A Computational Model of Craving and Obsession

A. DAVID REDISH[a] AND ADAM JOHNSON[b]

[a]*Department of Neuroscience, University of Minnesota, Minneapolis, Minnesota, USA*

[b]*Graduate Program in Neuroscience and the Center for Cognitive Sciences, University of Minnesota, Minneapolis, Minnesota, USA*

ABSTRACT: If addictions and problematic behaviors arise from interactions between drugs, reward sequences, and natural learning sytems, then an explanation of clinically problematic conditions (such as the self-administration of drugs or problem gambling) requires an understanding of the neural systems that have evolved to allow an agent to make decisions. We hypothesize a unified decision-making system consisting of three components—a situation recognition system, a flexible, planning-capable system, and an inflexible, habit-like system. In this article, we present a model of the planning-capable system based on a planning process arising from experimentally observed look-ahead dynamics in the hippocampus enabling a forward search of possibilities and an evaluation process in the nucleus accumbens. Based on evidence that opioid signaling can provide hedonic evalutation of an achieved outcome, we hypothesize that similar opioid-signaling processes evaluate the value of expected outcomes. This leads to a model of craving, based on the recognition of a path to a high-value outcome, and obsession, based on a value-induced limitation of the search process. This theory can explain why opioid antagonists reduce both hedonic responses and craving.

KEYWORDS: craving; obsession; addiction; hippocampus; nucleus accumbens; opioid signaling; opiates; dopamine

INTRODUCTION

We start from the assumption that neural systems have evolved to allow an agent to make decisions that will allow it to survive and procreate. This means that if we want to understand action-selection processes that lead to clinically problematic situations, such as self-administration of drugs[1–3] or the continued pursuit of problematic behaviors such as gambling,[4–7] we need to

Address for correspondence: A. David Redish, Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA. Voice: 612-626-3738; fax: 651-626-5009.
redish@ahc.umn.edu

Ann. N.Y. Acad. Sci. 1104: 324–339 (2007).
doi: 10.1196/annals.1390.014

first understand that natural learning system and how those addictive processes access it. Making optimal decisions requires calculations of the *expected utility* or *value* of taking specific actions in specific situations. Expected utility (or value) can be defined as the expected reward, taking into account the expected magnitude of the reward, the expected probability of receiving the reward, and the expected delay before receiving that reward.[8,9]

To predict the expected reward and the appropriate action to achieve that reward, the agent must first recognize the situation it is in. This recognition process is fundamentally a classification problem—this situation is like these and not like those. For example, if one is deciding whether or not to put a dollar in a soda machine, to predict the consequences of putting the dollar in and pushing the soda button, one needs to correctly recognize that one is in front of a soda machine, not in front of a bank ATM. In the psychology literature, this is referred to as accessing the correct *schema*. Importantly, one needs to recognize not only the general soda machine schema, but also to determine whether there are any specific situation cues available. For example, is this a Coke or a Pepsi machine? Is this machine more or less reliable than other machines?

Once one has identified the situation one is in, calculating the value of the actions to be taken requires some combination of cached memory and search of the possibilities.[10] In the computer science literature, this has been termed *depth of search*, and is a fundamental basis of heuristic reasoning. Interestingly, there is strong behavioral evidence that there are two systems in the mammalian brain with differing levels of search: (1) a flexible system, which is capable of being learned quickly, but is computationally expensive to use, and (2) an inflexible system, which can act quickly, but must be learned slowly. The flexible system allows the planning of multiple paths to achieve a goal and takes the expectation of that goal into account in its decision making. In contrast, the inflexible system simply retrieves the remembered action for a given situation.[11–14] The flexible system can be learned quickly because of its flexibility—knowing the existence of a potential path to a goal does not commit one to taking that path. However, the complexity of planning through those potential paths makes the flexible system computationally expensive. In contrast, the inflexible system must be learned slowly because it would be dangerous to commit to always taking an action in a situation until one knows that that action is the correct one. However, the limited search done in the inflexible (habit) system allows it to work quickly requiring only limited computational resources.

The existence of these two systems has been proposed in both the animal navigation (*cognitive map* vs. *route* strategies,[11,12] *place* vs. *response* strategies[13]) and learning theory literatures (*situation–outcome* (*S–O*) vs. *situation–action* (*S–A*) associations[14,15]).

In the navigation literature, the interaction of multiple navigation systems can be seen in how rats solve the classic single-T maze task.[13,16–19] Limited

training leads to a place strategy in which animals return to the same goal location when started from multiple starting points, even through this may require different actions. In contrast, extended training leads to a response strategy in which animals perform the same actions on entering the maze, even if that leads them to different goals. The place strategy depends on the integrity of the hippocampus and ventromedial striatum, whereas the response strategy depends on the integrity of the dorsolateral striatum.[13,18,19]

In the learning theory literature, the interaction of multiple learning systems can be seen in how rats respond to devaluation.[14,15,20] Classically, these differences are measured by first training an animal to take an action sequence leading to reward, and then, changing the value of the reward to the animal, usually in a different context. The value of a reward can be changed by providing excess amounts of the reward (satiation[14]) or by pairing the reward with an aversive stimulus, such as LiCl (devaluation[20,21]). Finally, the animal is provided the chance to take the action. If the action selection process takes into account the current value of the reward, then the animal will not respond, but if the action selection process is an association between the situation and the action (thus does not take into account the value of the reward), the animal will continue to respond. With extended training of a reliable association, animals switch from a devaluation-capable system to a devaluation-independent system.[14,22] The devaluation-capable system (S–O) is dependent on the integrity of the ventral striatum,[23,24] the prelimbic medial prefrontal cortex,[22] and the orbitofrontal cortex,[20,25] whereas the devaluation-independent system (S–A) is dependent on the integrity of the dorsal striatum[19,26,27] and the infralimbic cortex.[22,28]

This leads us to hypothesize a unified system incorporating three subsystems, a situation recognition system and two contrasting decision systems—a flexible, planning-capable system that accommodates multiple paths to goals and takes into account the value of potential outcomes, and an inflexible, habit-like system, which reacts with a single action to each situation and does not take into account the value of potential outcomes (see FIG. 1).

Both the planning-capable and habit-like systems require a recognition of the agent's situation. This recognition system entails a categorization process, which is likely to arise in cortical systems through competitive learning,[29–32] using content-addressable memory mechanisms.[33–35]

The first (flexible, planning) decision-making system requires recognition of a situation S, recognition of a means of achieving outcome O from situation S,

$$S \cdots \overset{(a)}{\rightarrow} O \tag{1}$$

as well as the evaluation of the value of achieving outcome O, which will depend on the agent's current needs N

$$E(V) = V(E(O), N) \tag{2}$$

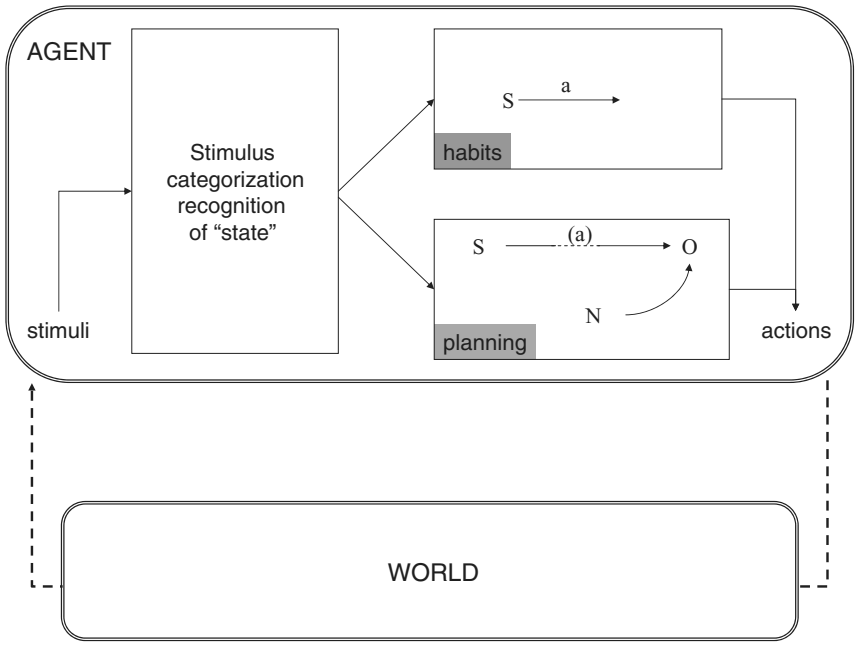

FIGURE 1. Three systems involved in decision making.

where $E(V)$ is the expectation of the value of taking action a in situation S, which is a function of the expected outcome* $E(O)$ and the needs of the agent N. Because the value of the outcome is calculated on the fly (online), that calculation can take into account the needs (N) of the agent.

The second (inflexible, habit) decision-making system entails a simple association between situation and action. Thus, the habit system requires recognition of a situation S, and a single, identified action to take within that situation. We describe this system with the simple formulation

$$S \xrightarrow{a} \tag{3}$$

Evaluation in the second (inflexible, habit) system entails a memory recall of the learned associated (cached[10]) value of taking action a in situation S,

$$E(V) = V(S, a) \tag{4}$$

most likely learned through temporal-difference reinforcement learning mechanisms.[8] These two systems, along with a situation-recognition component (S), form a unified theory of decision-making processes.

*Computationally, the outcome O is simply a future state S', but we refer to it as the "outcome" to emphasize the importance of the "completion of needs" that the new situation $O = S'$ will achieve.

While the $S \xrightarrow{a}$ system has been well modeled through the TDRL algorithms,[9,36–39] the mechanisms that underlie the $S \cdots \overset{(a)}{\rightarrow} O$ system are more controversial.[40] Following a recent suggestion that one possible difference between $S \xrightarrow{a}$ and $S \cdots \overset{(a)}{\rightarrow} O$ systems is the depth of search,[10] we propose a model of the $S \cdots \overset{(a)}{\rightarrow} O$ system based on a *consideration of possibilities* signal provided by the hippocampus.

A COMPUTATIONAL MODEL OF THE DEVALUATION-CAPABLE/MAP-NAVIGATION SYSTEM

The key to both devaluation and map navigation is the ability to consider the possible consequences of one's actions. This hypothesized mechanism needs three components: (1) a recognition of the situation at hand (S), (2) a process by which the system can calculate the expected consequences of taking available actions (retrieval of the $S \cdots \overset{(a)}{\rightarrow} O$ relationship), and (3) evaluation of the expected outcome ($E(V) = V(E(O), N)$). We hypothesize that the planning component is instantiated through hippocampal dynamics and the evaluation component is instantiated through processing in orbitofrontal cortex and through opioid signaling in the nucleus accumbens.

Planning

We have recently observed look-ahead dynamics in the hippocampal neural ensemble recordings of rats facing a high-cost choice.[41,42] Briefly, rats were trained to run a choice task in which they made choices to receive food. Rats, particularly early in the session, paused at high-cost choices and showed behavior reminiscent of vicarious trial and error (VTE[43–45]). Because of the spatial tuning of hippocampal pyramidal cells,[12,46] it is possible to reconstruct the position of the animal x from the firing pattern F using Bayesian reconstruction techniques.[47–50] The reconstructed distribution $P(x|F)$ tracked the animal well as the animal ran through the central path. When the animal paused, the reconstructed distribution moved out along one choice, and then the other, alternating a few times before the rat began moving again.

It is not known what effect these nonlocal planning signals seen in the hippocampus have on downstream structures, but it is known that other hippocampal processes representing nonlocal information (i.e., sharp wave ripple complexes occurring during slow-wave sleep in which replay of recent experiences is known to occur[51–54]) do have effects on downstream structures, such as nucleus accumbens.[55] Thus, it is likely that these representations could also be translated downstream, providing a potential planning signal (recognition of a potential $S \cdots \overset{(a)}{\rightarrow} O$ path) for decision making.

Alternative Structures Involved in Planning

Historically, planning and expectation of outcome have been associated with prefrontal structures,[56–58] and Daw *et al.*[10] have suggested the prefrontal cortex as the site of the $S \cdots \overset{(a)}{\rightarrow} O$ process. Hippocampus projects to medial prefrontal cortex,[59] and prefrontal structures have been observed to contain goal-related processes.[60,61] The relative roles of hippocampus and prefrontal cortex in planning remain to be elucidated.

In the motor control fields, the cerebellum has been hypothesized to be the site of "forward models" predicting the consequences of one's actions.[62–64] While the cerebellum has been identified in cognitive processes as well as motor,[65] the processes controlled by the cerebellum tend to be those with tightly controlled timing, likely controlled by highly specialized cerebellar circuits,[62,63,66,67] which once learned become inflexible. In contrast, the planning processes addressed above require flexible circuits capable of evaluating consequences over variable and longer time periods. Neither devaluation nor map navigation have been found to be dependent on cerebellar integrity.

EVALUATION

To evaluate the value of an outcome, the system needs a signal that recognizes hedonic value. Two structures that have been suggested to be involved in the evalutation of an outcome are the orbitofrontal cortex[20,68–72] and the ventral striatum.[9,73–76] Neurons in the ventral striatum show reward correlates,[75,77–81] and anticipate predicted reward.[77,78,82,83] The hippocampus projects to ventral striatum,[84–86] and ventral striatal firing patterns reflect hippocampal neural activity.[55,87] Neurons in the orbitofrontal cortex encode parameters relating to the value of potential choices.[68,69] Both fMRI,[70,71] and lesion[20,56,88,89] data have also implicated the orbitofrontal cortex in the evaluation of value. Anticipatory neural firing of goal-related information in orbitofrontal cortex is dependent on hippocampal integrity.[90]

Berridge and Robinson[91,92] suggest that hedonic signals ("liking") are carried by opioid signaling, as evidenced by the effect of opioid agonists and antagonists on taste reactivity. Consistent with these ideas, Levine and colleagues[93,94] report that opioid antagonists directly interfere with the reported qualia of hedonic pleasure associated with eating sweet, without interfering in taste discrimination.

There are multiple opioid receptor types in the mammalian brain (μ, κ, δ,[95–97]). Whereas μ-receptor agonists are rewarding, euphorigenic, and support self-administration, κ-receptor agonists are aversive, dysphoric, and interfere with self-administration.[95–103†] μ-receptor antagonists block

†The role of δ receptors is more controversial.[96,102,104]

self-administration and conditioned approach to drug cues, but blocking the other opioid receptors (κ, δ) do not.[95,102,103] Each receptor type is associated with a preferential endogenous opioid signaling peptide (μ: β-endorphin, the endomorphins; κ: dynorphin, δ: the enkephalins).[96,99,105] These data suggest that the opioid system is well situated to provide a direct evaluation of an event: rewarding signals via μ receptors and aversive signals via κ receptors.

It is important to differentiate hedonic rewards and costs from reinforcement and aversion.[37,91,92,106] Reinforcement and aversion entail changes reflecting changes in expectation (i.e., the *value prediction error* term in temporal difference learning[8,9]). If one correctly predicts the hedonic pleasure provided by a reward, then one's value prediction error signal is zero, even though one presumably still feels that hedonic pleasure on achieving the reward. While euphoria and dysphoria have been associated with opioid signals,[95–103] reinforcement signals have been associated with dopamine.[9,91,92,106,107]

If endogenous opioids signal the actual hedonic evaluation of an achieved outcome, then when faced with potential outcome signals arriving from the hippocampus, one might expect similar proccesses to evaluate the value of expected outcomes. This predicts that the effect of hippocampal planning signals on accumbens structures will be to trigger evaluative processes similar to those that occur in response to actual achieved outcomes. This has immediate consequences for craving and obsession.

IMPLICATIONS

Craving

Craving is the intense desire for something. It is, fundamentally, a subjective, internal feeling, and may or may not always be reflected in external actions. In the terminology presented above, craving is the recognition that there is a pathway to a high-value outcome. This expectation can only occur in the $S \cdots\overset{(a)}{\rightarrow} O$ (planning) system; the $S \xrightarrow{a}$ (habit) system does not include a recognition of the expected outcome. Because the flexible (planning) system only entails the recognition that an action can lead to a potential path to a goal and does not entail a commitment to action, craving is not necessarily going to produce action selection. In the $S \cdots\overset{(a)}{\rightarrow} O$ planning system, when the forward planning (hippocampal) component reaches a goal that is evaluated to have a high value, this will produce a strong desire to achieve that goal. We suggest that the psychological effect of that recognition is to produce "craving."

Obsession

It is important to remember that the forward search component of the $S \cdots\overset{(a)}{\rightarrow} O$ system requires a memory retrieval process. This search process

entails the exploration of multiple consequences from situation S. Oversensitization of a single $S \cdots\overset{(a)}{\rightarrow} O$ relationship is likely to limit the exploration of possibilities, which would appear as a cognitive blinding to alternatives. Sensitization of an $S \cdots\overset{(a)}{\rightarrow} O$ relation would also mean that when an animal is returned to situation S, it is more likely to remember that it can reach outcome O, which would make it more likely to remember the existence of outcome O, thus more likely to experience craving in situation S. Craving would then lead to a recurring search of the same $S \cdots\overset{(a)}{\rightarrow} O$ path, which would appear as cognitive blinding or obsession.

DISCUSSION: PREDICTIONS AND OPEN QUESTIONS

In this article, we have proposed a model of craving based on a computational theory of planning processes,[10] which we have suggested arise from an interaction between a consideration-of-possibilities process involving hippocampus and an evaluative process involving nucleus accumbens or orbitofrontal cortex. Essentially, this produces an outcome-expectancy[108,109] model of craving:[110,112] craving entails recognition that there is a means of achieving a highly charged positive outcome (or of relieving a highly charged negative outcome). This model is consistent with new interpretations of Pavlovian conditioning as a memory-of-expectations process.[113] This process is fundamentally an associative memory process in that it requires the memory that there is a path to outcome O from situation S. Thus, it suggests that craving should involve structures involved in memory, particularly working memory, such as frontal cortex[114,115] and hippocampus.[12,116] Craving should also involve structures involved in the evaluation of future rewards, such as orbitofrontal cortex (OFC)[68,117] and nucleus accumbens.[118–120] Evidence from cue-induced craving responses in addicts supports these hypotheses.[112,121–123] The theory also provides immediate explanations for why opioid antagonists can be used to block craving, and makes predictions about a hippocampal role in devaluation.

Competitive opioid antagonists have been used clinically to reduce craving.[98,124–126] The model of the planning system laid out above provides an immediate explanation for this effect: when the predictive component of the planning system identifies the completion of an $S \cdots\overset{(a)}{\rightarrow} O$ pathway and a potential means of achieving an outcome, the evaluative component will release reward signals (i.e., endogenous opioids), identifying the value of that outcome for evaluative purposes. As noted above, the identification of a pathway to high reward leads to craving for that reward. The hypothesis that reward signals are released on recognition of a pathway to a high-value outcome implies that blocking those reward signals would not only dampen the subjective hedonic value of receiving reward, but would also dampen craving for those rewards. If that reward signal is based on opioid signaling, then this may

explain why opioid antagonists such as naltrexone or nalmefene can reduce craving.

Addiction has been proposed to entail a transition from exploratory use, to (in some users) the development of strong desires (craving), followed in some users by a strong, habitual use in which the user loses control of the drug use.[127–130] This sequence follows the sequence of normal learning. Flexible, map-based, devaluation-capable strategies are learned first;[12,13,17] but with repeatable, regular experience, animals switch to automated, inflexible, route-based, devaluation-resistant strategies.[12–14,17,28,131] In animals, drug-seeking also first involves more ventromedial aspects of striatum[132,133] and later involves the more dorsolateral aspects.[133,134] This theory predicts that drug addiction should progress through a flexible strategy based on intense craving to an inflexible, habit-based strategy, which is independent of craving.

This unified hypothesis leads to important open questions and predictions. An important, but as yet unresolved question is: How well does the map/route differentiation in the navigation literature[11,12] translate to the devaluation/nondevaluation distinction?[14] In the navigation literature, the key difference between map- and route-based strategies is flexibility. Map strategies are highly flexible, allowing paths around obstacles,[11,12] and journeys to the same location from different starting points.[13,16] In contrast, route strategies are highly inflexible, requiring the same paths under the same conditions.[11–13,135] In early maze experiments, overtrained rats were found to run full speed into novel obstacles[136,137] or off shortened tracks.[138] In the devaluation literature, the key difference lies in the inclusion of the outcome in action selection. $S{-}O$ strategies entail a consideration of the outcome, while $S{-}A$ strategies do not. Anatomically, map learning is critically dependent on the hippocampus.[11,12] However, Corbit and Balleine[139] found that hippocampal lesions had no effect on devaluation. Importantly, these lesions were partial and occurred before training. Similarly sized partial lesions that occurred before training have little or no effect on place finding in the Morris water maze,[140] which is the classic hippocampal-dependent navigation task. Ostlund and Balleine[141] report that hippocampal lesions after training devastate devaluation learning, as it does place finding in the Morris water maze.[140,142]

The crucial test of this hypothesis, however, is the prediction that similar opiate signaling will occur in response to both veridical inputs (reflecting real receipt of reward/punishment, leading to euphoria/dysphoria) and to hypothetical inputs (reflecting planning, leading to craving/dread). These hypotheses could be tested with simultaneous recordings of hippocampus and ventral striatum.

ACKNOWLEDGMENTS

This work was supported by a Career Development Award from the University of Minnesota TTURC (to ADR, NCI/NIDA P50 DA01333), by a graduate

fellowship from the Center for Cognitive Sciences at the University of Minnesota (to AJ, T32HD007151), by a Fulbright scholarship (to AJ), as well as by NIMH R01-MH06829 and by the Land Grant Professorship program at the University of Minnesota. We thank Daniel Smith, Carolyn Fairbanks, Jadin Jackson, Zeb Kurth-Nelson, Suck-Won Kim, Steve Jensen, and Paul Schrater for helpful discussions.

REFERENCES

1. VOLKOW, N. & T.-K. LI. 2005. The neuroscience of addiction. Nature Neurosci. **8:** 1429–1430.
2. KALIVAS, P.W. & N.D. VOLKOW. 2005. The neural basis of addiction: a pathology of motivation and choice. Am. J. Psychiatry **162:** 1403–1413.
3. O'BRIEN, C.P., N. VOLKOW & T.-K. LI. 2006. What's in a word? Addiction versus dependence in DSM-V. Am. J. Psychiatry **163:** 764–765.
4. CUSTER, R.L. 1984. Profile of the pathological gambler. J. Clin. Psychiatry **45:** 35–38.
5. WAGENAAR, W.A. 1988. Paradoxes of Gambling Behaviour. Hillsdale, NJ.
6. PETRY, N.M. 2006. Should the scope of addictive behaviors be broadened to include pathological gambling? Addiction **101:** 152–159.
7. POTENZA, M.N. 2006. Should addictive disorders include non-substance-related conditions? Addiction **101:** 142–151.
8. SUTTON, R.S. & A.G. BARTO. 1998. Reinforcement Learning: An Introduction. MIT Press. Cambridge, MA.
9. DAW, N.D. 2003. Reinforcement learning models of the dopamine system and their behavioral implications. Ph.D. thesis, Carnegie Mellon University. Pittsburgh, PA.
10. DAW, N.D., Y. NIV & P. DAYAN. 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neurosci. **8:** 1704–1711.
11. O'KEEFE, J. & L. NADEL. 1978. The Hippocampus as a Cognitive Map. Clarendon Press. Oxford.
12. REDISH, A.D. 1999. Beyond the Cognitive Map: From Place Cells to Episodic Memory. MIT Press. Cambridge, MA.
13. PACKARD, M.G. & J.L. MCGAUGH. 1996. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiol. Learn. Mem. **65:** 65–72.
14. BALLEINE, B.W. & A. DICKINSON. 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology **37:** 407–419.
15. DICKINSON, A. 1980. Contemporary Animal Learning Theory. Cambridge University Press. New York.
16. TOLMAN, E.C., B.F. RITCHIE & D. KALISH. 1946. Studies in spatial learning. II. Place learning versus response learning. J. Exp. Psychol. **36:** 221–229.
17. RESTLE, F. 1957. Discrimination of cues in mazes: a resolution of the 'place-vs-response' question. Psychol. Rev. **64:** 217–228.
18. BARNES, C.A., L. NADEL & W.K. HONIG. 1980. Spatial memory deficit in senescent rats. Can. J. Psychol. **34:** 29–39.

19. YIN, H.H. & B.J. KNOWLTON. 2004. Contributions of striatal subregions to place and response learning. Learn. Mem. **11:** 459–463.
20. SCHOENBAUM, G., M. ROESCH & T.A. STALNAKER. 2006. Orbitofrontal cortex, decision making, and drug addiction. Trends Neurosci. **29:** 116–124.
21. NELSON, A. & S. KILLCROSS. 2006. Amphetamine exposure enhances habit formation. J. Neurosci. **26:** 3805–3812.
22. KILLCROSS, S. & E. COUTUREAU. 2003. Coordination of actions and habits in the medial prefrontral cortex of rats. Cerebral Cortex **13:** 400–408.
23. SCHOENBAUM, G., T.A. STALNAKER & M.R. ROESCH. 2006. Ventral striatum fails to represent bad outcomes after cocaine exposure. Soc. Neurosci. Abstr. Program No. 485.16. 2006 Neuroscience Meeting Planner. Atlanta, GA: Society for Neuroscience, 2006. Online.
24. CORBIT, L.H., J.L. MUIR & B.W. BALLEINE. 2001. The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. J. Neurosci. **21:** 3251–3260.
25. SCHOENBAUM, G. *et al*. 2006. Encoding changes in orbitofrontal cortex in reversal-impaired aged rats. J. Neurophysiol **95:** 1509–1517.
26. YIN, H.H., B. KNOWLTON & B.W. BALLEINE. 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur. J. Neurosci. **19:** 181–189.
27. YIN, H.H., B.J. KNOWLTON & B.W. BALLEINE. 2006. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav. Brain Res. **166:** 189–196.
28. COUTUREAU, E. & S. KILLCROSS. 2003. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. Behav. Brain Res. **146:** 167–174.
29. GROSSBERG, S. 1976. Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. Biol. Cyber. **23:** 121–134.
30. RUMELHART, D.E. & J.L. MCCLELLAND, Eds. 1986. PDP: Explorations in the Microstructures of Cognition. Vol. 1. Foundations. MIT Press. Cambridge, MA.
31. ARBIB, M., Ed. 1995. The Handbook of Brain Theory and Neural Networks. MIT Press. Cambridge, MA.
32. REDISH, A.D. 2005. Implications of the temporal difference reinforcement learning model for addiction and relapse. Neuropsychopharmacology **30:** S27–S28.
33. HEBB, D.O. 1949. The Organization of Behavior. Wiley, New York. Reissued 2002. LEA.
34. HOPFIELD, J.J. 1982. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA **79:** 2554–2558.
35. HERTZ, J., A. KROGH & R.G. PALMER. 1991. Introduction to the Theory of Neural Computation. Addison-Wesley. Reading, MA.
36. MONTAGUE, P. R., P. DAYAN & T.J. SEJNOWSKI. 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. **16:** 1936–1947.
37. REDISH, A.D. 2004. Addiction as a computational process gone awry. Science **306:** 1944–1947.
38. DOYA, K. 2000. Reinforcement learning in continuous time and space. Neur. Comput. **12:** 219–245.
39. DAW, N.D., A.C. COURVILLE & D.S. TOURETZKY. 2006. Representation and timing in theories of the dopamine system. Neur. Comput. **18:** 1637–1677.

40. DAYAN, P. & B.W. BALLEINE. 2002. Reward, motivation, and reinforcement learning. Neuron **36:** 285–298.
41. JOHNSON, A. & A.D. REDISH. 2005. Observation of transient neural dynamics in the rodent hippocampus during behavior of a sequential decision task using predictive filter methods. Acta Neurobiol. Exp. **65:** 103.
42. JOHNSON, A. & A.D. REDISH. 2006. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point: a possible mechanism for the consideration of alternatives. Program No. 574.2. 2006 Neuroscience Meeting Planner. Atlanta, GA: Society for Neuroscience, 2006. Online.
43. MEUNZINGER, K.F. 1938. Vicarious trial and error at a point of choice I. a general survey of its relation to learning efficiency. J. Genet. Psychol. **53:** 75–86.
44. TOLMAN, E.C. 1939. Prediction of vicarious trial and error by means of the schematic sowbug. Psychol. Rev. **46:** 318–336.
45. HU, D. & A. AMSEL. 1995. A simple test of the vicarious trial-and-error hypothesis of hippocampal function. Proc. Natl. Acad. Sci. **92:** 5506–5509.
46. O'KEEFE, J. 1976. Place units in the hippocampus of the freely moving rat. Exp. Neurol. **51:** 78–109.
47. RIEKE, F. *et al*. 1997. Spikes. MIT Press. Cambridge, MA.
48. ZHANG, K. *et al*. 1998. Interpreting neuronal population activity by reconstruction: Unified framework with application to hippocampal place cells. J. Neurophysiol. **79:** 1017–1044.
49. BROWN, E.N. *et al*. 1998. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. J. Neurosci. **18:** 7411–7425.
50. JOHNSON, A., J. JACKSON & A.D. REDISH. In press. Measuring distributed properties of neural representations beyond the decoding of local variables—implications for cognition. *In* Mechanisms of Information Processing in the Brain: Encoding of Information in Neural Populations and Networks. C. HÖLSCHER & M.H.J. MUNK, Eds. Cambridge University Press. Cambridge, UK.
51. WILSON, M.A. & B.L. MCNAUGHTON. 1994. Reactivation of hippocampal ensemble memories during sleep. Science **265:** 676–679.
52. KUDRIMOTI, H.S., C.A. BARNES & B.L. MCNAUGHTON. 1999. Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. J. Neurosci. **19:** 4090–4101.
53. NÁDASDY, Z. *et al*. 1999. Replay and time compression of recurring spike sequences in the hippocampus. J. Neurosci. **19:** 9497–9507.
54. LEE, A.K. & M.A. WILSON. 2002. Memory of sequential experience in the hippocampus during slow wave sleep. Neuron **36:** 1183–1194.
55. PENNARTZ, C.M.A. *et al*. 2004. The ventral striatum in off-line processing: ensemble reactivation during sleep and modulation by hippocampal ripples. J. Neurosci. **24:** 6446–6456.
56. BECHARA, A. 2005. Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. Nature Neurosci. **8:** 1458–1463.
57. ZERMATTEN, A. *et al*. 2005. Impulsivity and decision making. J. Nerv. Ment. Dis. **193:** 647–650.
58. OWEN, A.M. 1997. Cognitive planning in humans: neuropsychological, neuroanatomical and neuropharmacological perspectives. Progr. Neurobiol. **53:** 431–450.
59. JONES, M.W. & M.A. WILSON. 2005. Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. PLOS Biol. **3:** e402.

60. JUNG, M.W. *et al*. 1998. Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. Cerebral Cortex **8:** 437–450.
61. HOK, V. *et al*. 2005. Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. Proc. Natl. Acad. Sci. USA **102:** 4602–4607.
62. HIKOSAKA, O. *et al*. 1998. Differential roles of the frontal cortex, basal ganglia, and cerebellum in visuomotor sequence learning. Neurobiol. Learn. Mem. **70:** 137–149.
63. DOYA, K. 1999. What are the computations of the cerebellum, the basal ganglia, and the cerebral cortex? Neur. Networks **12:** 961–974.
64. MIALL, R.C. 1998. The cerebellum, predictive control and motor coordination. Novartis Found. Symp. **218:** 272–284.
65. SEIDLER, R.D. *et al*. 2002. Cerebellum activation associated with performance change but not motor learning. Science **296:** 2043–2046.
66. DOYON, J. *et al*. 1998. Role of the striatum, cerebellum and frontal lobes in the automatization of a repeated visuomotor sequence of movements. Neuropsychologia **36:** 625–641.
67. ITO, M. 2000. Mechanisms of motor learning in the cerebellum. Brain Res. **886:** 237–245.
68. PADOA-SCHIOPPA, C. & J.A. ASSAD. 2006. Neurons in the orbitofrontal cortex encode economic value. Nature **441:** 223–226.
69. SCHOENBAUM, G. & M. ROESCH. 2005. Orbitofrontal cortex, associative learning, and expectancies. Neuron **47:** 633–636.
70. VOLKOW, N.D. & J.S. FOWLER. 2000. Addiction, a disease of compulsion and drive: Involvement of the orbitofrontal cortex. Cerebral Cortex **10:** 318–325.
71. O'DOHERTY, J. *et al*. 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nature Neurosci. **4:** 95–102.
72. FEIERSTEIN, C.E. *et al*. 2006. Representation of spatial goals in rat orbitofrontal cortex. Neuron **60:** 495–507.
73. MOGENSON, G.J., D.L. JONES & C.Y. YIM. 1980. From motivation to action: Functional interface between the limbic system and the motor system. Progr. Neurobiol. **14:** 69–97.
74. MOGENSON, G.J. & M. NIELSEN. 1984. Neuropharmacological evidence to suggest that the nucleus accumbens and subpallidal region contribute to exploratory locomotion. Behav. Neural Biol. **42:** 52–60.
75. LAVOIE, A.M. & S.J.Y. MIZUMORI. 1994. Spatial-, movement- and reward-sensitive discharge by medial ventral striatum neurons in rats. Brain Res. **638:** 157–168.
76. O'DOHERTY, J. *et al*. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science **304:** 452–454.
77. MARTIN, P.D. & T. ONO. 2000. Effects of reward anticipation, reward presentation, and spatial parameters on the firing of single neurons recorded in the subiculum and nucleus accumbens of freely moving rats. Behav. Brain Res. **116:** 23–38.
78. MIYAZAKI, K. *et al*. 1998. Reward-quality dependent anticipation in rat nucleus accumbens. NeuroReport **9:** 3943–3948.
79. CARELLI, R.M., S.G. IJAMES & A.J. CRUMLING. 2000. Evidence that separate neural circuits in the nucleus accumbens encode cocaine versus "natural" (water and food) reward. J. Neurosci. **20:** 4255–4266.
80. CARELLI, R.M. 2002. Nucleus accumbens cell firing during goal-directed behaviors for cocaine vs. 'natural' reinforcement. Physiol. Behav. **76:** 379–387.

81. CARELLI, R.M. & J. WONDOLOWSKI. 2003. Selective encoding of cocaine versus natural rewards by nucleus accumbens neurons is not related to chronic drug exposure. J. Neurosci. **23:** 11214–11223.
82. YUN, I.A. *et al*. 2004. The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues. J. Neurosci. **24:** 2923–2933.
83. SCHULTZ, W. *et al*. 1992. Neuronal activity in monkey ventral striatum related to the expectation of reward. J. Neurosci. **12:** 4595–4610.
84. MCGEORGE, A.J. & R.L. FAULL. 1989. The organization of the projection from the cerebral cortex to the striatum in the rat. Neuroscience **29:** 503–537.
85. FINCH, D.M. 1996. Neurophysiology of converging synaptic inputs from rat prefrontal cortex, amygdala, midline thalamus, and hippocampal formation onto single neurons of the caudate/putamen and nucleus accumbens. Hippocampus **6:** 495–512.
86. SWANSON, L.W. 2000. Cerebral hemisphere regulation of motivated behavior. Brain Res. **886:** 113–164.
87. MARTIN, P.D. 2001. Locomotion towards a goal alters the synchronous firing of neurons recorded simultaneously in the subiculum and nucleus accumbens of rats. Behav. Brain Res. **124:** 19–28.
88. SCHOENBAUM, G., A.A. CHIBA & M. GALLAGHER. 1999. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J. Neurosci. **19:** 1876–1884.
89. CHUDASAMA, Y. *et al*. 2003. Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: differential effects on selectivity, impulsivity and compulsivity. Behav. Brain Res. **146:** 105–119.
90. DAVIS, J.B. *et al*. 2006. Hippocampal dependence of anticipatory neuronal firing in the orbitofrontal cortex of rats learning an odor-sequence memory task. Soc. Neurosci. Abstr. Program No. 66.7. 2006 Neuroscience Meeting Planner. Atlanta. GA: Society for Neuroscience, 2006. Online.
91. BERRIDGE, K.C. & T.E. ROBINSON. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Rev. **28:** 309–369.
92. BERRIDGE, K.C. & T.E. ROBINSON. 2003. Parsing reward. Trends Neurosci. **26:** 507–513.
93. ARBISI, P.A., C.J. BILLINGTON & A.S. LEVINE. 1999. The effect of naltrexone on taste detection and recognition threshold. Appetite **32:** 241–249.
94. LEVINE, A.S. & C.J. BILLINGTON. 2004. Opioids as agents of reward-related feeding: a consideration of the evidence. Physiol. Behav. **82:** 57–61.
95. DE VRIES, T.J. & T.S. SHIPPENBERG, 2002. Neural systems underlying opiate addiction. J. Neurosci. **22:** 3321–3325.
96. HERZ, A. 1997. Endogenous opioid systems and alcohol addiction. Psychopharmacology **129:** 99–111.
97. HERZ, A. 1998. Opioid reward mechanisms: a key role in drug abuse? Can. J. Physiol. Pharmacol. **76:** 252–258.
98. MEYER, R. & S. MIRIN. 1979. The Heroin Stimulus. Plenum. New York.
99. CHAVKIN, C., I.F. JAMES & A. GOLDSTEIN. 1982. Dynorphin is a specific endogenous ligand of the kappa opioid receptor. Science **215:** 413–415.
100. MUCHA, R.F. & A. HERZ. 1985. Motivational properties of kappa and mu opioid receptor agonists studied with place and taste preference conditioning. Psychopharmacology **86:** 274–280.

101. BALS-KUBIK, R., A. HERZ & T. SHIPPENBERG. 1989. Evidence that the aversive effects of opioid antagonists and κ-agonists are centrally mediated. Psychopharmacology **98:** 203–206.
102. MATTHES, H.W.D. *et al*. 1996. Loss of morphine-induced analgesia, reward effect, and withdrawal symptoms in mice lacking the μ-opioid-receptor gene. Nature **383:** 819–823.
103. KIEFFER, B.L. 1999. Opioids: first lessons from knockout mice. Trends Pharmacol. Sci. **20:** 19–26.
104. BROOM, D.C. *et al*. 2002. Nonpeptidic δ-opioid receptor agonists reduce immobility in the forced swim assay in rats. Neuropsychopharmacology **26:** 744–755.
105. ZADINA, J.E. *et al*. 1997. A potent and selective endogenous agonist for the μ-opiate receptor. Nature **386:** 499–502.
106. SCHULTZ, W. 2002. Getting formal with dopamine and reward. Neuron **36:** 241–263.
107. MONTAGUE, P.R. *et al*. 1995. Bee foraging in uncertain environments using predictive hebbian learning. Nature **377:** 725–728.
108. TOLMAN, E.C. 1948. Cognitive maps in rats and men. Psychol. Rev. **55:** 189–208.
109. BOLLES, R.C. 1972. Reinforcement, expectancy, and learning. Psychol. Rev. **79:** 394–409.
110. MARLATT, G.A. 1985. Cognitive factors in the relapse process. *In* Relapse Prevention. G.A. MARLATT & J.R. GORDON, Eds.: 128–200. Guilford. New York.
111. GOLDMAN, M.S., S.A. BROWN & B.A. CHRISTIANSEN. 1987. Expectancy theory: Thinking about drinking. *In* Psychological Theories of Drinking and Alcoholism. H.T. BLAINE & K.E. LEONARD, Eds.: 181–226. Guilford. New York.
112. TIFFANY, S.T. 1999. Cognitive concepts of craving. Alcohol Res. Health **23:** 215–224.
113. RESCORLA, R.A. 1988. Pavlovian conditioning: It's not what you think it is. Am. Psychol. **43:** 151–160.
114. GOLDMAN-RAKIC, P.S., S. FUNAHASHI & C.J. BRUCE. 1990. Neocortical memory circuits. Cold Spring Harbor Symposia on Quant. Biol. LV: 1025–1038.
115. FUSTER, J.M. 1997. The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe, 3rd ed. Lippincot-Raven. Philadelphia, PA.
116. OLTON, D.S. & R.J. SAMUELSON. 1976. Remembrance of places passed: spatial memory in rats. J. Exp. Psych.: Anim. Behav. Processes **2:** 97–116.
117. ROESCH, M.R., A.R. TAYLOR & G. SCHOENBAUM. 2006. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron **51:** 509–520.
118. KALIVAS, P.W., N. VOLKOW & J. SEAMANS. 2005. Unmanageable motivation in addiction: A pathology in prefrontal-accumbens glutamate transmission. Neuron **45:** 647–650.
119. ANAGNOSTARAS, S.G., T. SCHALLERT & T.E. ROBINSON. 2002. Memory processes governing amphetamine-induced psychomotor sensitization. Neuropsychopharmacology **26:** 703–715.
120. LI, Y., M.J. ACERBO & T.E. ROBINSON. 2004. The induction of behavioural sensitization is associated with cocaine-induced structural plasticity in the core (but not shell) of the nucleus accumbens. Eur. J. Neurosci. **20:** 1647–1654.
121. CHILDRESS, A.R. *et al*. 1993. Cue reactivity and cue reactivity interventions in drug dependence. NIDA Res. Monogr. **137:** 73–94.
122. GRANT, S. *et al*. 1996. Activation of memory circuits during cue-elicited cocaine craving. Proc. Natl. Acad. Sci. **93:** 12040–12045.

123. HOMMER, D.W. 1999. Functional imaging of craving. Alcohol Res. Health **23:** 187–196.
124. KIEFER, F. & K. MANN. 2005. New achievements and pharmacotherapeutic approaches in the treatment of alcohol dependence. Eur. J. of Pharmacol. **526:** 163–171.
125. O'BRIEN, C.P. 2005. Anticraving medications for relapse prevention: a possible new class of psychoactive medications. Am. J. Psychiatry **162:** 1423–1431.
126. GRANT, J.E. *et al*. 2006. Multicenter investigation of the opioid antagonist nalmefene in the treatment of pathological gambling. Am. J. Psychiatry **163:** 303–312.
127. ALTMAN, J. *et al*. 1996. The biological, social and clinical bases of drug addiction: commentary and debate. Psychopharmacology **125:** 285–345.
128. LOWINSON, J.H. *et al*., Eds. 1997. Substance Abuse: A Comprehensive Textbook 3rd ed. Williams and Wilkins. Baltimore.
129. ROBBINS, T.W. & B.J. EVERITT. 1999. Drug addiction: bad habits add up. Nature **398:** 567–570.
130. EVERITT, B.J. & T.W. ROBBINS. 2005. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nature Neurosci. **8:** 1481–1489. Corrected online after print.
131. SCHMITZER-TORBERT, N.C. & A.D. REDISH. 2002. Development of path stereotypy in a single day in rats on a multiple-T maze. Archives Italiennes de Biologie **140:** 295–301.
132. ITO, R. *et al*. 2000. Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. J. Neurosci. **20:** 7489–7495.
133. LETCHWORTH, S.R. *et al*. 2001. Progression of changes in dopamine transporter binding site density as a result of cocaine self-administration in rhesus monkeys. J. Neurosci. **21:** 2799–2807.
134. ITO, R. *et al*. 2002. Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. J. Neurosci. **22:** 6247–6253.
135. REDISH, A.D. & D.S. TOURETZKY. 1998. The role of the hippocampus in solving the Morris water maze. Neural Computation **10:** 73–111.
136. WATSON, J.B. 1907. Kinaesthetic and organic sensations: their role in the reactions of the white rat to the maze. Psychol. Rev. **8:** 43–100.
137. CARR, H. & J.B. WATSON. 1908. Orientation in the white rat. J. Comp. Neurol. Psychol. **18:** 27–44.
138. DENNIS, W. 1932. Multiple visual discrimination in the block elevated maze. J. Comp. Physiol. Psychol. **13:** 391–396.
139. CORBIT, L.H. & B.W. BALLEINE. 2000. The role of the hippocampus in instrumental conditioning. J. Neurosci. **20:** 4233–4239.
140. MOSER, M.-B. & E.I. MOSER. 1998. Distributed encoding and retrieval of spatial memory in the hippocampus. J. Neurosci. **18:** 7535–7542.
141. OSTLUND, S.B. & B.W. BALLEINE. 2004. Post-training neurotoxic lesions of the dorsal hippocampus disrupt goal-directed actions. Soc. Neurosci. Abstr. Program no. 897.17. 2004 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2004. Online.
142. MORRIS, R.G.M. 1981. Spatial localization does not require the presence of local cues. Learn. Motiv. **12:** 239–260.