Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex

Kyle S. Smith*,†, Arti Virkud*, Karl Deisseroth*, and Ann M. Graybiel*†

*McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; and †Department of Bioengineering, Stanford University School of Medicine, Stanford, CA 94305

Contributed by Ann M. Graybiel, September 18, 2012 (sent for review August 20, 2012)

Habits tend to form slowly but, once formed, can have great stability. We probed these temporal characteristics of habitual behaviors by intervening optogenetically in forebrain habit circuits as rats performed well-ingrained habitual runs in a T-maze. We trained rats to perform a maze habit, confirmed the habitual behavior by devaluation tests, and then, during the maze runs (ca. 3 s), we disrupted population activity in a small region in the medial prefrontal cortex, the infralimbic cortex. In accordance with evidence that this region is necessary for the expression of habits, we found that this cortical disruption blocked habitual behavior. Notably, however, this blockade of habitual performance occurred on line, within an average of three trials (ca. 9 s of inhibition), and as soon as during the first trial (<3 s). During subsequent weeks of training, the rats acquired a new behavioral pattern. When we again imposed the same cortical perturbation, the rats regained the suppressed maze-running that typified the original habit, and, simultaneously, the more recently acquired habit was blocked. These online changes occurred within an average of two trials (ca. 6 s of infralimbic inhibition). Measured changes in generalized performance ability and motivation to consume reward were unaffected. This immediate toggling between breaking old habits and returning to them demonstrates that even semiautomated behaviors are under cortical control and that this control occurs online, second by second. These temporal characteristics define a framework for uncovering cellular transitions between fixed and flexible behaviors, and corresponding disturbances in pathologies.

learning | limbic | time | reinforcement | neuroplasticity

Overtrained T-Maze Running Is Habitual. We used a navigational T-maze task (Fig. 1A) similar to one developed to examine neural activity in the striatum during habit formation (13–15). We tracked the learning curves of multiple sets of rat subjects (n = 10). In fully trained sessions of ca. 40 trials, the rats were required to initiate maze runs in response to a warning cue, run down the maze, and turn right or left, depending on an auditory instruction cue, to receive one of two rewards (chocolate milk or sucrose solution). For each rat, each reward type was assigned to one end arm, and entry into an incorrect arm resulted in no reward. Rats were trained to a criterion of statistically significant performance accuracy (72.5% correct; criterion training stage 5; Fig. 1B) and then were overtrained for 10 or more additional sessions. Peak performance accuracy was high at ~90% correct (Fig. 1B).

We applied a reward-devaluation protocol as a quantitative test for habit formation after this extensive training (16). We devalued one of the two maze rewards in home-cage sessions and then tested in the maze experiments for habitual running to the end arm baited with the now-devalued reward, relative to runs to the normally valued reward in the other end arm. To induce devaluation, we paired one reward with a nauseogenic dose of lithium chloride in home-cage sessions (17, 18). In home-cage tests, we confirmed that this method produced an aversion to the devalued reward (Fig. 1C). After devaluation, the rats were placed in the maze to run the task again, but the end arms were not baited. This procedure allowed us to determine whether the animals would still run when instructed to the side with the now-devalued reward, suggesting that this running was habit-driven.

The behavior of the control rats (n = 4) established that overtrained T-maze running was normally habitual. In the probe test after reward devaluation, these rats continued running to the devalued goal when so instructed, just as they had in the session before devaluation (Fig. 1D). Their runs to the nondevalued goal were unchanged and accurate as well (Fig. 1D). This insensitivity to reward devaluation confirmed that the training procedures produced ingrained, habitual maze runs.

To test for the effects of IL cortical activity on this habitual behavior, we used the strategy of perturbing the IL cortex during the unrewarded probe session. We applied the intervention exclusively during the performance of the runs to ensure that it was online contribution of cortical activity to the control of habitual behavior, but also, because of its unique repeatability, to track the effects of such brief, seconds-long manipulations over weeks of subsequent performance. To our surprise, we found that despite the automaticity of habitual behavior, it is subject to online control by the prefrontal cortex.

The authors declare no conflict of interest.

1To whom correspondence may be addressed. E-mail: kyless@mit.edu or graybiel@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216264109/-/DCSupplemental.
Online. We did not include either the prerun periods, the periods during reward delivery, or intertrial intervals. In the experimental rats (n = 6), an AAV-5 viral vector encoding eNpHR3.0 fused to EYFP, and targeted preferentially to glutamatergic pyramidal neurons through the calcium/calmodulin-dependent protein kinase IIα promoter (AAV5-CaMKIIα-eNpHR3.0-EYFP) (12), was injected bilaterally into the IL cortex 1–3 mo before behavioral training (Fig. 2A). Anterograde viral labeling was later observed in the medial striatum but not in the sensorimotor striatum, confirming a lack of a direct corticostriatal projection from the IL cortex to the dorsolateral region of the striatum known to also be essential for habit expression (Fig. 2B) (3, 10). Two of the control rats were given laser exposure following injections of control virus lacking halorhodopsin (AAV5-CaMKIIα-EYFP). The other two were injected with halorhodopsin-containing virus, but the laser was inactive.

Optical Stimulation of Halorhodopsin-Expressing IL Neurons Inhibits Spike Activity. IL manipulation was accomplished by delivery of 593.5-nm-wavelength laser light through optical fibers implanted bilaterally so as to extend to the dorsal edge of the targeted IL cortex (Fig. 2A). Partial inhibition of IL activity was verified in two behaving rats fitted with a head stage carrying tetrodes and optical fibers by applying equivalent illumination parameters to those that would be used in the maze task (Fig. 2C–G). Neuronal spike activity was suppressed by 56% on average during illumination in 11 units, and the suppression was consistent over 40+ trials. This protocol resulted in a more temperate, as well as temporally specific, suppression of spike activity than the full blockade of activity produced, presumably, by lesions or drug infusions. In six units, we observed increased firing rates during light delivery, and these units were often recorded on the same tetrode as others showing inhibition (Fig. 2F). Thus, the optogenetic inhibition targeted to pyramidal cells influenced local microcircuitry (19, 20), yielding a time-locked disruption of population spike activity (inhibition:excitation ratio, 1.8:1).

IL Perturbation Blocks Habits Online. We began by disrupting the IL cortex online while the rats performed the postdevaluation probe test (Fig. 3). Light-on occurred immediately after gate-opening, and light-off occurred immediately after goal-reaching (ca. 3 s of light per trial; <2 min total per day). This within-run treatment produced a dramatic effect: the rats acted as though they had not acquired the habit of running to the devalued goal. As shown in Fig. 3D, the rats with IL perturbation sharply reduced their runs to the end arm that would have had the devalued reward—runs characterized made by the overtrained control rats (Fig. 3E). Instead, during the intervention, the rats had a propensity to run to the nondevalued end arm (Fig. 3A and B). Such avoidance behavior was seldom observed in the control group (Fig. 3B).

The devaluation sensitivity induced by IL perturbation was evident almost immediately (Fig. 4C). It took on average of three light trials for rats to begin avoiding the devalued goal where instructed there. In one animal, the avoidance was present during the first trial. This onset of habit blockade corresponded to an average within-run illumination time of ca. 7–9 s, with the most immediate effect within 3 s in the single rat. The online IL intervention did not have a generalized effect on the performance ability of the rats, however, because they ran accurately and seemingly automatically with IL illumination when they were instructed to run to the nondevalued reward side (Fig. 3D). These experiments suggested that IL cortical activity was required online, during the actual maze runs, in order for the expression of running behavior as a habit.

Replacement Habit Forms with Postdevaluation Training. We next analyzed the behavior of the rats when they again received reward for correct performance in subsequent days of maze training (Figs. 4D and E). In accordance with classic findings (3, 17, 18, 21), all of the rats, including the control animals, avoided the devalued end arm after being reexposed to the reward that had been devalued (Figs. 4A and B and 5A and B), and they almost never consumed it (Fig. 5C). Thus, the control rats required actual contact with the devalued reward on the maze to trigger a loss of habitual behavior that rats with a disrupted IL cortex already exhibited in the unrewarded probe session.

Under IL perturbation, when the rats avoided the now-devalued goal, they almost never stopped the task. Instead, they continued to run to the nondevalued end arm (Figs. 4B and 5B). These “wrong-way runs” increased in frequency over days, despite the fact that the rats had not been instructed to go to that end arm and did not ever receive reward for these runs. The rats showed no overt anticipatory behavior such as licking or signs of distress at the lack of reward. In control rats, the high frequency of wrong-way runs lasted for as long as we tested (>3 wk). Moreover, these runs also appeared immune to the modest loss of in-maze aversion to the devalued outcome that appeared to occur during this time, as indicated by a small recovery of in-maze drinking of the devalued reward when the rats did run to it (Fig. 5C). This pattern of behavior suggested that the rats developed a new habit in these postdevaluation days, namely of always running to the nondevalued end arm.

IL Perturbation Blocks Replacement Habit and Returns Original Behavior. The nearly immediate loss of habitual behavior during online perturbation of IL activity suggested that switching off IL cortex might be like flipping an off-switch for habitual behavior, consistent with prior work (8, 9). To evaluate this possibility further, we extended the laser experiments after the original probe test, gradually increasing the time for up to a month. The IL cortex was first disrupted during one to two rewarded sessions immediately after the initial probe test intervention [postprobe stage (PP)1 (first laser-on days after probe); Figs. 4 and 5]. This intervention produced no detectable

**Fig. 1.** T-maze training and habitual behavior of control rats. (A) Protocol for task training. (B) Performance across training stages for control rats. (C) Home-cage reward drinking before (Pre) and after (Post) devaluation. (D) Performance during the last session before devaluation (Left) and then during probe test after (Right), for control rats (solid, cued to devalued goal; dashed, cued to nondevalued goal); NS, not significantly different. *P < 0.05. Data are presented as means ± SEM throughout.
effect on behavior; both groups of rats equally avoided the devalued end arm on most trials. Nor did further IL perturbation introduced at up to 6 d after the initial procedure change the behavior of the rats (n = 2), which remained stable for up to 14 d [stages PP2–PP5 (laser-off days); Figs. 4–6]. This result suggested that once the habitual behavior had been blocked, and the IL-disrupted rats were outcome-guided in behavior like the control rats, further IL disruption failed to affect maze runs.

We obtained sharply different results when we extended the time before imposing the further perturbation (Figs. 4–6). When we again disrupted the IL cortex online 2–3 wk after the initial disruption, on days 13 (n = 4), 15 (n = 1), or 21 (n = 1) (stage PP6), the effect was immediate and surprising: the rats now avoided running to the devalued reward on this prior day, as so cued 6 times on average and drank it 2.17 times on average before the late IL perturbation, ran to the devalued reward when so cued (Fig. 5A–D). This change in behavior was abrupt. These rats, on the day before the late IL perturbation, ran to the devalued reward when so cued 6 times on average and drank it 2.17 times on average (~0.7 mL), at most twice consecutively (Figs. 4D and 6A and B). By contrast, on the following day with IL illumination, the rats ran to the devalued reward and drank it earlier in the session, quickly reached an equivalent number as on the prior day, surpassed it by the fifth such trial on average, and then kept going: they ran there and drank 14.5 times on average (~4.4 mL), often in long repeated sequences of consecutive runs and drinks (Fig. 6B). Simultaneously, this late IL perturbation blocked the new wrong-way running habit that had developed. Both effects occurred within seconds, and within few trials, as had the original habit blockade. Control rats rarely ran to the devalued reward on the equivalent test trials (mean runs and drinks, 1.25 times) and, instead, continued to mainly avoid the devalued reward (Figs. 4–6, black lines).

This apparent return of the original habitual behavior after the late IL disruption remained during subsequent laser-off days (stages PP7–PP8, up to 4 test days; Figs. 4 and 5). When we then administered another laser session (stage PP9), this treatment fully returned the original habitual behavior, which remained for as long as we tested (up to 20 d after the third silencing, stage PP10) (Figs. 4 and 5). Over the same length of time, control rats continued to avoid the devalued goal and to not consume it (Figs. 4 and 5).

Day-by-day analysis of the running patterns suggested that a tipping point for return of the original learned behavior occurred between 6 and 13 d after the initial probe test, following which the effect of IL perturbation changed from blocking the initial habit to seemingly promoting it (Fig. 6C). This reversal of the effect of IL perturbation corresponded to the time when the new wrong-way runs had been repeated over a number of sessions. The coordinate effects on the original and new behavior of
the rats suggested that IL perturbation might have turned off both the initial habit and the new habit and that blockade of the second habit might, thus, have uncovered the original habitual behavior.

**IL Perturbation Does Not Affect Generalized Motivation to Consume the Devalued Reward or Taste-Aversion Memory.** We tested the possibility that the IL manipulation was changing a generalized motivation to drink or the associated taste-aversion memory irrespective of maze habits by examining home-cage drinking of the devalued reward on test days following the late IL perturbation (around stage PP9). The cages were in the maze room to provide context similarity (Fig. 5E). On sequential test days, light was delivered or not delivered while the reward was freely available to the rats in their cages (n = 6). This IL manipulation had no effect on drinking of the devalued reward (Fig. 5E, “Late”), which was also similar to home-cage drinking assessed at an equivalent time in control rats (Fig. 5E). We likewise tested the effect of IL perturbation on in-cage drinking just after devaluation and again found no effect (Fig. 5E, “Early”). This lack of effect sharply contrasted to the major effect of IL intervention on drinking after correct performance in the maze experiment proper, both at the first IL intervention and at the late interventions (Figs. 5C and 6B). Thus, if IL perturbation was affecting appetitive motivation, it was doing so only in the maze.

The in-cage test showed that both the control and IL-manipulated rats drank more at the late time than they did right after devaluation, with control rats reaching 35% (ca. 10 mL) of pre-devaluation drinking. This partial recovery also contrasted sharply with the consistently low levels of drinking in the maze during the same time period: in the maze, the rats still rejected drinking over half of the times that they ran to it (Figs. 5C and 6A). Thus, even though the incentive value of the reward at home was partly recovering over time, it had remained weak in the maze normally. These disconnects between in-cage and in-maze drinking accord with the strong context-dependency of habits.

**Discussion**

Our findings demonstrate that well-ingrained habits can be controlled online by optogenetic manipulation of a specific site in the prefrontal cortex, the IL cortex. Strikingly, this control was effective even when exerted only during performance of the behavior. Moreover, although optogenetic inhibition of the IL cortex could block expression of an ingrained habit, further IL inhibition could return this habitual behavior if enough time had elapsed to allow the formation of a new habitual behavior. These effects occurred within seconds of online performance time.

These findings carry implications for the temporal dynamics and online scope of behavioral control exerted by the neocortex over habitual behavior. First, despite the seemingly automatic behavior typified by habits, their behavioral automaticity requires ongoing permissive or supervisory activity in the medial prefrontal cortex. Second, this prefrontal control appears to favor new habits if there is a competition between old and new habits; IL perturbation can block an old habit abruptly but can also abruptly bring back an old habitual behavior by blocking the newer one. Third, online manipulation of the medial prefrontal cortex affects the expression of habitual behavior almost immediately (within a performance cycle or two, totaling only

![Fig. 3](image-url) **Fig. 3.** Optogenetic perturbation of IL cortex blocks habits. (A) IL-illumination protocol. (B) Equivalent performance across training for halorhodopsin-treated (red) and control (dashed gray) rats before devaluation and silencing. (C) Home-cage drinking pre- and postdevaluation for IL-halorhodopsin (IL-halo) rats. (D) Performance during last session before devaluation (Left) and probe test after (Right), for rats with IL-halo (solid, cued to devalued goal; dashed, cued to nondevalued goal). (E) Performance during the probe session for control and IL-halo rat groups. Interaction of goal value and rat group on probe: *P < 0.05; *P < 0.01; **P < 0.001. NS, not significantly different.

![Fig. 4](image-url) **Fig. 4.** IL perturbation reinstates original habit and simultaneously breaks replacement habit. (A) Percentage of trials in which rats instructed to run to the devalued goal went correctly in which also drank the reward (IL-halo group; red; control group; black; see Fig. 5 for runs and drinks separately). Shown are five sessions leading up to reward devaluation, the unrewarded probe session (only correct runs shown), and the postdevaluation stages (PP1; initial light-on session(s) or first postprobe session for controls; PP2-PP9; sessions without light; PP10; light-on session or equivalent for control; PP9: light-on session; PP10: final session without light). *P < 0.05; **P < 0.01; *P < 0.001 between groups within stage; differences lacking symbols indicate not significantly different. (B) Percentage of wrong-way runs to nondevalued goal. (C) Trial-by-trial plot for the IL-halo group and control group, in which rats ran the wrong way during probe session (i.e., blockade). For each trial, wrong-way run was scored “0,” and then scores were averaged (e.g., “0.5”); half of rats ran the wrong way and half ran correctly, “1”; all rats ran the wrong way. (D) Trial plot for IL-halo and control groups, in which rats ran to the devalued reward when instructed and drank it, during stage PP6 (i.e., habit reinstatement).
seconds of disrupted neuronal activity) when exerted only during performance, not during pretask anticipation and planning or during postperformance reinforcement or consolidation times.

These results support the findings of classic work on the habit system suggesting that in rodents, the IL cortex is necessary for habit expression (7–9) and point to the startling extent and potency of control exerted by this small cortical region. Our experiments also generated the unanticipated finding that IL perturbation can have the opposite effect when applied later to the same animals, when they had acquired a new habit: it can result in the expression of the apparently same habitual behavior that earlier IL perturbation originally blocked. Runs to the nondevalued goal, and home-cage drinking, were unaffected by the intervention, ruling out generalized effects on performance ability, motivation to drink, or devaluation memory.

One interpretation of these findings is that after devaluation, the original habit lost access to behavior, but its representation was maintained in the brain, and the late IL perturbation unmasked it. This conclusion is in good accordance with evidence, dating from the time of Pavlov (22), that when a habit is broken, it is not forgotten; rather, a new one replaces it. By this view, the IL cortex might serve, in part, as an online executive controller that favors newly acquired habits over old strategies (8, 11). Certainly, the IL cortex did not appear to act as a simple bidirectional on/off switch for habitual behavior: when we performed the second manipulation before the 1- to 2-wk period after which the new habit was well developed, the perturbation did not reinstate the original habit (nor did it block the emerging new behavior). We take this result to suggest that the reinstatement was locked to the blockade of the second habit. The view that the IL cortex supervises newly established habits over old strategies, even habits, is consistent with evidence that the IL cortex helps maintain current response tendencies when they compete with prior ones (7, 11, 23).

An alternative view is that the late perturbation could have returned the rats to a state of value-driven behavior. There is usually a close coupling of reward-proximal stimuli and actions to current value (7, 21, 24, 25); the fact that the rats drank the devalued reward after this late manipulation of IL suggests that the reward might no longer have been aversive. However, this interpretation must face three issues. First, in-maze drinking was consistently low throughout testing in control rats and was low in the IL-halorhodopsin rats up to the putative “reinstatement.” Even when these rats ran to the devalued reward, they drank it <50% of the time; they had the opportunity to drink more on the maze but did not. Thus, the devalued reward was consistently aversive on the maze up to the point of IL-perturbation. However, then, during the late IL intervention, pursuit of this devalued reward jumped far above this level. Second, the IL-halorhodopsin rats had experience with the devalued reward in sessions before IL perturbation on the few trials they drank it, and yet this experience failed to evoke a return of the original

Fig. 5. Maze running and reward drinking during IL perturbation. (A and B) Percentage of correct performance on trials with instruction to devalued goal (A) or to nondevalued goal (B). *P < 0.05; **P < 0.01; ***P < 0.001; differences lacking symbols indicate not significantly different. (C and D) Percentage of drinking of devalued reward (C) and nondevalued reward (D) when the trial was run correctly. The inverse is runs to the devalued goal without drinking. (E) Home-cage drinking of devalued reward for IL-halo group in two light-on and light-off sessions after devaluation (“Early” test, conducted around PP1–PP2), and in light-on and light-off sessions after habit reinstatement (“Late” test, conducted around PP9). *P < 0.05 compared with predevaluation intake (“Pre,” dashed gray line); **P < 0.05 compared with intake just after devaluation (“Post”). (Right) Home-cage drinking of devalued reward for control group (around PP9).
behavior. Only during the IL intervention, and only in the later period of testing (PP6), did rats continue to pursue the devalued goal beyond a few samples. Third, the return of pursuing the devalued goal was nearly immediate, as soon as the very first laser trial, suggesting that the runs were based on a stored value rather than rats changing their performance within the session after contact with the devalued reward.

The maze behavior we analyzed here constituted a complex habit, with components related to the two different rewards, only one of which was devalued. Our findings suggest, as a favored interpretation, that the return of the original behavior of running to the devalued goal was a readjustment of a component of the larger behavioral repertoire learned by the rats and may have influenced the coupling of the components and subcomponents of running and drinking as well.

This evidence, collectively, places the temporal control of habit expression in a paradoxical context: despite the apparent automatic performance of habits, classically considered as outcome-independent and noncognitive, the prefrontal cortex still monitors ongoing contingencies time step by time step, and does so with the capacity to reverse the semiautomatic behavioral expression. Our findings further raise the possibility that scripts for alternate habitual behaviors are somehow stored when not expressed and that they can be unmasked if IL activity is disturbed, suggesting coordination with other brain regions. Our experiments do not settle how this coordination is accomplished or whether it is accomplished exclusively online. Still, our anatomical evidence from using the viral vector to trace the connections from the IL-injection sites confirmed that this region did not have detectable direct connections with the sensorimotor striatum but, rather, with regions (10) that should give the IL cortex direct access to circuits involved in flexibility and reinforcement as well as addiction (e.g., via projections to prelimbic neocortex and to the medial and ventral striatum) (3, 4, 6, 7, 15, 23, 25, 26), and also to habit-promoting circuits through the central amygdala (27). Each could be important for the IL habit-toggling function. These observations raise the question of whether the IL cortex, or its human brain homolog, could similarly control addictions or states in which behavioral flexibility and behavioral fixity are out of balance, as seen in major neurologic and neuropsychiatric disorders (5, 6, 23, 26, 28–30). Evoking fast, robust, and enduring behavioral change by targeting the IL cortex or its homolog could be of substantial value in treating disease states in a range of clinical settings, in addition to serving as a powerful approach to studying the real-time making and breaking of normal habitual behaviors.

**Materials and Methods**

Rats were trained on a T-maze task requiring them to respond to auditory instruction cues by turning into maze end arms to receive reward (~0.3 mL chocolate milk or sucrose, each paired with a distinct cue). Training proceeded over daily sessions through task acquisition (72.5% accuracy) and through 10+ additional overtraining sessions. For reward devaluation, rats received three pairings of free home-cage intake with lithium chloride injection and were later returned to the task for an unrewarded probe session and rewarded sessions. Task events were controlled by computer software (MED-PC; Med Associates). Behavior was monitored by in-maze photoembs and an overhead video camera. For optogenetic manipulation, injections of AAV5-eNpHR3.0-CaMKII–EYFP or AAV5-CaMKII–EYFP were made bilaterally into the IL cortex, and bilateral dual-ferrule optical fibers were implanted to terminate at the dorsal aspect of the IL cortex. To perturb reward-related signaling, yellow light (2.5–5 mW) was delivered from the warning cue to goal arrival (ca. 3 s). In tests to analyze spiking dynamics, neuronal activity was recorded from 12 to 24 tetrodes while light was delivered (3-s-on/10-s-off pulses). Immunostaining and Nissl-staining procedures were used to label tetrode tracks, fiber optic cannulae tracks, and YFP-positive neurons. ANOVA and neuronal spike distribution statistics were used to assess basal ganglia and behavioral activity changes, with significance set at \( P < 0.05 \). Also see SI Materials and Methods.

**ACKNOWLEDGMENTS**

We thank Christine Keller-McGandy, Alex McWhinnie, Dan Loring, Henry F. Hall, David Wu, and Asia Kubota, and Dordaneh Sugano for their technical help. This work was supported by National Institutes of Health (NIH) Grants R01 MH063797 to (A.M.G.) and F32 MH085454 (to K.S.S.), by the Stanley H. and Sheila G. Sydney Fund (A.M.G.), and grants from the NIH, Defense Advanced Research Projects Agency, and the Gatesy Foundation (to K.D.).