



High on habits

Monica R. F. Hilário and Rui M. Costa*

Section on In Vivo Neural Function, Laboratory for Integrative Neuroscience, NIAAA, NIH, Bethesda, MD, USA

The neural circuits involved in learning and executing goal-directed actions, which are governed by action-outcome contingencies and sensitive to changes in the expected value of the outcome, have been shown to be different from those mediating habits, which are less dependent on action-outcome relations and changes in outcome value. Extended training, different reinforcement schedules, and substances of abuse have been shown to induce a shift from goal-directed performance to habitual performance. This shift can be beneficial in everyday life, but can also lead to loss of voluntary control and compulsive behavior, namely during drug seeking in addiction. Although the brain circuits underlying habit formation are becoming clearer, the molecular mechanisms underlying habit formation are still not understood. Here, we review a recent study where Hilario et al. (2007) established behavioral procedures to investigate habit formation in mice in order to investigate the molecular mechanisms underlying habit formation. Using those procedures, and a combination of genetic and pharmacological tools, the authors showed that endocannabinoid signaling is critical for habit formation.

Keywords: striatum, endocannabinoids, dopamine, goal-directed, habits

Edited by:

Sidney A. Simon, Duke University, USA

Reviewed by:

Barry Setlow, Texas A&M University, USA

Albino J. Oliveira-Maia, Duke University Medical Center, USA; Universidade do Porto, Portugal

*** Correspondence:**



Rui Costa received his D.V.M. from the Technical University of Lisbon. He performed his Ph.D. with Dr. Alcino Silva at UCLA in the field of learning and memory, investigating the molecular and cellular mechanisms underlying the learning disabilities in NF1. After postdoctoral work with Dr. Miguel Nicolelis at Duke University where he established multi-site neuronal recordings in awake behaving mice, he became Section Chief at NIH in 2006, and Adjunct of the Champalimaud Neuroscience Program at IGC in 2007. His laboratory studies the neurobiology of action in health and disease. costarui@mail.nih.gov

INTRODUCTION

Goal-directed actions allow us to respond in an efficient way to changing situations. However, the continuous control and attention they demand can result in an unnecessary expenditure of our cognitive resources and can thus be prejudicial in some situations. In situations where the behavior is repeated regularly for a long time without major changes in the incentive value of the outcome, or situations where we cannot manipulate the probability of obtaining an outcome irrespective of the strategy employed, rules and habits can be advantageous. However, habitual behavior when taken to an extreme is associated with loss of control and with maladaptive behavior, such as drug seeking in addiction or compulsivity. Therefore, understanding the molecular and circuit mechanisms underlying habit formation can be important to prevent or treat these disorders.

It is known that different cortico-basal ganglia circuits support the learning and execution of goal-directed actions and habits (Balleine and Dickinson, 1998; Corbit and Balleine, 2003;

Killcross and Coutureau, 2003; Yin and Knowlton, 2004; Yin et al., 2004, 2005a,b, 2006). However, much less is known about the molecular bases of habit formation. In an effort to identify the molecular substrates involved in habit formation Hilario et al. (2007) tailored behavioral paradigms to study goal-directed actions and habit formation in mice. They confirmed that different schedules of reinforcement bias mice towards goal-directed actions or habits by using devaluation by sensory-specific satiety to test for habitual behavior (Hilario et al., 2007). They also introduced a novel assay that measures **generalization of actions** to novel manipulanda similar to those where animals were trained. Using these paradigms, they investigated the role of endocannabinoid signaling through CB1 receptors in habit formation, by employing both genetic and pharmacological tools (Hilario et al., 2007).

Here we review the study by Hilario et al. (2007), starting by defining the concepts of goal-directed and habitual behavior that were used in that study, and the foundations for the experi-

Action generalization

In the context of the task described, and in contrast to discrimination, a similar action is performed on a novel manipulandum which was never paired with the outcome.

Goal-directed action

The performance of a goal-directed action is dependent upon its consequences. Operationally, their execution is sensitive to changes in the incentive value of the outcome, and to the contingency between the execution of the action and getting the outcome.

Habit

Habits are insensitive to changes in outcome value and to the contingency between action and outcome.

It is usually entertained that the performance of habits is elicited by antecedent stimuli or situations, rather than by the expectancy of the consequences of the behavior.

Devaluation

Procedure to reduce the value of the outcome for an animal. It can be achieved in several manners, including by conditioned taste aversion, or by sensory-specific satiety.

mental design adopted by the authors. We also discuss the details of the behavioral paradigms adapted to study habit formation in mice. Finally, we explore the rationale behind the hypothesis involving endocannabinoids in habit formation, and the data indicating that endocannabinoid signaling through CB1 receptors is necessary for habit formation.

GOAL-DIRECTED ACTIONS AND HABITS

The study of how we learn actions and what drives them has been the focus of neuroscience and behavioral science for some time. However, the field has struggled not only with the identification of the circuits and the cellular and molecular bases supporting actions, but also with the definitions of what **goal-directed actions** and **habits** are (or if they differ at all). For the most part of last century learned actions were reduced to a stimulus-response (S-R) relation, and learning was perceived as a consequence of the continuous strengthening or weakening of the S-R relation by the use of reinforcements (Hull, 1943). Even though researchers like Tolman (1948, 1949) proposed that animals could use information they learned flexibly and use cognitive maps, and Von Holst proposed alternatives to the dominant view of behavior as a chain of reflexes emanating from Sherrington's work (Creed et al., 1932; Von Holst, 1973), for a long time behaviorists relied mostly on observational methods, and excluded intentionality, expectation or internal representation of the value of the outcome because they were considered subjective variables. However, in the later part of the 20th century Dickinson and Rescorla developed experimental tools to investigate if instrumental behavior was being performed because of its consequences or not (Adams, 1982; Adams and Dickinson, 1981; Colwill and Rescorla, 1985).

To investigate if actions were habitual (governed by a S-R relation) or goal-directed they asked if the actions were dependent on the expected value of the outcome, by introducing a **devaluation** test (Adams, 1982; Adams and Dickinson, 1981; Colwill and Rescorla, 1985). In this devaluation test rats were trained using an operant box to get access to food rewards, and after training the expected value of the reinforcements was manipulated by decreasing the value of the food (typically food poisoning). By comparing the number of responses when the food was devalued versus when it was not, they were able to distinguish experimentally habits as behavior impervious to devaluation, and goal-directed actions as sensitive to devaluation. Another test used to investigate if actions were goal-directed examined whether

the behavior was dependent on the contingency between the performance of the action and earning the outcome (Corbit et al., 2002; Dickinson et al., 1996; Hammond, 1980). Briefly, if the contingency between one of the actions and the outcome was decreased (degraded), rats would decrease the performance of that action specifically. These studies established that goal-directed actions are sensitive to changes in the expected value of the outcome and the contingency between the action and the outcome (A-O); while habitual behaviors are insensitive to changes in outcome value and contingency between action and outcome, suggesting they are governed by S-R relations (Balleine and Dickinson, 1994). These were the definitions adopted by Hilario et al. (2007) in their study.

Adams and Dickinson noticed not only that overtraining on a particular schedule could produce a transition from goal-directed behavior to habits, but also that different schedules of reinforcement differentially predisposed for habit formation (Adams, 1982; Adams and Dickinson, 1981). Specifically, the use of random ratio training schedules produced goal-directed behavior in rats, while the use of random interval schedules promoted habitual behavior (Adams and Dickinson, 1981; Dickinson, 1985; Dickinson et al., 1983). Balleine and Dickinson (1998) later used these procedures to start to examine the neural circuits involved in goal-directed behavior and habits. These behavioral assays have been very useful to investigate the neural circuits and the cellular and molecular mechanisms involved in goal-directed actions and habits (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Corbit et al., 2003; Coutureau and Killcross, 2003; Faure et al., 2005; Hilario et al., 2007; Nelson and Killcross, 2006; Yin et al., 2004, 2005b).

INVESTIGATING GOAL-DIRECTED ACTIONS AND HABITS IN MICE

Genetically engineered mice can be very useful to investigate the role of specific genes in a particular behavior, and to visualize or manipulate the circuits involved in that behavior. To investigate the molecular mechanisms of habit formation Hilario et al. (2007) adapted the experimental procedures previously used in rats (Adams, 1982; Adams and Dickinson, 1981; Corbit and Balleine, 2003), and developed new ones in mice (Hilario et al., 2007). Using an operant box where a particular action could be performed to obtain a specific outcome, they trained mice with two reinforcers: either regular "chow" pellets or sucrose. One reinforcer was delivered in the operant chamber contingent upon lever pressing (the outcome of the action of lever pressing), and the other reinforcer was presented

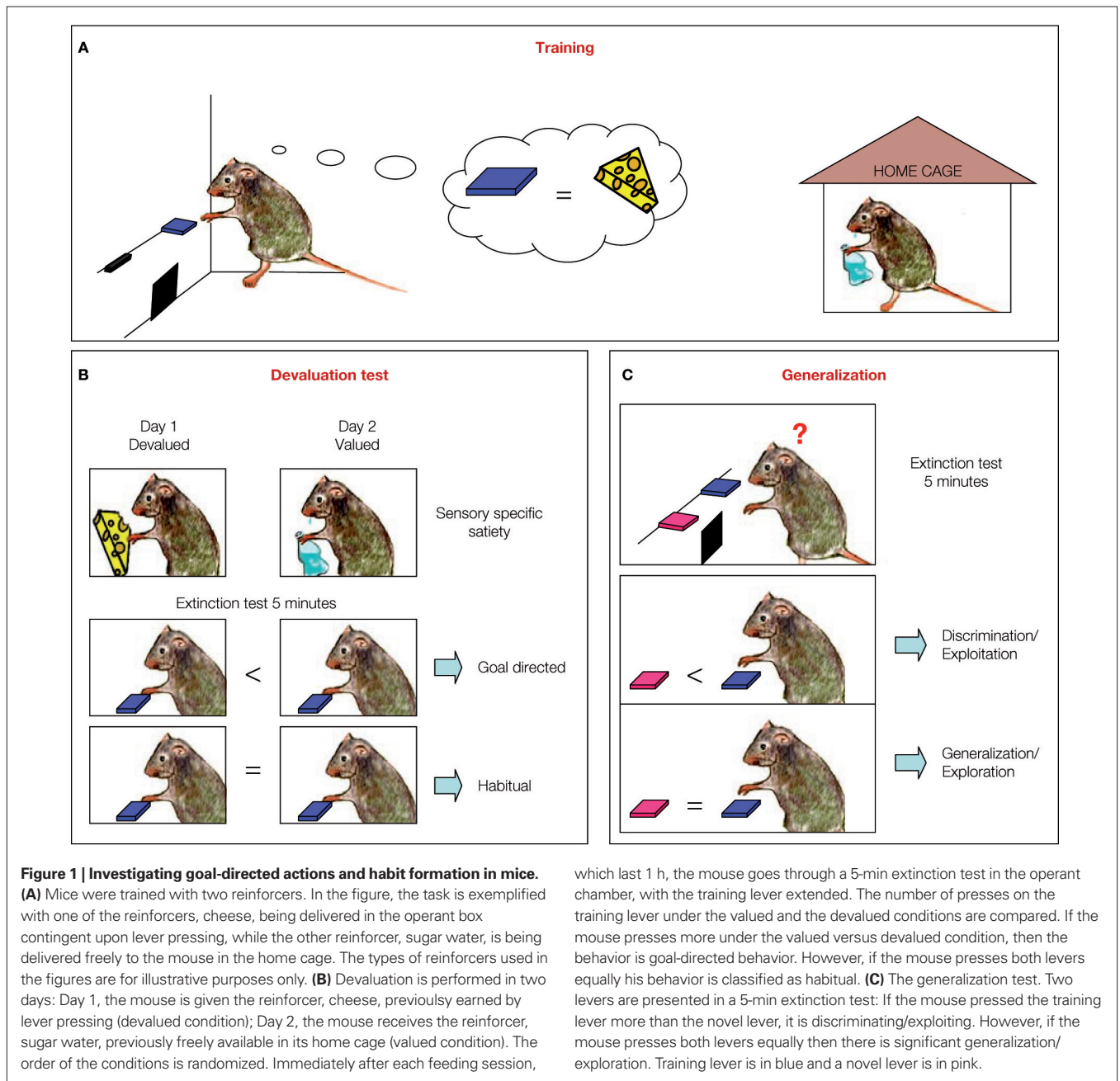
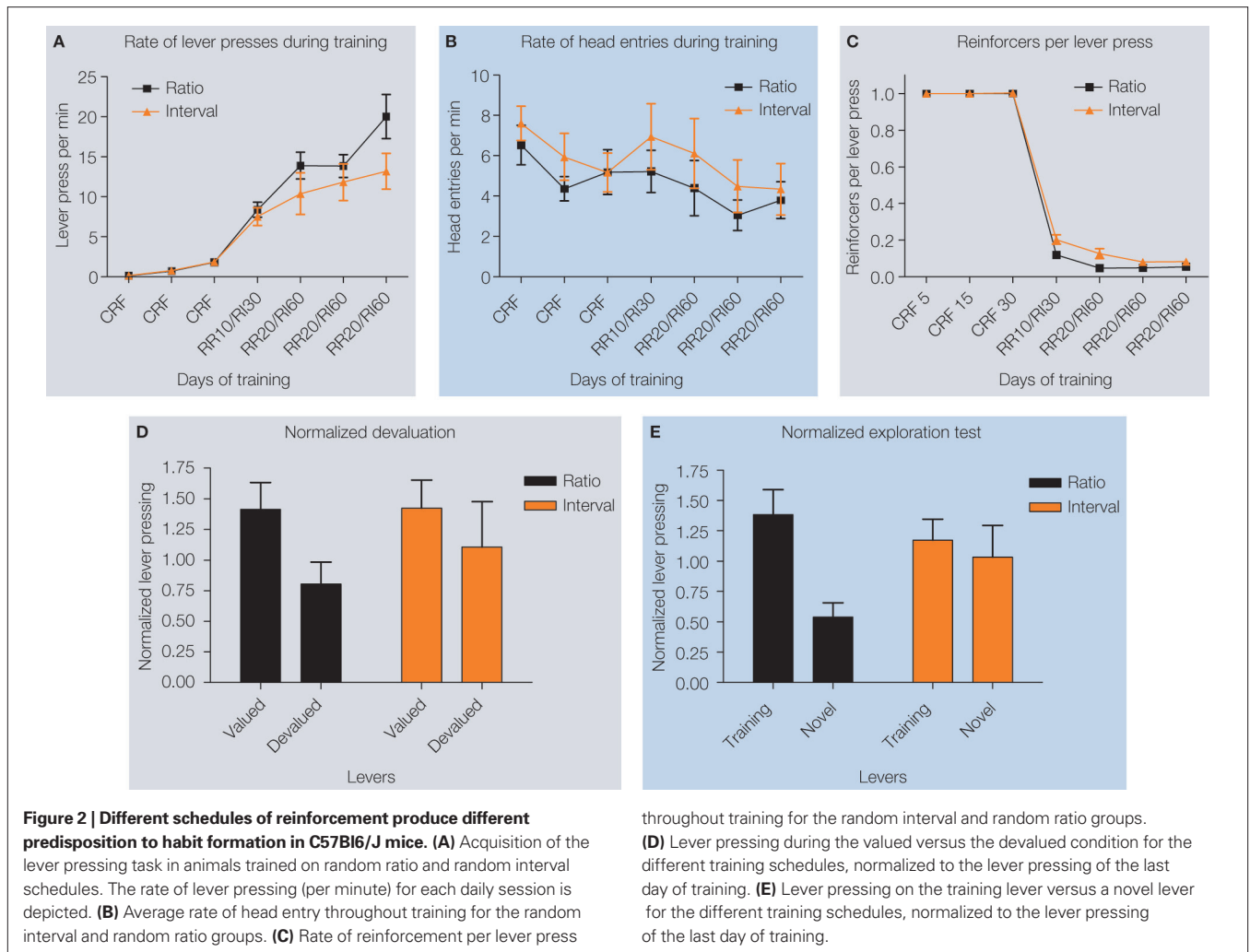


Figure 1 | Investigating goal-directed actions and habit formation in mice. (A) Mice were trained with two reinforcers. In the figure, the task is exemplified with one of the reinforcers, cheese, being delivered in the operant box contingent upon lever pressing, while the other reinforcer, sugar water, is being delivered freely to the mouse in the home cage. The types of reinforcers used in the figures are for illustrative purposes only. (B) Devaluation is performed in two days: Day 1, the mouse is given the reinforcer, cheese, previously earned by lever pressing (devalued condition); Day 2, the mouse receives the reinforcer, sugar water, previously freely available in its home cage (valued condition). The order of the conditions is randomized. Immediately after each feeding session,

which last 1 h, the mouse goes through a 5-min extinction test in the operant chamber, with the training lever extended. The number of presses on the training lever under the valued and the devalued conditions are compared. If the mouse presses more under the valued versus devalued condition, then the behavior is goal-directed behavior. However, if the mouse presses both levers equally his behavior is classified as habitual. (C) The generalization test. Two levers are presented in a 5-min extinction test: If the mouse pressed the training lever more than the novel lever, it is discriminating/exploiting. However, if the mouse presses both levers equally then there is significant generalization/exploration. Training lever is in blue and a novel lever is in pink.

non-contingently in their home cage and used as a control for the devaluation test (Figure 1A). After training the mice under a continuous reinforcement schedule to establish the relation between lever pressing and outcome delivery, animals were divided into two different groups: one group was trained under a random ratio schedule while the other group under a random interval schedule of reinforcement. Mice trained under a random ratio schedule of reinforcement received one reinforcer after a certain number of presses (on average every 20 lever presses in Hilario et al., 2007), whereas mice trained under a random interval schedule

received a reinforcer upon the first press after a certain interval had elapsed since the last reinforcer was earned (60 s on average in this study). During training, random ratio animals had a tendency to show higher rates of lever pressing than random interval animals, which is consistent with a strategy to maximize the number of reinforcers per press in the different schedules (Dickinson et al., 1983) (Figure 2A). For random ratio animals the more they press the more they earn, while for random interval animals the best strategy is to press at a rate matching the reinforcement rate in time. Despite the differences in pressing rate observed, Hilario



et al. (2007) matched training schedules so that the number of reinforcements, the reinforcement rate per lever press, and the reinforcement rate per time were relatively similar between ratio and interval trained animals (Figure 2B,C).

To determine if lever pressing in mice trained under different schedules was goal-directed or habitual, the effects of devaluation by sensory-specific satiety were examined during tests in extinction (Figure 1B). During this type of devaluation test, the outcome that was earned contingent upon lever pressing was devalued by satiating the animals with it before the extinction test (devalued), and the performance of the animal was compared to the control situation in which the animals were satiated with the reinforcer they got for free in their home cage (valued). This test allowed them to examine how much the lever pressing action was dependent upon the expected value of the outcome that was earned contingently upon lever pressing, and controlled for the motivational effects of general satiety. As expected,

during the devaluation test, random ratio-trained animals responded significantly less during the devalued condition than during the valued condition. Conversely, random interval-trained animals were insensitive to changes in value during the test, and pressed equally during the valued and devalued conditions, indicating that they were habitual (Figure 2D). Because random interval trained animals pressed less during training and during the test, they examined if the different sensitivity to devaluation of animals trained on ratio and interval schedules could be explained by a floor effect, i.e. that the random interval trained animals would not show devaluation because they could not decrease their lever pressing further. This was not the case, since when the performance was normalized to the amount of pressing during the last training day the same results were observed. Furthermore, no correlation was found between lever pressing during training or testing and the amount of devaluation for each of the training schedules. On the contrary, there was a

significant negative correlation between the total number of lever presses during devaluation and the amount of devaluation in interval schedule trained animals, indicating that animals that pressed less were the ones that devalued more. Therefore, Hilario et al. (2007) confirmed previous observations in rats that random interval schedules favor habit formation while random ratio schedules favor goal-directed behavior (Adams, 1982; Dickinson, 1985; Dickinson et al., 1983), and showed that these schedules of reinforcement can be used to study habit formation in mice.

Hilario et al. (2007) also introduced a new assay which investigates how much the animals explore or generalize to a novel lever. This test was designed based on the assumption that the shift from goal directed responding to habitual responding corresponds to a shift from actions being driven by the expected value of the outcome and the contingency between action and outcome (A-O relation) to actions being elicited by antecedent stimuli (S-R relation) (Balleine and Dickinson, 1994) (**Figure 1C**). They reasoned that if in habitual animals the response is being elicited by antecedent stimuli, then if they would be given a choice between pressing the training lever or a novel lever that is similar to the training lever but just in a different location, the mouse will show a tendency to generalize and thus press the novel lever. Conversely if goal-directed actions are being driven by the relation between the action and the outcome, the mouse should press more on the training lever and very little on the novel lever, which was never paired with the outcome. They showed for the first time that random interval schedules known to promote habit formation favor relatively more **exploration** of a novel lever in relation to those mice trained under random ratio schedules, which favored discrimination of the actions and exploitation of the reinforced lever (**Figure 2E**).

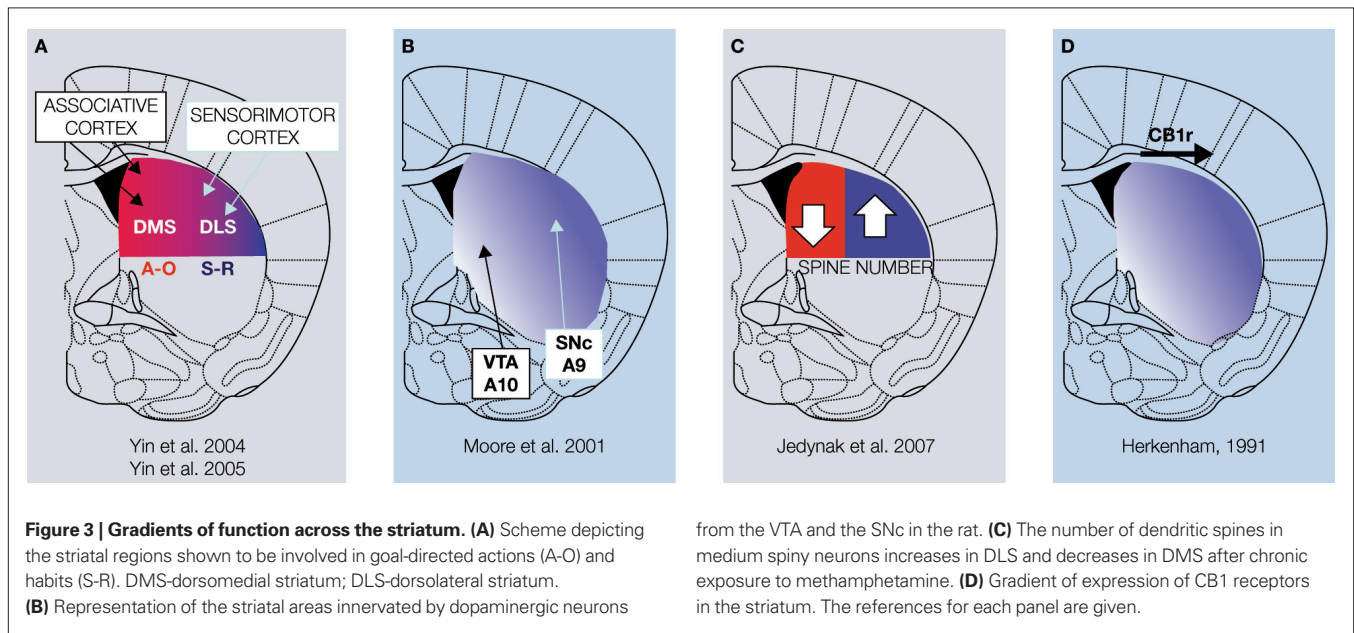
These results suggest that, in ratio trained animals, behavior is governed by the action-outcome relation, while in random interval trained animals, behavior is governed more by a stimulus-response relation. Hilario et al. (2007) concluded that the reinforcement schedules could be presented as useful tools in studying the molecular, cellular, and circuit mechanism of goal-directed actions and habit formation in mice. Furthermore, they suggested that the generalization/exploration test could be a complement to the devaluation test in mutant animals that may have different metabolism, different sensitivities to satiety, or different sensitivities to food reward. However, it still remains to be determined if the processes and

the neural substrates underlying generalization/exploration in the two-lever choice test and the insensitivity to changes in value in the devaluation test, are similar or different.

PARALLEL CORTICO-BASAL GANGLIA LOOPS AND GRADATION OF FUNCTION ACROSS THE STRIATUM

The neuroanatomical circuits that support goal-directed actions have been shown to differ from those supporting habitual behavior. Parallel cortico-basal ganglia loops seem to be critical for learning actions in a different manner. While the limbic loops that stream through the Nucleus Accumbens seem to mediate responses in relation to specific stimuli (stimulus outcome relations or pavlovian to instrumental transfer), loops that course through the dorsal striatum seem to be more involved in operant behavior (Parkinson et al., 2002; Setlow et al., 2002; Wiltgen et al., 2007; Yin et al., 2004, 2005b). Although the dorsal striatum in rodents is not divided clearly into caudate and putamen, it does have a medial-lateral gradient of connectivity which is similar (but not identical) to the caudate (ventromedial), and putamen (dorsolateral) connectivity in primates (McFarland and Haber, 2000; Voorn et al., 2004). The medial portion of the dorsal striatum, which extends ventrally to the limits of accumbens has been shown to receive most of its input from the associative areas of the cortex, (like the caudate), while the dorsolateral striatal region receives input from the sensorimotor areas of the cortex (like the putamen) (Voorn et al., 2004). The associative cortico-basal ganglia circuits involving the dorsomedial striatum (Yin et al., 2005a,b), the pre-limbic cortex (Balleine and Dickinson, 1998; Corbit and Balleine, 2003), and the mediadorsal thalamus (Corbit et al., 2003) have been shown to support the learning and performance of goal-directed behavior, but do not affect habit formation. In contrast, the dorsolateral or sensorimotor striatum (Yin et al., 2004) and the infralimbic cortex (Killcross and Coutureau, 2003) have been shown to support the formation of habits (**Figure 3A**). Interestingly, the different corticostriatal loops interact with each other (Kasanez et al., 2008). Given this, the shift from goal-directed behavior to habitual behavior in interval trained animals has been proposed to reflect a competition between the dorsomedial and the dorsolateral striatum (Yin et al., 2006), which are involved in these different types of learning respectively.

Cocaine self-administration in primates has been shown to progressively activate the limbic, associative and sensorimotor areas of the striatum



(Porrino et al., 2004), and administration of cocaine in rats induced a shift in task-related activity from ventromedial to dorsolateral striatum (Takahashi et al., 2007). Interestingly, the projection of dopaminergic neurons to striatum also follows an interesting gradient with dopaminergic neurons projecting from the substantia nigra pars compacta (A9) targeting more the dorsolateral striatum, and dopaminergic neurons projecting from the ventral tegmental area (A10) targeting more the ventromedial striatum, nucleus accumbens (Moore et al., 2001), and frontal cortices (Figure 3B). Consistently, lesions of the nigrostriatal input to the dorsolateral striatum (Faure et al., 2005), and infusion of dopamine into the ventral medial prefrontal cortex seem to impair habits and favor goal-directed behavior (Hitchcott et al., 2007).

The dopamine transporter (DAT), the main target of cocaine, is highly expressed in the dorsolateral striatum, and less expressed in more medial and ventral regions of the striatum and in the pre-frontal cortex, where Catechol-O-methyl transferase (COMT) is more prevalent (Arbuthnott and Wickens, 2007; Matsumoto et al., 2003). Sensitization with amphetamine, which also acts on the dopamine transporter, can increase dendritic spine density in medium spiny neurons (MSNs) in the dorsolateral striatum (Jedynak et al., 2007), which is necessary for habit formation, and at the same time decrease spine density in the dorsomedial striatum, which is critical for goal-directed instrumental behavior (Figure 3C). Consistently, amphetamine sensitization favors a shift from goal-directed to habitual

behavior (Nelson and Killcross, 2006; Nordquist et al., 2007).

In addition, LTP was found to occur more easily in the dorsomedial striatum, while LTD has been shown to be easier to induce in the dorsolateral striatum (Partridge et al., 2000). Interestingly, striatal LTD was found to depend on CB1 receptor activation, the primary molecular target in the brain of endocannabinoids (Gerdeman and Lovinger, 2001; Gerdeman et al., 2002). Endocannabinoid release in the striatum has been shown to be modulated by dopamine signaling (Giuffrida et al., 1999; Kreitzer and Malenka, 2005; Yin and Lovinger, 2006). Intriguingly, recent studies have shown that amphetamine sensitization depends on endocannabinoid signaling through CB1 receptors in the dorsal striatum (Corbille et al., 2007), which raises the possibility that the effects of amphetamine in predisposing for habit formation could be mediated by endocannabinoid signaling. Furthermore, the expression of CB1 receptors across the striatum displays a medial-lateral gradient of increased expression, with the highest expression in the dorsolateral striatum (Gerdeman et al., 2003; Herkenham et al., 1991), which has been shown to be necessary for habit formation (Figure 3D). Moreover, signaling through the cannabinoid receptor type 1 (CB1) has been implicated in reward and addiction (Caille et al., 2007; Casadio et al., 1999; Cossu et al., 2001; De Vries et al., 2001; Di Marzo et al., 2001; Gerdeman et al., 2003; Hansson et al., 2007; Houchi et al., 2005; Sanchis-Segura et al., 2004; Wang et al., 2003). This long line of evidence may suggest a possible role of endocannabinoid signaling in habit formation.

ENDOCANNABINOID SIGNALING IS CRITICAL FOR HABIT FORMATION

To study if habit formation is dependent upon endocannabinoid signaling, Hilario et al. (2007) employed mice with genetically targeted mutations in the *CB1* gene (Zimmer et al., 1999). Three groups of mice, wild-type (WT), *CB1*^{+/-}, and *CB1*^{-/-} littermates, were trained on a random interval schedule, previously shown by the authors to promote habitual behavior. Hilario et al. (2007) demonstrated that, independent of the genotype, all animals were capable of learning to press for reinforcements in a similar manner (Figure 4A). However, when tested on the devaluation test, while WT mice showed insensitivity to change in value of the outcome and thus habitual behavior, both *CB1*^{+/-}, and *CB1*^{-/-} mutants showed sensitivity to sensory-specific satiety, suggesting that their actions were still goal-directed (Figure 4B). These results were further confirmed using the exploration/generalization test. During the choice test, WT mice pressed equally the training lever and a novel lever similar to the training lever

(generalization/exploration) suggesting that their actions were habitual. However, *CB1*^{-/-} mutant mice pressed preferentially the training lever suggesting that their actions were driven by the relation between action and outcome (discrimination/exploitation) (Figure 4C).

CB1 receptors have been shown to be important for development, feeding behavior, and reward (Caille et al., 2007; Di Marzo et al., 2001; Sanchis-Segura et al., 2004). To prevent conclusions that could be confounded by possible chronic developmental or behavioral abnormalities in the *CB1* knockout mice, Hilario et al. (2007) ran another set of experiments using acute pharmacological blockade of *CB1* receptors. *CB1* receptors were blocked specifically during the random interval schedule training sessions with two different doses of the *CB1* receptor antagonist AM251 (Figure 4D). The devaluation and generalization tests that followed were performed in the absence of drug. Hilario et al. (2007) showed that the mice injected with the *CB1* antagonist during training were still sensitive to manipulations of outcome value and displayed

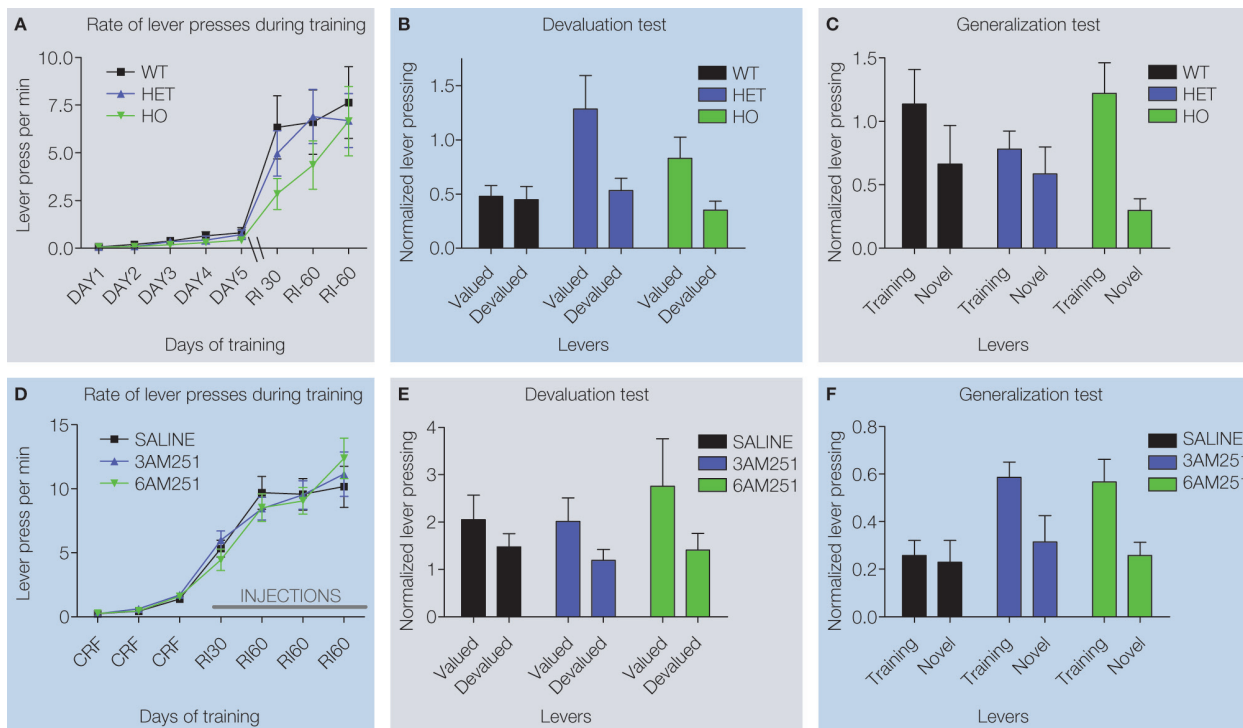


Figure 4 | Decreased predisposition for habit formation in *CB1* mutant mice and in C57Bl6/J mice injected with *CB1* antagonists during random interval training. (A) Acquisition of the lever pressing task in WT, *CB1*^{+/-} and *CB1*^{-/-} mice trained in a random interval schedule. The rate of lever pressing (per minute) for each daily session is depicted. (B) Normalized lever pressing during the valued versus the devalued condition for WT, *CB1*^{+/-} and *CB1*^{-/-} mice. (C) Lever pressing (normalized) on the training lever versus a novel lever in WT, *CB1*^{+/-} and *CB1*^{-/-} mice. (D) Acquisition of the lever pressing task for animals

injected with saline, 3 mg/kg AM251 or 6 mg/kg AM251. The rate of lever pressing (per minute) for each daily session is depicted. Note that animals were only injected during RI-30 and RI-60 training. (E) Normalized lever pressing during the valued versus the devalued condition for mice injected with saline, 3 mg/kg AM251 or 6 mg/kg AM251. (F) Lever pressing (normalized) on the training lever versus a novel lever in mice injected with saline, 3 mg/kg AM251 or 6 mg/kg AM251. The devaluation and generalization tests were performed without drug treatment.

a higher tendency to exploit the trained lever, while animals injected with saline during training were habitual in both tests (Figure 4E,F). These results indicate that CB1 activation is necessary during training but not during testing, and that the decreased predisposition observed in CB1 knockout mice is not likely attributable to developmental abnormalities or altered CB1 signaling during feeding on the devaluation test.

To summarize, genetic knockout and pharmacological blockade of CB1 receptors consistently impaired habit formation and the development of a stimulus-response behavioral pattern, providing evidence for the critical role of endocannabinoid signaling in habit formation.

WHERE AND HOW IS ENDOCANNABINOID SIGNALING NECESSARY FOR HABIT FORMATION

Hilario et al. (2007) showed that endocannabinoid signaling through CB1 receptors is critical for habit formation. This finding opens new lines of questioning, such as where and how CB1 signaling operates to promote habit formation. Endocannabinoids in the brain can function as retrograde messengers, modulating the release of different neurotransmitters, and producing short-term and long-term depression of excitatory and inhibitory transmission (Gerdeman et al., 2002; Kreitzer and Regehr, 2001; Wilson and Nicoll, 2001; Yin and Lovinger, 2006). Although CB1 receptors are one of the most-abundant G-protein coupled receptors in the brain and are expressed almost ubiquitously, we have already described the dorsolateral striatum as a good candidate for the “where” question. In the dorsolateral striatum CB1 receptors could serve to decrease “competing” glutamatergic inputs to MSNs by inducing depression at these synapses (Gerdeman et al., 2002; Huang et al., 2001). However, CB1 receptor activation is also important for the depression of inhibitory inputs in the dorsolateral striatum (Adermark and Lovinger, 2007), suggesting it could potentially reduce lateral inhibition between MSNs or reduce inhibition of MSNs by fast-spiking interneurons. Interestingly, a combination of depression of “competing” excitatory inputs and reduction in lateral inhibition could facilitate the firing of groups of neurons that are preferentially connected, like a cell assembly (Carrillo-Reid et al., 2008), with less interference from the cortex and competing cell assemblies in the striatum. CB1 mediated long-term depression in the striatum is expressed by a decrease in presynaptic release probability, which is manifested by a decrease in amplitude of spontaneous excitatory postsynaptic currents, but also by

an increase in paired pulse facilitation (a second afferent stimulation given within a certain time window of the first produces a larger response). Therefore, another interesting possibility is that endocannabinoid signaling through CB1 receptors acts as a filter to increase signal to noise, since after the induction of pre-synaptic depression the postsynaptic neuron would listen preferentially to bursts of inputs rather than single inputs.

CB1 is also expressed heavily in the distal terminals of the MSNs from the direct and indirect pathway, which synapse onto the substantia nigra pars reticulata and the globus pallidus, respectively (Sanudo-Pena et al., 1999). Therefore, since MSNs are inhibitory projection neurons, it is possible that endocannabinoid signaling through CB1 receptor activation is necessary to disinhibit basal ganglia nuclei downstream of the striatum. Another intriguing possibility is that CB1 mediated signaling modulates the strength of excitatory and inhibitory synaptic inputs onto dopaminergic neurons (Lupica and Riegel, 2005; Szabo et al., 2002). It has been shown that endocannabinoids are released in response drugs of abuse (Caille et al., 2007), and that the transient increases in dopamine release by drugs of abuse are mediated by CB1 receptors (Cheer et al., 2007). Since CB1 receptor blockade diminishes the effects of several drugs of abuse on dopamine release (Cheer et al., 2007), one possibility is that endocannabinoid-mediated inhibition of GABA release onto dopamine neurons is necessary for dopaminergic neurons to increase firing and release dopamine onto downstream targets like the dorsolateral striatum, where dopamine has been shown to be necessary for habit formation (Faure et al., 2005; Nelson and Killcross, 2006; Szabo et al., 2002).

POSSIBLE APPLICATIONS

Hilario et al. (2007) demonstrated that endocannabinoid signaling is necessary for the development of habitual behavior. Precisely how endocannabinoids modulate striatal information processing *in vivo* and interact with other neurotransmitter systems, such as glutamate, acetylcholine, and dopamine, is still a matter for much needed research. If endocannabinoids are indeed involved in the balance of the neural mechanisms that underlie our vulnerability to develop habits, drug seeking behaviors, compulsions, or even other striatal-based pathologies, their understanding is of the utmost importance to the formulation of more adequate treatments. Because current research has suggested that the endocannabinoid system can control the dopamine system and vice versa, the blockade of CB1 receptors has been targeted as a potential therapeutic approach for pathological conditions that involve dopamine-

related imbalances. The drug Rimobant, a CB1 antagonist, has been employed in the treatment of addiction (Cahill and Ussher, 2007), and has been proposed to function by reducing the levels of dopamine in the motivation centers of the brain, which are triggered by addictive drugs. This drug class has been shown to induce a decrease in drug rewarding effects, to reduce the influence of drug-associated stimuli, and to lower the relapse rates of drugs such as opioids, cocaine, nicotine, ethanol and amphetamine (De Vries et al., 2001; Le Foll et al., 2008). It has also been proposed that manipulations of endocannabinoid signaling through CB1 could be beneficial in other striatal involving disorders like Parkinson's disease (Garcia-Arencibia et al., 2008; Kreitzer and Malenka, 2007). In the future, it will be important to investigate the brain region and cell types where CB1 signaling is required for its effects, to not only define how endocannabinoids contribute to normal behavior, but to also understand how therapies can be customized to specific pathologies.

CONCLUSION

Hilario et al. (2007) demonstrate that training paradigms using different reinforcement schedules are useful tools for studying the molecular, cellular, and circuit mechanisms of goal-directed actions and habit formation in mice. Furthermore, they introduced a novel experimental behavioral tool, the generalization/exploration assay, which can be used complementarily with devaluation and contingency degradation assays to measure behavioral changes during habit formation. Using these paradigms for examining habit formation in mice, the authors showed using genetic and pharmacological tools that endocannabinoid signaling through CB1 receptors is necessary at the time of training for habit formation to occur.

ACKNOWLEDGEMENTS

We thank C. Gremel and E. Dias-Ferreira for helpful comments on this manuscript. This work was supported by the NIAAA DICBR.

REFERENCES

- Adams, C. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol. Comp. Physiol. Psychol.* 34B, 77–98.
- Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.* 33, 109–122.
- Adermark, L., and Lovinger, D. M. (2007). Retrograde endocannabinoid signaling at striatal synapses requires a regulated postsynaptic release step. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20564–20569.
- Arbuthnott, G. W., and Wickens, J. (2007). Space, time and dopamine. *Trends Neurosci.* 30, 62–69.
- Balleine, B. W., and Dickinson, A. (1994). Motivational control of goal directed action. *Anim. Learn. Behav.* 22, 1–18.
- Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.
- Cahill, K., and Ussher, M. (2007). Cannabinoid type 1 receptor antagonists (rimonabant) for smoking cessation. *Cochrane Database Syst. Rev.*, CD005353.
- Caille, S., Alvarez-Jaimes, L., Polis, I., Stouffer, D. G., and Parsons, L. H. (2007). Specific alterations of extracellular endocannabinoid levels in the nucleus accumbens by ethanol, heroin, and cocaine self-administration. *J. Neurosci.* 27, 3695–3702.
- Carrillo-Reid, L., Tecuapetla, F., Tapia, D., Hernandez-Cruz, A., Galarraga, E., Drucker-Colin, R., and Bargas, J. (2008). Encoding network states by striatal cell assemblies. *J. Neurophysiol.* 99, 1435–1450.
- Casadio, A., Martin, K. C., Giustetto, M., Zhu, H., Chen, M., Bartsch, D., Bailey, C. H., and Kandel, E. R. (1999). A transient, neuron-wide form of CREB-mediated long-term facilitation can be stabilized at specific synapses by local protein synthesis. *Cell* 99, 221–237.
- Cheer, J. F., Wassum, K. M., Sombers, L. A., Heien, M. L., Ariansen, J. L., Aragona, B. J., Phillips, P. E., and Wightman, R. M. (2007). Phasic dopamine release evoked by abused substances requires cannabinoid receptor activation. *J. Neurosci.* 27, 791–795.
- Colwill, R. M., and Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *J. Exp. Psychol. Anim. Behav. Process.* 11, 120–132.
- Corbille, A. G., Valjent, E., Marsicano, G., Ledent, C., Lutz, B., Herve, D., and Girault, J. A. (2007). Role of cannabinoid type 1 receptors in locomotor activity and striatal signaling in response to psychostimulants. *J. Neurosci.* 27, 6937–6947.
- Corbit, L. H., and Balleine, B. W. (2003). The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res.* 146, 145–157.
- Corbit, L. H., Muir, J. L., and Balleine, B. W. (2003). Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur. J. Neurosci.* 18, 1286–1294.
- Corbit, L. H., Ostlund, S. B., and Balleine, B. W. (2002). Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J. Neurosci.* 22, 10976–10984.
- Cossu, G., Ledent, C., Fattore, L., Imperato, A., Bohme, G. A., Parmentier, M., and Fratta, W. (2001). Cannabinoid CB1 receptor knockout mice fail to self-administer morphine but not other drugs of abuse. *Behav. Brain Res.* 118, 61–65.
- Coutureau, E., and Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* 146, 167–174.
- Creed, R. S., Denny-Brown, D., Eccles, J. C., Liddell, E. G. T., and Sherrington, C. S. (1932). Reflex Activity of the Spinal Cord. Oxford, London.
- De Vries, T. J., Shaham, Y., Homberg, J. R., Crombag, H., Schuurman, K., Dieben, J., Vanderschuren, L. J., and Schoffelmeier, A. N. (2001). A cannabinoid mechanism in relapse to cocaine seeking. *Nat. Med.* 7, 1151–1154.
- Di Marzo, V., Goparaju, S. K., Wang, L., Liu, J., Batkai, S., Jarai, Z., Fezza, F., Miura, G. I., Palmiter, R. D., Sugiura, T., and Kunos, G. (2001). Leptin-regulated endocannabinoids are involved in maintaining food intake. *Nature* 410, 822–825.
- Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. London B308*, 67–78.
- Dickinson, A., Campos, J., Varga, Z. I., and Balleine, B. (1996). Bidirectional instrumental conditioning. *Q. J. Exp. Psychol. B* 49, 289–306.
- Dickinson, A., Nicholas, D. J., and Adams, C. D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q. J. Exp. Psychol.* 35B, 35–351.
- Faure, A., Haberland, U., Conde, F., and El Massioui, N. (2005). Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *J. Neurosci.* 25, 2771–2780.
- Garcia-Arencibia, M., Ferraro, L., Tanganelli, S., and Fernandez-Ruiz, J. (2008). Enhanced striatal glutamate release after the administration of rimonabant to 6-hydroxydopamine-lesioned rats. *Neurosci. Lett.* 438, 10–13.
- Gerdeman, G., and Lovinger, D. M. (2001). CB1 cannabinoid receptor inhibits synaptic release of glutamate in rat dorsolateral striatum. *J. Neurophysiol.* 85, 468–471.
- Gerdeman, G. L., Partridge, J. G., Lupica, C. R., and Lovinger, D. M. (2003). It could be habit forming: drugs of abuse and striatal synaptic plasticity. *Trends Neurosci.* 26, 184–192.
- Gerdeman, G. L., Ronesi, J., and Lovinger, D. M. (2002). Postsynaptic

- endocannabinoid release is critical to long-term depression in the striatum. *Nat. Neurosci.* 5, 446–451.
- Giuffrida, A., Parsons, L. H., Kerr, T. M., Rodriguez de Fonseca, F., Navarro, M., and Piomelli, D. (1999). Dopamine activation of endogenous cannabinoid signaling in dorsal striatum. *Nat. Neurosci.* 2, 358–363.
- Hammond, L. J. (1980). The effect of contingency upon the appetitive conditioning of free-operant behavior. *J. Exp. Anal. Behav.* 34, 297–304.
- Hansson, A. C., Bermudez-Silva, F. J., Malinen, H., Hyttia, P., Sanchez-Vera, I., Rimondini, R., Rodriguez de Fonseca, F., Kunos, G., Sommer, W. H., and Heilig, M. (2007). Genetic impairment of frontocortical endocannabinoid degradation and high alcohol preference. *Neuropsychopharmacology* 32, 117–126.
- Herkenham, M., Lynn, A. B., Johnson, M. R., Melvin, L. S., de Costa, B. R., and Rice, K. C. (1991). Characterization and localization of cannabinoid receptors in rat brain: a quantitative in vitro autoradiographic study. *J. Neurosci.* 11, 563–583.
- Hilario, M. R. F., Clouse, E., Yin, H. H., and Costa, R. M. (2007). Endocannabinoid signaling is critical for habit formation. *Front. Integr. Neurosci.* 1, doi: 10.3389/neuro.07/006.2007.
- Hitchcott, P. K., Quinn, J. J., and Taylor, J. R. (2007). Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cereb. Cortex.* 12, 2820–2827.
- Houchi, H., Babovic, D., Pierrefiche, O., Ledent, C., Daoust, M., and Naassila, M. (2005). CB1 receptor knockout mice display reduced ethanol-induced conditioned place preference and increased striatal dopamine D2 receptors. *Neuropsychopharmacology* 30, 339–349.
- Huang, C. C., Lo, S. W., and Hsu, K. S. (2001). Presynaptic mechanisms underlying cannabinoid inhibition of excitatory synaptic transmission in rat striatal neurons. *J. Physiol.* 532, 731–748.
- Hull, C. L., and Clark, L. (1943). Principles of behavior: an introduction to behavior theory. Appleton-Century-Crofts, New York.
- Jedynak, J. P., Uslander, J. M., Esteban, J. A., and Robinson, T. E. (2007). Methamphetamine-induced structural plasticity in the dorsal striatum. *Eur. J. Neurosci.* 25, 847–853.
- Kasanez, F., Riquelme, L. A., Della-Maggiore, V., O'Donnell, P., and Murer, M. G. (2008). Functional integration across a gradient of corticostriatal channels controls UP state transitions in the dorsal striatum. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8124–8129.
- Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408.
- Kreitzer, A. C., and Malenka, R. C. (2005). Dopamine modulation of state-dependent endocannabinoid release and long-term depression in the striatum. *J. Neurosci.* 25, 10537–10545.
- Kreitzer, A. C., and Malenka, R. C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 445, 643–647.
- Kreitzer, A. C., and Regehr, W. G. (2001). Cerebellar depolarization-induced suppression of inhibition is mediated by endogenous cannabinoids. *J. Neurosci.* 21, RC174.
- Le Foll, B., Forget, B., Aubin, H. J., and Goldberg, S. R. (2008). Blocking cannabinoid CB1 receptors for the treatment of nicotine dependence: insights from pre-clinical and clinical studies. *Addict. Biol.* 13, 239–252.
- Lupica, C. R., and Riegel, A. C. (2005). Endocannabinoid release from mid-brain dopamine neurons: a potential substrate for cannabinoid receptor antagonist treatment of addiction. *Neuropharmacology* 48, 1105–1116.
- Matsumoto, M., Weickert, C. S., Akil, M., Lipska, B. K., Hyde, T. M., Herman, M. M., Kleinman, J. E., and Weinberger, D. R. (2003). Catechol O-methyltransferase mRNA expression in human and rat brain: evidence for a role in cortical neuronal function. *Neuroscience* 116, 127–137.
- McFarland, N. R., and Haber, S. N. (2000). Convergent inputs from thalamic motor nuclei and frontal cortical areas to the dorsal striatum in the primate. *J. Neurosci.* 20, 3798–3813.
- Moore, A. E., Cicchetti, F., Hennen, J., and Isacson, O. (2001). Parkinsonian motor deficits are reflected by proportional A9/A10 dopamine neuron degeneration in the rat. *Exp. Neurol.* 172, 363–376.
- Nelson, A., and Killcross, S. (2006). Amphetamine exposure enhances habit formation. *J. Neurosci.* 26, 3805–3812.
- Nordquist, R. E., Voorn, P., de Mooij-van Malsen, J. G., Joosten, R. N., Pennartz, C. M., and Vanderschuren, L. J. (2007). Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. *Eur. Neuropsychopharmacol.* 17, 532–540.
- Parkinson, J. A., Dalley, J. W., Cardinal, R. N., Bamford, A., Fehner, B., Lachenal, G., Rudarakanchana, N., Halkerston, K. M., Robbins, T. W., and Everitt, B. J. (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behav. Brain Res.* 137, 149–163.
- Partridge, J. G., Tang, K. C., and Lovinger, D. M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J. Neurophysiol.* 84, 1422–1429.
- Porrino, L. J., Lyons, D., Smith, H. R., Daunais, J. B., and Nader, M. A. (2004). Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. *J. Neurosci.* 24, 3554–3562.
- Sanchis-Segura, C., Cline, B. H., Marsicano, G., Lutz, B., and Spanagel, R. (2004). Reduced sensitivity to reward in CB1 knockout mice. *Psychopharmacology (Berl.)* 176, 223–232.
- Sanudo-Pena, M. C., Tsou, K., and Walker, J. M. (1999). Motor actions of cannabinoids in the basal ganglia output nuclei. *Life Sci.* 65, 703–713.
- Setlow, B., Gallagher, M., and Holland, P. C. (2002). The basolateral complex of the amygdala is necessary for acquisition but not expression of CS motivational value in appetitive Pavlovian second-order conditioning. *Eur. J. Neurosci.* 15, 1841–1853.
- Szabo, B., Siemes, S., and Wallmichrath, I. (2002). Inhibition of GABAergic neurotransmission in the ventral tegmental area by cannabinoids. *Eur. J. Neurosci.* 15, 2057–2061.
- Takahashi, Y., Roesch, M. R., Stalnaker, T. A., and Schoenbaum, G. (2007). Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. *Front. Integr. Neurosci.* 1, nihpa51247.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
- Tolman, E. C. (1949). There is more than one kind of learning. *Psychol. Rev.* 56, 144–155.
- Von Holst, E. (1973). The Behavioural Physiology of Animals and Man. Coral Gables, University of Miami Press.
- Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W., and Pennartz, C. M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci.* 27, 468–474.
- Wang, L., Liu, J., Harvey-White, J., Zimmer, A., and Kunos, G. (2003). Endocannabinoid signaling via cannabinoid receptor 1 is involved in ethanol preference and its age-dependent decline in mice. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1393–1398.
- Wilson, R. I., and Nicoll, R. A. (2001). Endogenous cannabinoids mediate retrograde signalling at hippocampal synapses. *Nature* 410, 588–592.
- Wiltgen, B. J., Law, M., Ostlund, S., Mayford, M., and Balleine, B. W. (2007). The influence of Pavlovian cues on instrumental performance is mediated by CaMKII activity in the striatum. *Eur. J. Neurosci.* 25, 2491–2497.
- Yin, H. H., and Knowlton, B. J. (2004). Contributions of striatal subregions to place and response learning. *Learn Mem.* 11, 459–463.
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22, 505–512.
- Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav. Brain Res.* 166, 189–196.
- Yin, H. H., and Lovinger, D. M. (2006). Frequency-specific and D2 receptor-mediated inhibition of glutamate release by retrograde endocannabinoid signaling. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8251–8256.
- Zimmer, A., Zimmer, A. M., Hohmann, A. G., Herkenham, M., and Bonner, T. I. (1999). Increased mortality, hypoactivity, and hypoalgesia in cannabinoid CB1 receptor knockout mice. *Proc. Natl. Acad. Sci. U.S.A.* 96, 5780–5785.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 September 2008; accepted: 06 November 2008; published: 15 December 2008.

Citation: *Front. Neurosci.* (2008) 2, 2:208–217. doi: 10.3389/neuro.01.030.2008

Copyright: © 2008 Hilário and Costa. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.