task schematic. A trial began when a rat entered the food magazine. After a 5-second interval, a brief light stimulus was presented, pseudorandomly, in one of five different nosepoke ports on the opposite wall. If the rat poked the lit aperture, it received a food pellet followed by a 5-second timeout ("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") the house light extinguished briefly to signify failure, the rat received no reward and a 5-second timeout.

(I), outcomes, following the plotting conventions of (F-G). Mid-striatal stimulation did not increase premature responses or decrease accuracy, i.e. it did not show an impulsivity effect. It did significantly increase omissions ($\beta = 0.448$, $p < 0.001$) and increase RT ($\beta = 49$ ms, $p < 0.001$), both of which are opposite to what would be expected from impulsive behavior.

All p-values represent Wald Z-tests of model parameters from generalized linear mixed models (GLMMs), see Methods for formulae and Tables S1-S3 (Set-Shift) and S17-S20 (5CSRTT) for statistical details.

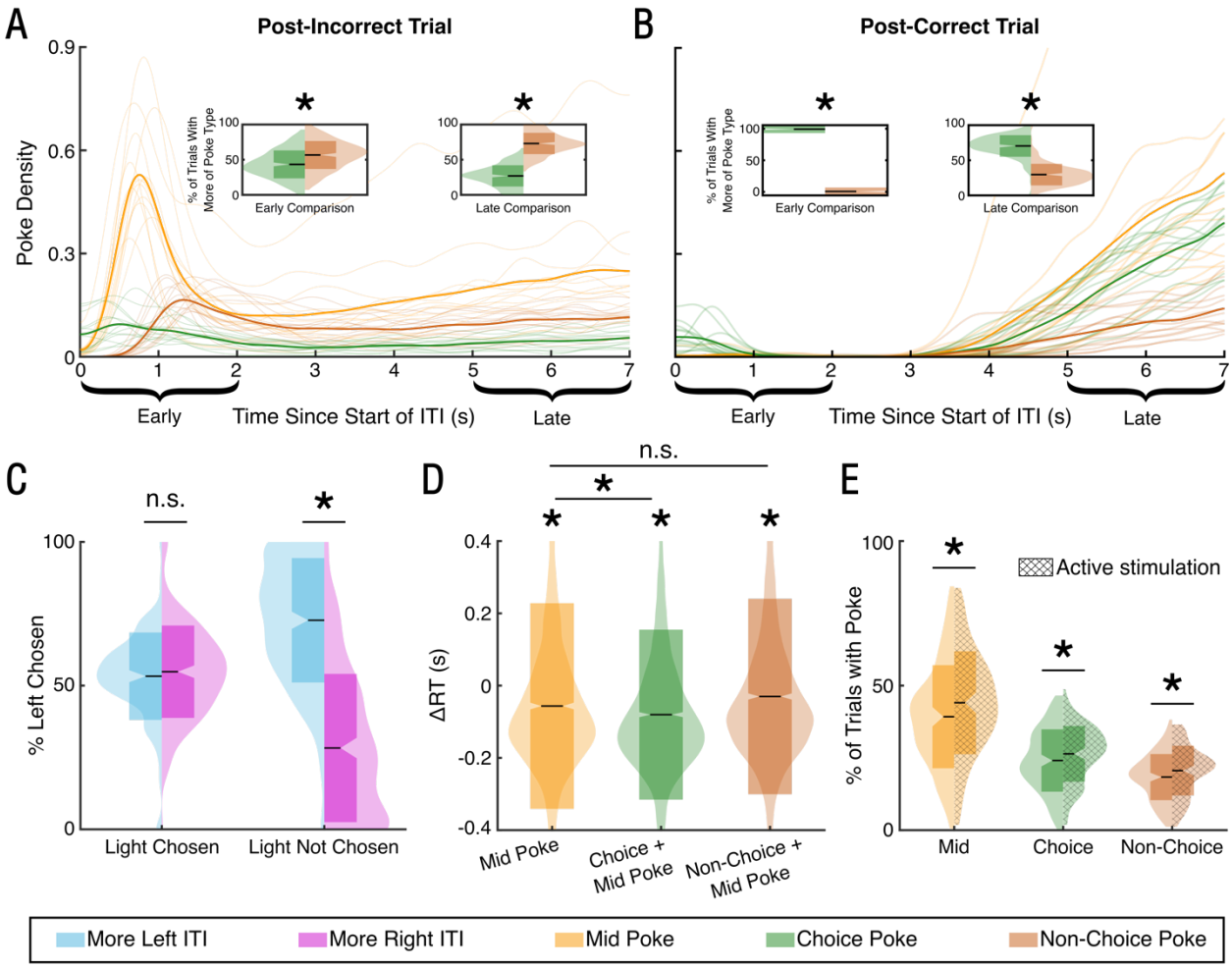

Figure 2. Precommitment related behaviors during inter-trial intervals (ITIs).

(A-B) probability of poking the mid (trial initiation) and side ports following a decision. Thick lines show the mean over all animals and sessions, while light shaded lines in background show individual animals. Each plot is a kernel density estimate of the distribution, using a Gaussian kernel with 0.2s standard deviation. "Choice" and "Non-choice" pokes are pokes in the ports that were and were not chosen in the preceding trial, respectively. We specifically emphasize "Early" (first 2 seconds) and "Late" (last 2 seconds) behavior. Following an incorrect response (A), rats showed an Early ITI peak in poking that leveled to a stable rate for the remainder of the inter-trial interval. Pokes in the side ports were significantly more likely to match the option that was not chosen in the preceding trial, both in the Early (left inset (A), β=0.53, p<0.001, Table S22) and Late ITI (right inset (A), β=1.73, p<0.001, Table S22). These resemble regretful correction and rehearsal, respectively. Following a correct response (B), rats made fewer Early pokes (β =-0.24, p<0.001, Table S21), but

increased their poking in the Late period (β =0.12, p<0.001, Table S21). Pokes in the side ports were significantly more likely to match the option that was rewarded in the preceding trial both in the Early ITI (left inset (B), β=8.27, p<0.001, Table S23) and the Late ITI (right inset (B), β=1.89, p<0.001, Table S23). These are consistent with rehearsal of a rewarded option.

(C-D) Effects of ITI behavior on the subsequent trial. In these panels, “Choice” and “Non-Choice” now refer to the upcoming trial, rather than the preceding trial as in (A-B).

(C) ITI pokes affected subsequent trial behavior (precommitment), but only when the rat chose to ignore the cue light (e.g., following successful learning of 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. If the lit port was chosen, ITI pokes had no effect (β=-0.087,p=0.08, Table S24). If the lit port was not chosen, rats were much more likely to poke the port that they had rehearsed during the ITI (β=1.84,p<0.001, Table S24).

(D) ITI pokes in the late portion of the ITI affected subsequent decision times, also consistent with precommitment. Plots show the distribution of RTs conditioned on different types of ITI behavior. Pokes in the mid port significantly lowered RTs (yellow, β_1=-0.04,p<0.001, Table S25), but there was an even stronger effect (green, β_2=-0.01,p=0.02, Table S25) when the rat had rehearsed a full response sequence, namely a mid poke along with a poke into the side port that matched the subsequent choice. In contrast, there was no additive effect of a mid poke along with a poke into the side that was subsequently not chosen (orange, β_3=0.00, p=0.44, Table S25). * above individual distributions indicates 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 mid-striatal stimulation. Stimulation increased all poke types, both those consistent with precommitment (choice pokes)

(green, β=0.21, p<0.001, Table S26) and those that might represent general motivated behavior (mid (yellow, β=0.24,p<0.001, Table S26) and non-choice pokes (orange, β=0.09, p<0.007, Table S26)). * above each distribution reflects a significant difference between stimulation on and off.

All p-values represent Wald Z-tests of model parameters from generalized linear mixed models (GLMMs), see Methods for formulae.

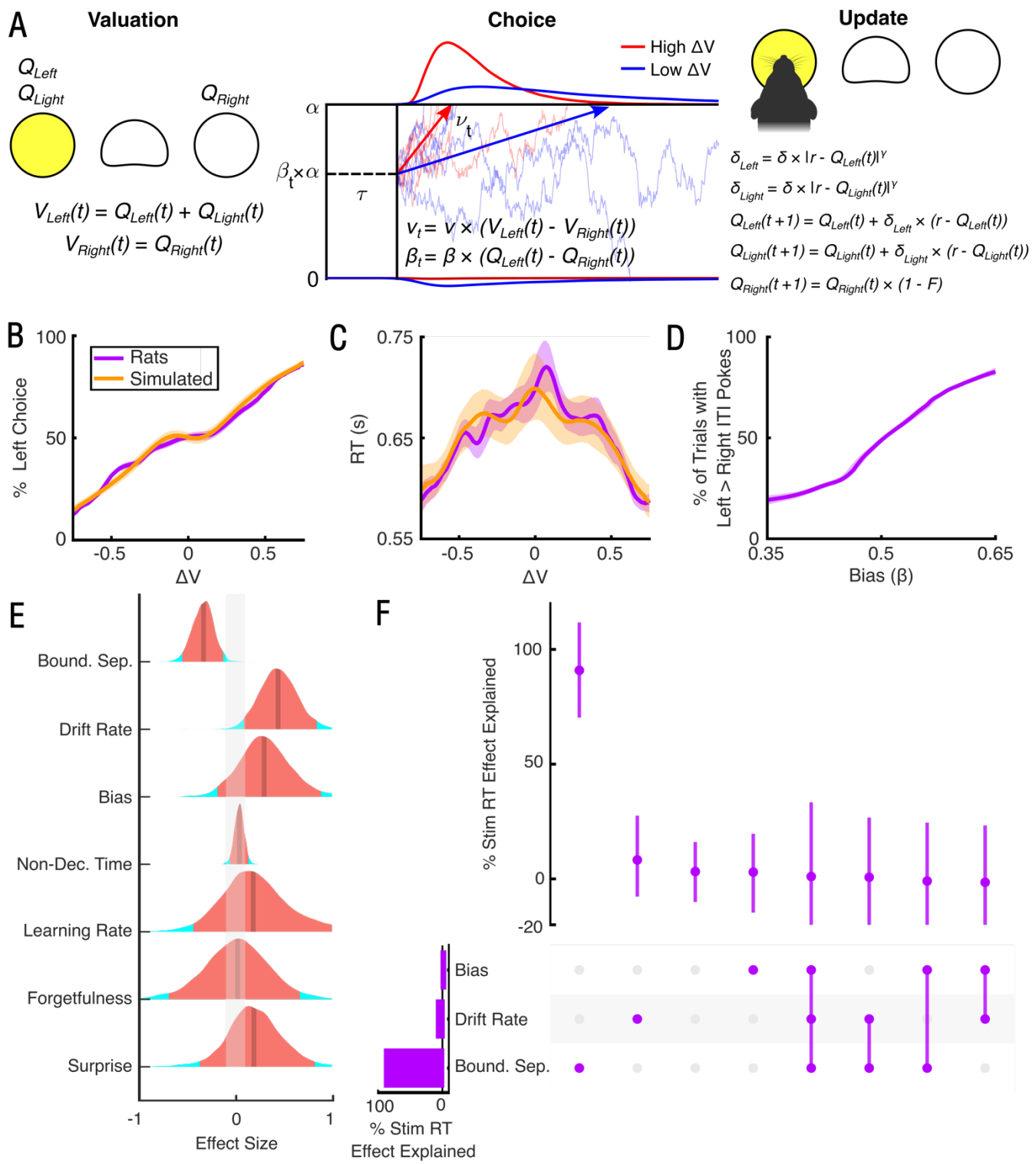

Figure 3. Mid-striatal stimulation decreases reaction times specifically by improving evidence processing and conflict resolution.

672 (A) Illustration of the Reinforcement Learning-Drift Diffusion Model (RLDDM) fit to Set-Shift behavior. The
673 RLDDM first assigns values V to the ports based on its underlying estimates Q of the current value of the sides
674 (left, right) and the light. The model then chooses a port based on a drift-diffusion process described by non-
675 decision time τ, drift rate ν, and decision boundary α. The difference in total value informs the drift rate, while
676 the difference between the sides alone informs the bias ß. Trials with high value difference are likely to have
677 shorter reaction times and more consistently correct choices (red vs. blue curves above). After the port choice,
678 values Q are updated based on a reinforcement learning process, with a non-linear scaling factor γ to diminish
679 small errors and accentuate large ones. The value of the unchosen option decays according to a forgetfulness
680 factor F. The model was fit hierarchically to the data from all rats, allowing individual-specific estimates of
681 model parameters and their change with stimulation.
682
683 (B-D) Posterior predictive simulations of behavioral outcomes for models (orange) and rat behavior (purple)
684 over 4000 hierarchical posterior draws; see Methods. Shading shows the 95% highest density interval across
685 animals and trials.
686
687 (B) Choice as a function of the difference in value between the left and right sides (ΔV), where higher X-axis
688 values represent a higher value for the left port. Consistent with expected psychometrics, rats and model
689 simulations were both substantially more likely to select the more valuable side, with a small plateau of
690 equiprobability around the $\Delta V=0$ equipoise point.
691
692 (C) RT as a function of the difference in value between the left and right sides, with same conventions as (B).
693 Reaction times were greater for smaller value differences than larger ones, for both rats and simulated model
694 data. This is consistent with greater decision difficulty as the choices become near-equivalent.
695

(D) Probability of an ITI poke as a function of the model's bias term. Higher X-axis values correspond to greater bias ß towards the left port, while the Y-axis shows the percentage of left vs. right pokes during the ITI. As the bias term increased, so did rats' tendency to rehearse the left port during the ITI. This bias term can therefore be treated as a smoothed estimate of the planning/rehearsal behaviors during ITIs (Figure 2). The simulated and empirical curves entirely overlap.

(E) Distributions of stimulation effect on model parameters over 4000 posterior draws. The median of the distribution is indicated by the solid line while the 95% highest density interval is in orange. The grey shaded area around zero represents the Region of Practical Equivalence (ROPE) for a null effect (effect size less than 0.1). Electrical stimulation of mid-striatum substantially decreased boundary separation α (d=-0.33, pd=0.999, 0% in ROPE), substantially increased drift rate ν (d=0.44, pd=0.993, 0.6% in ROPE), and marginally increased bias ß (d=0.30, pd=0.885, 15.4% in ROPE). None of the other parameters showed meaningful effects with stimulation. As described in the main text, the interaction of these parameter changes represents an increase in general cognitive control.

(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. The stimulation related decrease in reaction time was fully explained by unique contributions from boundary separation (91%, HDI=[70,112]) and drift rate (10%, HDI=[-8,28]) with negligible effects from the interaction between these or other parameters.

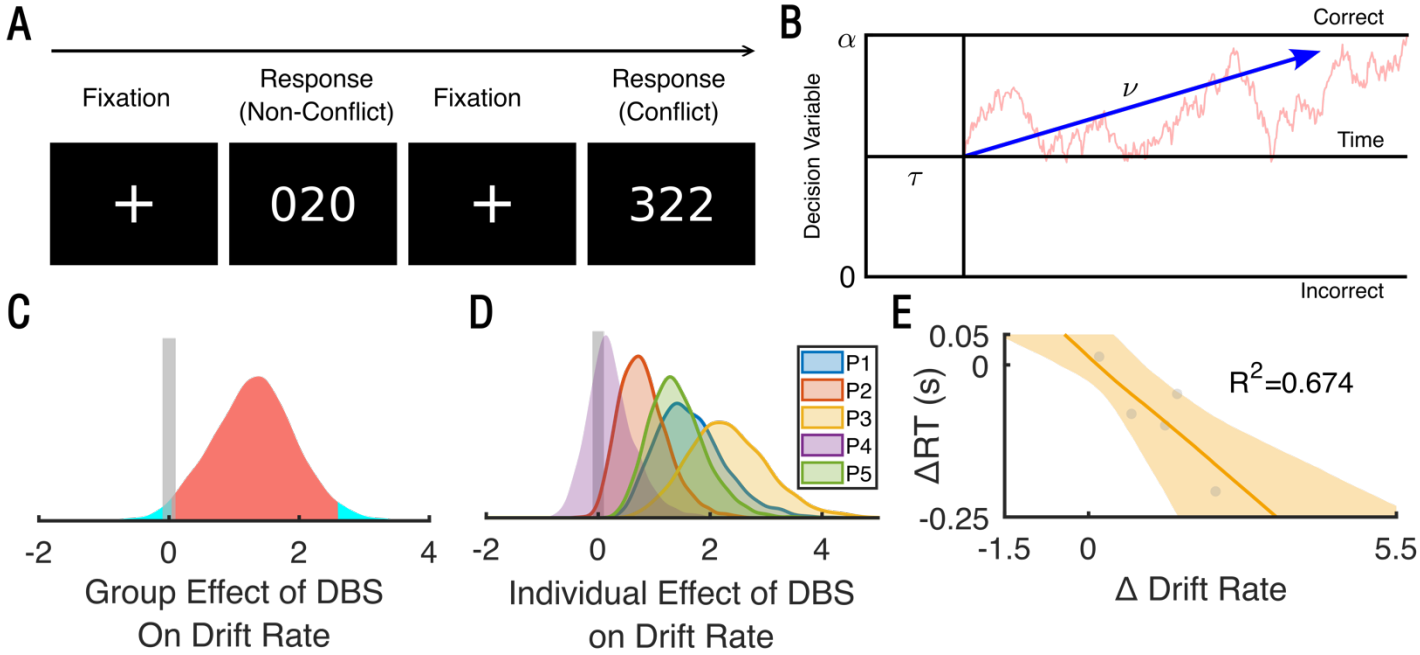

Figure 4. Changes in evidence processing identified in rats also explain the previously reported human effects. (A), schematic of the Multi Source Interference Task (MSIT). Participants choose the number that is different from the other two. On Conflict trials, the target number is out of position and flanked by distractors that are valid targets, leading to interference that requires cognitive control to suppress (*27*, *65*). In Non-Conflict trials, the target is in position and flankers are not targets. In (*27*, *28*), we reported that striatal/internal capsule stimulation improves RT on this task.

(B), model schematic. We fit a standard drift-diffusion model (DDM), exploring the potential effect of striatal stimulation on boundary separation α and/or drift rate ν. Based on Figure 3's results, we did not model potential effects on non-decision 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 while the 95% highest density interval is in orange. The grey shaded area around zero represents the Region of Practical Equivalence (ROPE) for a null effect (effect size less than 0.1). Electrical stimulation of mid-striatum substantially increased the drift rate (d=1.31 pd=0.978, 1.47% in ROPE).

(D), individual distributions of drift rate changes, over 4000 posterior draws, on the same scale as (C). In 4 of the 5 participants, stimulation substantially increased drift rate.

(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 panel 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 panel D) against the change in median RT.