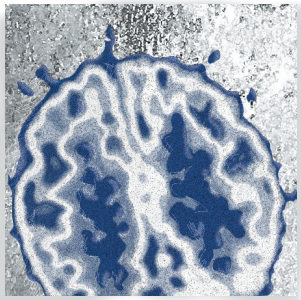
Habit formation

Kyle S. Smith, PhD; Ann M. Graybiel, PhD



Habits, both good ones and bad ones, are pervasive in animal behavior. Important frameworks have been developed to understand habits through psychological and neurobiological studies. This work has given us a rich understanding of brain networks that promote habits, and has also helped us to understand what constitutes a habitual behavior as opposed to a behavior that is more flexible and prospective. Mounting evidence from studies using neural recording methods suggests that habit formation is not a simple process. We review this evidence and take the position that habits could be sculpted from multiple dissociable changes in neural activity. These changes occur across multiple brain regions and even within single brain regions. This strategy of classifying components of a habit based on different brain signals provides a potentially useful new way to conceive of disorders that involve overly fixed behaviors as arising from different potential dysfunctions within the brain's habit network.

 Dialogues Clin Neurosci. 2016;18:33-43.

Keywords: *action chunking; action-outcome; addiction; basal ganglia; corticostriatal circuit; obsessive-compulsive disorder; repetitive behavior; stimulus-response*

Introduction

As the American philosopher William James wrote, habits make up a major part of our behavioral and cognitive lives.[1] The emphasis on experiment-based logic since that time and the enduring interest in habits in the research community have given us a rich set of approaches to study the brain basis of habit formation. For the most part, these measures center on behavioral tasks designed to test whether a learned response is driven by stimulus-response (SR) associations or by more cognitive or prospective processes. And yet, the "SR habits" so defined are hypotheses based on these measures, and each idea about them has its own potential limitations. We take here an alternate strategy: classifying habits into potential component features on the basis of new findings about the changes in patterns of neural activity that occur as simple habits are formed and broken, both within and across habit-related brain circuits. This framework reconsiders habits as being formed through multiple, simultaneously signaling processes in the brain.

Author affiliations: Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire, USA (Kyle S. Smith); McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA (Ann M. Graybiel)

Address for correspondence: Dr Kyle Smith, 6207 Moore Hall, Hanover, NH 03755, USA and Dr Ann Graybiel, 43 Vassar Street, Building 46, Room 6133, Cambridge, MA 02139, USA
(e-mail: kyle.s.smith@dartmouth.edu; graybiel@mit.edu)

Selected abbreviations and acronyms

AO	*action-outcome*
CeA	*central nucleus of the amygdala*
DLS	*dorsolateral striatum*
DMS	*dorsomedial striatum*
IL	*infralimbic*
OCD	*obsessive-compulsive disorder*
OFC	*orbitofrontal cortex*
PL	*prelimbic*
SNc	*substantia nigra compacta*
SR	*stimulus-response*

Classifying features of habit formation

Historical framing of habit formation

The historical definition of habits is that they are behaviors rooted in SR associations that have been acquired through learning based on reinforcement.[2-6] Most behavioral measures argued to reflect SR habits emphasize a lack of signs of cognitive influence. The SR associations are inferred from lack of evidence for purposeful or prospective behavior. For example, in one influential framework,[4,5] actions can become associated with expected outcomes (AO learning) through associative learning processes. This AO structure is demonstrated experimentally by showing that animals are sensitive to devaluation of the reward, for example, by pairing it with nauseogenic injections of lithium chloride. After tasting the reward as aversive in their home environment, subjects then will avoid the devalued goal when placed back in the task context, as though they had gained an aversive representation of the particular outcome for which they had previously worked and their behavior was guided by this negative outcome representation. With repeated experience in performing a behavior, or under particular task conditions, subjects can become insensitive to such devaluation procedures. Despite forming a lithium-induced aversion to the reward, animals will still work for it when in the task context. This insensitivity of behavior to the value of the outcome is suggested to reflect an underlying SR habit.[5,7] This framework also includes the criterion that habits are insensitive to changes in the contingency between an action and an outcome, for example, that habits are resistant to an omission schedule in which the action leads to reward cancellation.

The remarkable success of this framework is due in part to its utility in dissociating brain regions involved in AO versus SR behaviors. Studies on rodents and primates, including humans, have demonstrated that SR habits (exhibiting outcome independence) depend on brain structures including the dorsolateral striatum (DLS), dopaminergic neurons in the substantia nigra compacta (SNc), the infralimbic (IL) cortex, and the central nucleus of the amygdala (CeA).[8-14] By contrast, outcome-guided behaviors depend more on cognitive-associative circuits including the prelimbic (PL) cortex, orbitofrontal cortex (OFC), and dorsomedial striatum (DMS).[8-10,13,15,16]

Related work on stimulus-guided versus response-guided behaviors has uncovered similar brain networks for habits.[17] In this set of studies, based on maze navigation tasks, SR habits are inferred to exist when animals perform a set of learned actions rather than follow spatial cues in order to find rewards. In a plus-shaped maze, rats start from one arm and find food after turning into, for example, the right arm, and then the subjects are started from an opposing arm. Animals following an "egocentric" action plan will turn the same direction as they had previously (right in our example), whereas animals following a place strategy will follow the spatial cues to find where the food had been located.[17-20] In many conditions, animals initially start with a place strategy and then with training shift to a response strategy, taken as evidence of forming an SR habit.[17,19] Basal ganglia circuits, including the DLS and SNc dopaminergic neurons, are also implicated in the response strategy, as disruptions of their activity cause animals to favor a spatial strategy instead.[17,18,20-22]

Automaticity: action chunking and decline of deliberative behavior

Pioneering SR accounts of habit learning capture a great deal of the behavioral phenomena that arise as habits are formed in tasks, and certainly are valuable, yet the activity recorded in habit-related brain regions as habits are formed suggests that additional processes are at play. One dominant feature of neural activity in the basal ganglia is a pattern of activity that relates closely to how fluid and apparently nonpurposeful the behavior is, potentially by "chunking" the behavior to-

gether into a unit. Animals in a wide variety of tasks start with trial-and-error learning; under conditions in which task demands are stable, behavior becomes more rigid and consistent over the course of learning and practice. Several studies have characterized the neural correlates of this type of action automaticity in canonical habit-promoting brain regions, the DLS and IL cortex, and they find striking relationships to behavior and distinctions between these regions. Among these are a series of studies on rats running a T-shaped maze,[23-27] which we describe here. Rats wait at a starting gate, hear a warning cue, and then traverse the maze on opening of a gate. Part way through the run, the rats are exposed to one of two instructional cues (eg, auditory tones or tactile cues underfoot), instructing them to turn and enter the left or right T-arm. If they do it correctly, they receive a reward. If not, they receive nothing. Rats learn this task over weeks and reach an end-state of performing highly accurately and speedily. From training to overtraining, the rats also shift from being devaluation-sensitive (AO) to insensitive (SR).[26]

During this period of behavioral acquisition, cortico-basal ganglia circuits—long implicated in skill acquisition and habit formation—undergo changes in neural activity that map onto this shift into a relatively fixed running routine. For example, the predominant signal in the DLS that arises in medium spiny projection neurons as animals acquire the T-maze task is one in which the activity accentuates the boundaries of the maze runs. The majority of task-responsive neurons exhibit a burst of firing activity as the run is initiated or as the run is completed, or both, resulting in an ensemble representation of both the beginning and end of the run. Often there is an additional burst as the maze turn is completed. Non–task-related neurons become relatively quiet during the behavior. To the extent that this chunking pattern of activity within the DLS causally controls the habitual behaviors, which remains to be tested, habits may be encoded in the DLS by signals that help link the actions together into a chunk, with salient features being its initiation and termination,[28,29] just as working-memory processes can involve a chunking together of information (eg, phone numbers).[30] Both chunking and gathering together of elements of the entire sequence, called concatenation, can also be involved.[31]

Through a series of studies, we probed this DLS chunking pattern in relation to the behavior of the animals and in relation to which contingencies of the task are critical to its formation. One notable finding is that this pattern forms quite early in task learning, well before performance reaches asymptote and well before the behavior becomes insensitive to devaluation.[24,26] What this would suggest at face value is that the brain is built to favor more flexible decision-making processes as subjects learn task conditions, but somehow the habit system is nonetheless undergoing changes for later selection or dominance of the future habit. However, contrary to a view that the DLS is active but lacks influence over performance until later when a habit finally takes over, we uncovered a potential influence of this DLS chunking pattern on how deliberative a behavior is throughout essentially all stages of learning, both early (nonhabitual) and late (habitual). Dating back to the works of Tolman[32] and of Muenzinger,[33] researchers have recognized that animals display a sign of deliberative decision-making while performing maze tasks involving turn choices. Termed vicarious trial-and-error or deliberation, the behavior is seen as the temporary halting of a maze run with head turns toward possible maze arms before making a selection and turning.[33-35] These deliberations tend to be expressed often during trial-and-error learning and then decline to near zero levels as a behavior becomes well learned.[34] We found this transition also: animals deliberate on the majority of trials during early learning and then quit this behavioral sign of deliberation on most trials during overtraining on T-maze tasks.[26]

We found that the strength of the DLS chunking pattern correlates inversely with deliberations on a trial-by-trial basis: the stronger this pattern is, the less likely animals are to exhibit a deliberation during their run.[26] This correlation occurs during early learning phases as well, when animals are still devaluation-sensitive. Remarkably, the major DLS signal related to deliberations is the activity that occurs at the initiation of a maze run, and not activity that is present or absent during the deliberation itself. Thus, a strong burst of DLS activity as an animal begins its run correlates with a lower likelihood of a later deliberation, and weaker activity at run start correlates with more numerous instances of deliberation.[26]

This early DLS activity, like the late DLS activity near the end of runs, has parallels in the striatum and prefrontal cortex of primates as well.[36,37] For example, there is a similar relationship between action automaticity and the end-related signal in striatal activity.[37] In

one study, macaque monkeys were trained to perform a series of saccades to receive reward. The task involved many potential saccade sequences, and monkeys gradually formed stereotypic and efficient saccade strategies. Neurons in the striatum exhibited a clear chunking pattern of activity. The sharpness of the activity at the termination of the saccade sequences was highly correlated with the degree of stereotyped performance of the saccade sequences, and this activity encoded an integrated cost-outcome signal.[37] Thus, both the beginning and end activity in the sensorimotor striatum are closely related to how automatic and repetitive the performance of a given behavioral sequence is. These correlations suggest that the sensorimotor striatum carries a potentially active influence over behavior very early in the learning process, trial by trial, bestowing on behavior more automaticity the stronger the activity is as the behavior begins. On this point, in recent human neuroimaging work on decision-making processes for reward, a competition between cognitive and habit-like strategies has been shown to occur essentially at a trial-to-trial level, and even within sequential decision stages of a single trial.[38-41] Collectively, these findings support the idea that habits are not always an end-state of training, though that is when they may be most strongly expressed.

Importantly, this DLS chunking pattern appears to be relatively independent of many other aspects of behavior: whether the run is accurate (ie, rewarded) or not,[24-26] whether the run leads to a positive or negative outcome when experimenters manipulate the value of the reward,[26] and even whether animals encounter a sudden change in the identity of the instruction cue and must learn anew.[25] Under these conditions, we found the DLS chunking pattern to be stable. Furthermore, although DLS activity can be correlated with run speed, it can develop independence from speed, which is reduced sharply after task conditions are changed, despite no evidence that DLS activity changes concomitantly.[26] We noted this DLS stability when we, without warning, switched the cue identity from auditory to tactile in the task,[25] as well as when we exposed animals to a devalued reward for many sessions on the maze, allowing them to learn to avoid it.[26] The DLS pattern does, however, decay when all rewards lose their value,[26] when rewards are removed,[27] and when the contingency between the animals' acquired behavior and the acquired outcome expectancy is explicitly changed.[42] What these data indicate is that the DLS chunking pattern is probably operative in relation to executing a behavior automatically and nondeliberatively, that it tends to remain stable as long as familiar routines are performed and are at least partially reinforced, and that it may influence behavior essentially throughout the process of habit formation.

Such results fit with a large body of work on skill learning showing that the DLS is critical for motor skill acquisition and expression, suggesting that DLS activity may contribute to the stability and consistency of action repertoires. Whereas skills are thought to be a component of a habit, but distinct in many ways from what we regard as habits (ie, not always acquired through positive reinforcement), their structure nonetheless requires similar DLS-related circuits as do habits.[29] This is true even for fixed action patterns such as grooming in rodents.[43] Such similarity across types of repetitive behaviors raises the possibility that the DLS may in part be promoting the skill aspects of habits, or in other words, supporting them as sequences with structure and fluid expression.[31]

It will be of great interest to continue learning whether individually distinct types of neurons within the striatum (D_1- or D_2-receptor–expressing neurons; striosome or matrix projection neurons; different classes of interneurons) carry similar or different signals. Recently, Kubota et al[25] found that fast-spiking interneurons (putative γ-aminobutyric acid–mediated [GABAergic] interneurons) in the DLS also formed the begin-and-end chunking pattern in mice running a well-learned T-maze task. Moreover, when the modality of the instructional cue indicating which end-arm was baited was changed from auditory to tactile, these neurons developed a phasic, short-lived activity peak at the onset of the cue that was absent in the activity of DLS projection neurons. These results suggest that these interneurons function not only in maintaining action boundaries of the task, but also in registering task instruction changes to potentially aid behavioral flexibility. Work from the Costa laboratory has also evaluated the different signaling properties of dopamine D_1-receptor–expressing and D_2-receptor–expressing striatal projection neurons. Findings suggest that both types of neurons represent the onset of a well-learned action sequence (lever pressing), but that they may differently represent the step-wise progression of the actions as they are performed.[44,45] Recent work using a two-step nose-poke task in mice supports this view as well.[46] Finally, there

is strong evidence that D_2-receptor–expressing DLS neurons are critical for habit formation, based on use of the devaluation-insensitivity measure,[47-49] though D_1-receptor–expressing cells have not yet been exhaustively evaluated in these studies (but see ref 50).

Chunking activity elsewhere: distinct relations to habitual behavior

The chunking pattern is not unique to laboratory rats, but is also found in the DLS, and broader basal ganglia, of mouse and in corresponding regions in the striatum of macaque monkey prefrontal cortex and striatum during action sequences, in mouse SNc during action sequences, and in the HVC (formerly known as hyperstriatum ventrale, pars caudalis) of songbirds while singing.[36,37,45,51,52] Importantly, this pattern is not present in brain regions not thought to regulate habits from lesion or inactivation studies, including the DMS or PL cortex.[24,26] This growing body of data based on recordings of spike activity indicates that action chunking may be represented neurally across species, types of behaviors, and brain regions, and is a major—if not *the* major—way in which DLS represents habits.

Of note, this chunking pattern is also present in the superficial cortical layers of a medial prefrontal region known in rodents as the IL (infralimbic) cortex, which is also critical for habits but is not directly connected with the habit-related DLS.[10,14,26,53,54] Action chunking thus may be represented across multiple circuits simultaneously as habits are formed. However, the dynamics of the IL pattern are quite different from those of the DLS, suggesting a potentially distinct contribution of the IL cortex to habit formation. First, the beginning-and-end pattern forms late in the IL cortex, only as animals develop a consistency in their performance and an insensitivity to reward value during an overtraining period.[26] Second, also unlike the DLS, the IL pattern is not correlated on a trial-by-trial basis with deliberations, suggesting that the IL activity and deliberations may not be directly linked. Third, the IL pattern is exquisitely sensitive to changes in the task that require animals to change their behavior, whereas the DLS pattern is less sensitive.[26] Specifically, when we devalued one of the maze rewards, and animals changed their behavior to mainly running to the still-valued goal regardless of cueing, the IL pattern decayed rapidly. Then, as the animals rehearsed this new behavioral strategy over several weeks, the pattern reemerged as though to represent this new routine as a new habit.

We have extended this putative correlation to causal control by applying optogenetic manipulation of IL activity after genetically introducing light-sensitive proteins found in algae.[54,55] In our first study,[53] we found that inhibiting the IL activity only during maze runs after overtraining and reward devaluation immediately led animals to exhibit outcome sensitivity in conditions in which normal rats continued running for the devalued goal, by habit. Later, after 2 weeks of post-devaluation training during which the animals developed a new routine of always running to the still-valued goal, the same IL inhibition changed that behavior again: animals stopped performing this routine and instead reverted to their old habit of running when instructed to both devalued and valued goals. This set of findings suggests that the IL cortex operates as a strategy-scheduler of sorts, promoting newly acquired habits and behaviors at the expense of old ones that are being suppressed.[10,12,57] A functionally similar activity pattern has been found for blood-oxygenation-level–dependent (BOLD) activity in the inferior parietal cortex, suggesting that it could similarly help arbitrate between habitual and prospective cognitive processes.[39]

These findings pose an intriguing notion that parallel circuits exist for promoting habits—those rooted in the cortical-associative-limbic circuit (eg, IL cortex) and those in the basal ganglia (eg, DLS).[10] In this view, the IL cortex might promote habits by dampening or otherwise disrupting neural events related to prospection and flexibility in its target zones, including the DMS and the nucleus accumbens, or indirectly, interfacing with basal ganglia such as through connections with the CeA and onward to the SNc. In support of this possibility, Lingawi and Balleine[58] latter have shown that contralateral lesions to the anterior CeA and DLS suppress habit expression using the devaluation sensitivity measure, suggesting they interact for habits. It is possible that the IL connections with the amygdala facilitate this interaction.[58]

On this point, decision-making processes are supported by a range of brain circuits outside of the classic habit system, and deliberations themselves are correlated with interesting neural signals related to prospective cognitive processes in the OFC, hippocampus, and nucleus accumbens.[35,59,60] Among many remaining questions is whether habits involve a diminution of such

signals, or instead, involve accentuation of activity in regions like DLS and IL cortex that actively override these signals. The lack of deliberations when DLS activity is strong, and when animals have been overtrained on tasks in general, may support the former possibility—that deliberations are directly weakened as part of the habit formation process. In further support, activity in cognitive regions like the DMS declines simultaneously as animals are overtrained and DLS activity takes shape.[24]

An additional DLS role: outcome feedback

The above notion is not to say that this action automaticity and chunking is all that DLS does for habits; it is not. Other signals exist in the DLS during habit formation, with other relationships to behavior, further supporting the argument here that habits can be parsed into component processes. One appears, surprisingly, to be outcome feedback signaling.

Several laboratories have observed responses of striatal neurons to reward, including responses of neurons in the dorsal striatum.[61-63] Schmitzer-Torbert and Redish,[64] for example, found that a set of projection neurons in the dorsal striatum are engaged during maze runs, while another set of neurons are engaged only after the run is stopped and reward is being consumed. We observed these neurons in our T-maze task as well. In recent work, we found a population of neurons that entirely lack in-task responses, but that respond about a half of a second after the behavior is completed.[65] During learning, about half of these neurons tend to respond more after correct runs (during reward consumption), and the other half tend to respond after incorrect runs (when there is no reward). Though the population sizes of both of these subsets are similar during training, we find a striking shift during overtraining and habit formation: the number of neurons responding to errors after incorrect runs falls to near zero, whereas the number of neurons responding to rewards after correct runs increases proportionally. Thus, outcome signaling of errors is almost gone, but outcome signaling after correct responses remains strong. The lack of error responsivity as habits are acquired could contribute to a lack of error-corrective feedback that may render behaviors less sensitive to negative outcomes, while the maintained reward-related activity could help maintain habits from trial to trial, potentially signaling that rewards occurred as predicted. Moreover, the reward response appears to have a value component. When exposed to the maze after one reward is devalued, the response to the still-valued reward is greater than the response to the devalued reward when it is, on occasion, pursued. We highlight the fact that the temporal dynamics of the chunking-related and outcome-related neurons in the DLS are distinct: the bracketing pattern appearing early and the outcome signaling becoming strong later. Thus, the DLS appears to exhibit not only distinct signals for distinct aspects of habitual performance, but also distinct learning-related time courses when they form.

As noted, signals accentuating the beginning and end of saccadic eye movement sequences have also been found in recordings within the striatum and prefrontal neocortex of macaques.[36,37] Of special note in this work is that this bracketing pattern can be observed in self-trained monkeys as well as in monkeys trained on a cued saccade task and that the end peak includes an integrated cost-outcome signal that is highly correlated with the repetitiveness of the saccade sequences. It is likely that such signals exist also in rodents given the recognized homology of basal ganglia anatomy and function, underscoring the potential role of DLS in both task performance and outcome evaluation.

Revisiting the historical framework

The implication of this neural recording work is that habits—at least some habits—are not simply SR associations guiding a rat reflexively from point A to point B. Although the correlational nature of this work does warrant caution in such an interpretation, it raises the opportunity to consider behavioral characteristics of habits as not being limited to SR associations. By extension, brain regions associated with these characteristics (eg, prefrontal cortical region IL, and striatal region DLS) may not be required to encode an SR association as their principal contribution to habits.

Concerning the behavior measures themselves, lack of behavioral response to devaluation or contingency degradation is a negative result; SR is inferred when subjects do not exhibit goal-directed (AO) processes. In such conditions, we appreciate that evidence is strongly in favor of the brain site in question as being necessary for SR habits. However, other interpretations of insensitivity of behavior to outcome changes have been raised. These include an overly fixed knowledge of the learned

task conditions and routes to acquiring goals,[32] loss of associability of response-eliciting task stimuli due to reduced ability of the stimuli to call up information related to the perceptual details of the outcome,[6,64] a motivational attraction or value related to the action sequence itself,[12,66,67] and a level of motivation for reward that has become decoupled from the actual or perceived reward value.[68] SR behavior is similarly inferred in the maze studies by the fact that the animals follow a particular response routine rather than following external cues, a notion that has strong roots in research on response routines dating back over a century (eg, see ref 69), but that would meet with the same alternative interpretations. Thus, we argue the function of brain regions that promote these measures is a more open question than is often presumed.[10,12,70-72]

Let us take as an example the DLS, a canonical SR-learning system[8,13,17,73]: how do we reconcile its diverse neural signals with SR theory? The dominant task-bracketing pattern in DLS projection neuron activity is puzzling from an SR point of view. While promoting SR associations would conceivably override deliberative behaviors, it remains unclear why they would be manifested in DLS chunking as opposed to signals related to specific SR pairings, and particularly in the burst of DLS activity at action initiation and termination. Moreover, the stability of DLS activity in the face of major changes in the SR structure of the task, as noted above, suggests that this particular pattern may not reflect specific pairs of Ss and Rs. Although still a hypothesis for now, behavioral chunking may be one important underlying biological feature of a habit and could, itself, lead to outcome-insensitivity and response-based maze behaviors, thus effectively standing in as an SR association but dissociable from SR details.[12,28,74] In this view, chunking provides a structure to sequential behaviors, and as such, step one will be linked to step two, and so forth, leading to behavior that is focused on the next action step (or the whole sequence) and not on the final reward outcome.[28] Alternatively, the behavior may be focused on the major action events, such as start, turn, and stop, with fewer "expert neurons" responding in relation to other task events.[27] Models by the Balleine group raise the possibility that such processing might occur as a form of prospective behavior, with the target at a given time being the next action step.[75,76] Generally, for well-learned behaviors, the closer the rat gets to reward the more sensitive its behavior becomes to reward value,[6,13,77,78] which we ourselves observed in the T-maze task.[26] Thus, behaviorally and neurally, evidence suggests that action chunking can lead to habitual behaviors, with the initial action in the sequence carrying powerful influence over expression of the full habit and showing the greatest resistance to change when the action sequence is no longer a valued course of action.[26] This hypothesis leaves open the function of the outcome feedback signals that coexist in the DLS, which are novel enough to require further research before firm hypotheses can be made. Nevertheless, the change in their signaling during habit formation to favor correct over error outcomes is likely to be related to habit maintenance in important ways.

We also note that firing patterns reflective of SR associations have been difficult to demonstrate in recording studies focused on the striatum. Several studies have reported on activity in the striatum in rats performing SR tasks involving discrete stimuli paired with discrete responses; these studies do have the caveat that the cognitive versus habitual nature of performance is generally not assessed. For example, Stalnaker et al[79] observed that 20% of recorded neurons in the DLS fire during a certain response if it was preceded by a certain cue, which would seem to represent an SR association. However, the same proportion of neurons representing SR signals this way were found in the DMS, a region that is thought to oppose habits. Similarly, in our T-maze task, response-specific firing representations in projection neurons appear not to be different between the DMS and DLS.[24] Thorn et al[24] found that a similar 15% to 35% proportion of recorded neurons in the DMS and DLS exhibited preferential firing during one of the two T-maze turns. The activity of these neurons also did not predict the turn direction of the animal, nor did the proportion of these turn-specific neurons change over the course of training and habit formation. Such findings raise the possibility that the habit-promoting functions of DLS may not be expressed in these types of signals, or, if they are, that some process is required to promote their function in the DLS but not in the DMS as habits are formed. Moreover, studies have suggested that the DLS neurons lack responses to predictive stimuli when movement factors are ruled out.[80] If this lack of stimulus representation is true of most task conditions and species, it would suggest that the DLS represents the response (R-feature) that is somehow combined with the stimulus (S-feature) elsewhere

to form SR links. It is possible that SR associations are represented in other patterns of spike or oscillatory activity in these same brain regions, and that they are present in other brain regions or are compiled through circuit connections across areas. Yet, in all, it remains unknown whether manipulations to the DLS that have been shown to disrupt habits (eg, lesions, inactivations) are effective because they disrupt the DLS chunking activity, the DLS outcome feedback activity, both, or other potential signals (eg, from interneurons). If the DLS is inhomogeneous with respect to habit-related activity based on these different signaling processes, the hypothesis is that manipulations specific to those signals would produce different deficits in habit, for instance, a return of deliberative decision making and loss of action structure (blocking the chunking pattern) versus increased sensitivity to changes in specific rewards values or negative consequences (blocking reward-related activity). Identifying features of habits in relation to their neural correlate—in DLS, IL cortex, and elsewhere—will open up testable hypotheses such as these, which could prove useful in understanding the overall structure of a habit.

Implications for "disorders of habit"

Excessive and overly fixed behavioral routines are symptoms in many disorders, including addictions, obsessive-compulsive disorder (OCD), and autism-spectrum disorders. Links to dysfunctional corticostriatal circuits have been made for each of these.[29,81-88] For the most part, there is little consensus that habits are equivalent to or generative of symptoms in such disorders, though research has made progress in understanding the extent to which abnormally strong habits are part of the problem.

Addiction, for example, is a complex disorder involving changes in brain activity across hypothalamic, amygdalar, mesolimbic, cortical, and basal ganglia circuits. Different "failure modes," including potential failures in the motivational,[89] homeostatic,[90] and impulse-control systems,[91] can be thought of as different possible routes toward the same end-state of a compulsive, unhealthy behavioral pattern.[90,92,93] There is also strong evidence that addicted individuals and animal models of addiction exhibit habit-like tendencies in their drug-taking rituals and in their compulsive persistence in drug-taking in the presence of drug cues and drug seeking despite negative consequences.[72] These features are linked to the DLS and its dopamine input, in particular[72,92-94] with the thought that they reflect a failure mode of an overly strong SR drug-seeking habit.[92] However, the evidence above raises the possibility for different failure modes within the habit system itself as potentially contributing to such behavioral compulsion. These failures could include overly strong chunking-related activity in the DLS or IL cortex, loss of error-corrective signaling in the DLS, or inflexibility in the IL-related habit-promoting process. Each possibility remains tenable, we speculate, though none have been evaluated as yet.

OCD presents a challenging distinction in that the compulsive behaviors are thought to be driven more by negative reinforcement (avoiding a bad outcome, or an outcome perceived as bad) than by positive reinforcement. Yet, here too, the habit system is implicated.[83,93,97] For example, OCD sufferers working to avoid an aversive wrist shock would continue to do so more than controls even when they saw that the shock was "devalued" by the experimenter unplugging the electrical stimulator.[82] Corticostriatal connections have similarly been implicated in the compulsion behaviors in OCD, in human patients[98] and rodent models.[84,85] Such findings are important to consider in the context of related animal work showing that habits form more rapidly during or after a state of stress,[99-101] or in negative reinforcement conditions.[102] It remains to be seen whether, under such conditions, the DLS, IL cortex, or other habit-related regions of the brain have abnormal signaling, though as with addiction models, this is a testable possibility.

Conclusion

Findings from basic neuroscience research on habits are broadening our understanding of how habits arise from changes in neural activity in the brain. Our view is that the dynamics of activity we and others observe in key habit-promoting brain regions suggest that many reward-seeking habits could involve multiple signaling mechanisms in the brain. With further research into the casual roles of these signals, as well as work to uncover other signals that may exist in the wider habit-related brain circuitry, this possibility can be put to the test. At present, however, the available findings lead us to the view that habits are multifaceted, not simple SR behaviors, and that abnormal habits are possibly multifaceted as well. Classification of habits in terms of features rec-

ognizable in neural activity patterns should be useful as research efforts continue to wrestle with understanding the aspects of brain function that are distorted in cases of compulsive behavior. ❑

Acknowledgements: We thank Yasuo Kubota for comments on this manuscript. This work was supported by grants from the National Institutes of Health (R01 MH060379, R01 NS025529 and R01 EY012848 to AMG; F32 MH08545 to KSS), from the Whitehall Foundation 2014-05-77 (to KSS), and from the Nancy Lurie Marks Family Foundation (to AMG).

REFERENCES

1. James W. *The Principles of Psychology.* Vol 1. New York, NY: Cosimo; 1890.
2. Hull CL. *Principles of Behavior.* New York, NY: Appleton-Century Crofts; 1943.
3. Thorndike EL. *Animal Intelligence: an Experimental Study of the Associative Processes in Animals.* New York, NY: Macmillan; 1898.
4. Dickinson A. Actions and habits: the development of behavioral autonomy. *Philos Trans R Soc Lond B Biol Sci.* 1985;308(1135):67-78.
5. Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology.* 1998;37(4-5):407-419.
6. Holland PC. Cognitive versus stimulus-response theories of learning. *Learn Behav.* 2008;36(3):227-241.
7. Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Quart J Exp Psychol B.* 1982;34(2):77-98.
8. Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology.* 2010;35(1):48-69.
9. Yin HH, Ostlund SB, Balleine BW. Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur J Neurosci.* 2008;28(8):1437-1448.
10. Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex.* 2003;13(4):400-408.
11. Faure A, Haberland U, Conde F, El Massioui N. Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *J Neurosci.* 2005;25(11):2771-2780.
12. Smith KS, Graybiel AM. Investigating habits: strategies, technologies and models. *Front Behav Neurosci.* 2014;8:39.
13. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. *Nat Rev Neurosci.* 2006;7(6):464-476.
14. Hitchcott PK, Quinn JJ, Taylor JR. Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cereb Cortex.* 2007;17(12):2820-2827.
15. Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun.* 2013;4:2264.
16. Schoenbaum G, Roesch M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron.* 2005;47(5):633-636.
17. Packard MG. Exhumed from thought: basal ganglia and response learning in the plus-maze. *Behav Brain Res.* 2009;199(1):24-31.
18. Packard MG, McGaugh JL. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem.* 1996;65(1):65-72.
19. Tolman EC, Ritchie BF, Kalish D. Studies in spatial learning: V. Response learning versus place learning by the non-correction method. *J Exp Psychology.* 1947;37(4):285-292.
20. McDonald RJ, White NM. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav Neurosci.* 1993;107(1):3-22.
21. Wang LP, Li F, Wang D, Xie K, Shen X, Tsien JZ. NMDA Receptors in dopaminergic neurons are crucial for habit learning. *Neuron.* 2011;72(6):1055-1066.
22. Lee AS, Duman RS, Pittenger C. A double dissociation revealing bidirectional competition between striatum and hippocampus during learning. *Proc Natl Acad Sci U S A.* 2008;105(44):17163-17168.
23. Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. Building neural representations of habits. *Science.* 1999;286(5445):1745-1749.
24. Thorn CA, Atallah HE, Howe MW, Graybiel AM. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron.* 2010;66(5):781-795.
25. Kubota Y, Liu J, Hu D, et al. Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *J Neurophysiol.* 2009;102(4):2142-2160.
26. Smith KS, Graybiel AM. A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron.* 2013;79(2):361-374.
27. Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature.* 2005;437(7062):1158-1161.
28. Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem.* 1998;70(1-2):119-136.
29. Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci.* 2008;31:359-387.
30. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev.* 1956;101(2):343-352.
31. Graybiel AM, Grafton ST. The striatum: where skills and habits meet. *Cold Spring Harb Perspect Biol.* 2015;7(8):a021691.
32. Tolman EC. *Purposive Behavior in Animals and Men.* New York, NY: Century; 1932.
33. Muenzinger KF. Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *Pedagog Semin J Genetic Psychol.* 1938;53(1):75-86.
34. Tolman EC. Cognitive maps in rats and men. *Psychol Rev.* 1948;55(4):189-208.
35. van der Meer M, Kurth-Nelson Z, Redish AD. Information processing in decision-making systems. *Neuroscientist.* 2012;18(4):342-359.
36. Fujii N, Graybiel AM. Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science.* 2003;301(5637):1246-1249.
37. Desrochers TM, Amemori K, Graybiel AM. Habit learning by naive macaques is marked by response sharpening of striatal neurons representing the cost and outcome of acquired action sequences. *Neuron.* 2015;87(4):853-868.
38. Wunderlich K, Dayan P, Dolan RJ. Mapping value based planning and extensively trained choice in the human brain. *Nat Neurosci.* 2012;15(5):786-791.
39. Liljeholm M, Dunne S, O'Doherty JP. Differentiating neural systems mediating the acquisition vs. expression of goal-directed and habitual behavioral control. *Eur J Neurosci.* 2015;41(10):1358-1371.
40. Daw ND, Niv Y, Dayan P. Actions, policies, values, and the basal ganglia. In: Bezard E, ed. *Recent Breakthroughs in Basal Ganglia Research*. Hauppauge, NY: Nova Science Publishers; 2005:91-106.
41. Dolan RJ, Dayan P. Goals and habits in the brain. *Neuron.* 2013;80(2):312-325.
42. Regier PS, Amemiya S, Redish AD. Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. *J Neurophysiol.* 2015;114(3):1399-1416.
43. Aldridge JW, Berridge KC, Rosen AR. Basal ganglia neural mechanisms of natural movement sequences. *Can J Physiol Pharmacol.* 2004;82(8-9):732-739.
44. Cui G, Jun SB, Jin X, et al. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature.* 2013;494(7436):238-242.
45. Jin X, Tecuapetla F, Costa RM. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat Neurosci.* 2014;17(3):423-430.
46. Rothwell PE, Hayton SJ, Sun GL, Fuccillo MV, Lim BK, Malenka RC. Input- and output-specific regulation of serial order performance by corticostriatal circuits. *Neuron.* 2015;88(2):345-356.
47. Yu C, Gupta J, Chen JF, Yin HH. Genetic deletion of A_{2A} adenosine receptors in the striatum selectively impairs habit formation. *J Neurosci.* 2009;29(48):15100-15103.

Formación de hábitos

Los hábitos, tanto buenos como malos, están generalizados en la conducta animal. A través de estudios psicológicos y neurobiológicos se han desarrollado importantes sistemas para comprender los hábitos. Este trabajo nos entrega una rica comprensión acerca de las redes cerebrales que estimulan los hábitos, y también nos ayuda a comprender lo que constituye una conducta habitual en oposición a una conducta que es más flexible y eventual. A partir de estudios que utilizan métodos de registro neural hay evidencia creciente que sugiere que la formación de hábitos no es un proceso simple. Se revisa esta evidencia y se toma la posición que los hábitos podrían ser esculpidos a partir de múltiples cambios disociables en la actividad neural. Estos cambios ocurren a través de múltiples regiones cerebrales e incluso dentro de regiones cerebrales localizadas. Esta estrategia de clasificar los componentes de un hábito en base a diferentes señales cerebrales proporciona una nueva forma potencialmente útil de concebir los trastornos que implican conductas excesivamente establecidas como resultado de diferentes disfunciones potenciales en la red cerebral de hábitos.

La formation des habitudes

Les habitudes, les bonnes comme les mauvaises, sont omniprésentes dans le comportement animal. Grâce à des études psychologiques et neurobiologiques, d'importants cadres de travail ont été développés pour les comprendre. Ce travail permet une excellente compréhension des réseaux cérébraux qui peuvent favoriser ces habitudes et nous aide aussi à comprendre ce qui constitue un comportement habituel en opposition à un comportement plus flexible et plus prospectif. D'après un nombre croissant de données d'études cliniques utilisant des méthodes d'enregistrement neuronales, la formation des habitudes n'est pas un processus simple. Nous analysons ces données et supposons que les habitudes pourraient être façonnées à partir de nombreuses modifications dissociables de l'activité neuronale. Ces changements interviennent dans de multiples régions cérébrales et même dans des régions cérébrales localisées. Cette stratégie de classification des composants d'une habitude basée sur différents signaux cérébraux apporte une nouvelle façon potentiellement utile de concevoir les troubles qui impliquent des comportements excessivement ancrés comme résultant de dysfonctions potentielles différentes dans le réseau cérébral des habitudes.

48. Shan Q, Christie MJ, Balleine BW. Plasticity in striatopallidal projection neurons mediates the acquisition of habitual actions. *Eur J Neurosci.* 2015;42(4):2097-2104.
49. Corbit LH, Nie H, Janak PH. Habitual responding for alcohol depends upon both AMPA and D2 receptor signaling in the dorsolateral striatum. *Front Behav Neurosci.* 2014;8:301.
50. O'Hare JK, Ade KK, Sukharnikova T, et al. Pathway-specific striatal substrates for habitual behavior. *Neuron*. 2016;89(3):472-429.
51. Jin X, Costa RM. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature.* 2010;466(7305):457-462.
52. Fujimoto H, Hasegawa T, Watanabe D. Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci.* 2011;31(27):10023-10033.
53. Coutureau E, Killcross S. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav Brain Res.* 2003;146(1-2):167-174.
54. Smith KS, Virkud A, Deisseroth K, Graybiel AM. Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proc Natl Acad Sci U S A.* 2012;109(46):18932-18937.
55. Fenno L, Yizhar O, Deisseroth K. The development and application of optogenetics. *Ann Rev Neurosci.* 2011;34:389-412.
56. Bernstein JG, Boyden ES. Optogenetic tools for analyzing the neural circuits of behavior. *Trends Cogn Sci.* 2011;15(12):592-600.
57. Barker JM, Taylor JR, Chandler LJ. A unifying model of the role of the infralimbic cortex in extinction and habits. *Learn Mem.* 2014;21(9):441-448.
58. Lingawi NW, Balleine BW. Amygdala central nucleus interacts with dorsolateral striatum to regulate the acquisition of habits. *J Neurosci.* 2012;32(3):1073-1081.
59. Stott JJ, Redish AD. A functional difference in information processing between orbitofrontal cortex and ventral striatum during decision-making behaviour. *Philos Trans R Soc Lond B Biol Sci.* 2014;369(1655). doi:10.1098/rstb.2013.0472.
60. Johnson A, Redish AD. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci.* 2007;27(45):12176-12189.
61. Apicella P, Ljungberg T, Scarnati E, Schultz W. Responses to reward in monkey dorsal and ventral striatum. *Exp Brain Research.* 1991;85(3):491-500.
62. Cromwell HC, Hassani OK, Schultz W. Relative reward processing in primate striatum. *Exp Brain Res.* 2005;162(4):520-525.
63. Hikosaka O, Sakamoto M, Usui S. Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J Neurophysiol.* 1989;61(4):814-832.
64. Schmitzer-Torbert N, Redish AD. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *J Neurophysiol.* 2004;91(5):2259-2272.
65. Smith KS, Graybiel AM. Habit formation coincides with shifts in reinforcement representations in the sensorimotor striatum. *J Neurophysiol*. 2016. Jan 6. Epub ahead of print.
66. Holland PC, Wheeler DS. Representation-mediated food aversions. In: Reilly S, Schachtman T, eds. *Conditioned Taste Aversion: Behavioral and Neural Processes*. Oxford, UK: Oxford University Press; 2009:196-225.
67. Berridge KC. Wanting and liking: observations from the neuroscience and psychology laboratory. *Inquiry (Oslo).* 2009;52(4):378.
68. Glickman SE, Schiff BB. A biological theory of reinforcement. *Psychol Rev.* 1967;74(2):81-109.

69. Robinson TE, Berridge KC. Addiction. *Annu Rev Psychol.* 2003;54:25-53.
70. Carr H, Watson JB. Orientation in the white rat. *J Comp Neurol Psychol.* 1908;18(1):27-44.
71. de Wit S, Barker RA, Dickinson AD, Cools R. Habitual versus goal-directed action control in Parkinson disease. *J Cogn Neurosci.* 2011;23(5):1218-1229.
72. Desmurget M, Turner RS. Motor sequences and the basal ganglia: kinematics, not habits. *J Neurosci.* 2010;30(22):7685-7690.
73. Seger CA, Spiering BJ. A critical review of habit learning and the basal ganglia. *Front Syst Neurosci.* 2011;5:66.
74. Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci.* 2005;8(11):1481-1489.
75. Dezfouli A, Balleine BW. Habits, action sequences and reinforcement learning. *Eur J Neurosci.* 2012;35(7):1036-1051.
76. Dezfouli A, Lingawi NW, Balleine BW. Habits as action sequences: hierarchical action control and changes in outcome value. *Philos Trans R Soc Lond B Biol Sci.* 2014;369(1655). doi:10.1098/rstb.2013.0482
77. Nelson A, Killcross S. Amphetamine exposure enhances habit formation. *J Neurosci.* 2006;26(14):3805-3812.
78. Morgan MJ. Resistance to satiation. *Animal Behavior.* 1974;22:449-466.
79. Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G. Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. *Front Integr Neurosci.* 2010;4:12.
80. Root DH, Tang CC, Ma S, Pawlak AP, West MO. Absence of cue-evoked firing in rat dorsolateral striatum neurons. *Behav Brain Res.* 2010;211(1):23-32.
81. Everitt BJ, Belin D, Economidou D, Pelloux Y, Dalley JW, Robbins TW. Review. Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. *Philos Trans R Soc Lond B Biol Sci.* 2008;363(1507):3125-3135.
82. Gillan CM, Robbins TW. Goal-directed learning and obsessive-compulsive disorder. *Philos Trans R Soc Lond B Biol Sci.* 2014;369(1655). doi:10.1098/rstb.2013.0475.
83. Gillan CM, Apergis-Schoute AM, Morein-Zamir S, et al. Functional neuroimaging of avoidance habits in obsessive-compulsive disorder. *Am J Psychiatry.* 2015;172(3):284-293.
84. Ahmari SE, Spellman T, Douglass NL, et al. Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science.* 2013;340(6137):1234-1239.
85. Burguiere E, Monteiro P, Feng G, Graybiel AM. Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science.* 2013;340(6137):1243-1246.
86. Shepherd GM. Corticostriatal connectivity and its role in disease. *Nat Rev Neurosci.* 2013;14(4):278-291.
87. Di Martino A, Kelly C, Grzadzinski R, et al. Aberrant striatal functional connectivity in children with autism. *Biol Psychiatry.* 2011;69(9):847-856.
88. Gruner P, Anticevic A, Lee D, Pittenger C. Arbitration between action strategies in obsessive-compulsive disorder. *Neuroscientist.* 2015 Jan 20. Epub ahead of print.
89. Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev.* 1993;18(3):247-291.
90. Koob GF. Negative reinforcement in drug addiction: the darkness within. *Curr Opin Neurobiol.* 2013;23(4):559-563.
91. Kalivas PW, Volkow ND. The neural basis of addiction: a pathology of motivation and choice. *Am J Psychiatry.* 2005;162(8):1403-1413.
92. Redish AD, Jensen S, Johnson A. A unified framework for addiction: vulnerabilities in the decision process. *Behav Brain Sci.* 2008;31(4):415-437; discussion 437-487.
93. Voon V, Derbyshire K, Ruck C, et al. Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry.* 2015;20(3):345-352.
94. Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behav Brain Res.* 2009;199(1):89-102.
95. Willuhn I, Burgeno LM, Everitt BJ, Phillips PE. Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proc Natl Acad Sci U S A.*109(50):20703-20708.
96. Jonkman S, Pelloux Y, Everitt BJ. Differential roles of the dorsolateral and midlateral striatum in punished cocaine seeking. *J Neurosci.* 2012;32(13):4645-4650.
97. Gillan CM, Papmeyer M, Morein-Zamir S, et al. Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am J Psychiatry.* 2011;168(7):718-726.
98. Banca P, Voon V, Vestergaard MD, et al. Imbalance in habitual versus goal directed neural systems during symptom provocation in obsessive-compulsive disorder. *Brain.* 2015;138(pt 3):798-811.
99. Dias-Ferreira E, Sousa JC, Melo I, et al. Chronic stress causes frontostriatal reorganization and affects decision-making. *Science.* 2009;325(5940):621-625.
100. Braun S, Hauber W. Acute stressor effects on goal-directed action in rats. *Learn Mem.* 2013;20(12):700-709.
101. Schwabe L, Wolf OT. Stress prompts habit behavior in humans. *J Neurosci.* 2009;29(22):7191-7198.
102. Asem JS, Holland PC. Immediate response strategy and shift to place strategy in submerged T-maze. *Behav Neurosci.* 2013;127(6):854-859.