A neurocomputational hypothesis for nicotine addiction

Boris S. Gutkin[†], Stanislas Dehaene[‡], and Jean-Pierre Changeux[†§]

[†]Récepteurs et Cognition, Unité de Recherche Associée, Centre National de la Recherche Scientifique 2184, Institut Pasteur, 75015 Paris, France; and [‡]Cognitive Neuroimaging, Institut National de la Santé et de la Recherche Médicale–Commissariat à l'Energie Atomique, Unit 562, Service Hospitalier Frédéric Joliot, 91401 Orsay, France

Contributed by Jean-Pierre Changeux, November 30, 2005

We present a hypothetical neurocomputational model that combines a set of neural circuits at the molecular, cellular, and system levels and accounts for several neurobiological and behavioral processes leading to nicotine addiction. We propose that combining changes in the nicotinic receptor response, expressed by mesolimbic dopaminergic neurons, with dopamine-gated learning in action-selection circuits, suffices to capture the acquisition of nicotine addiction. We show that an opponent process enhanced by persistent nicotine-taking renders self-administration rigid and habitual by inhibiting the learning process, resulting in long-term impairments in the absence of the drug. The model implies distinct thresholds on the dosage and duration for the acquisition and persistence of nicotine addiction. Our hypothesis unites a number of prevalent ideas on nicotine action into a coherent formal network for further understanding of compulsive drug addiction.

computational model | reward

Tobacco addiction is a multistage process involving persistent cycles of chronic smoking (1, 2). It is tied to long-lasting effects on mesolimbic dopaminergic (DA) pathways by nicotine, the main addictive substance in tobacco smoke (2, 3). Although the pharmacological target of nicotine is now well identified (4), how nicotine binding translates into addictive behavior remains enigmatic. Particularly puzzling is the ease with which nicotine addiction is acquired and resists despite its limited (or even negative) hedonic impact (3).

Both reinforcement learning and opponent processes have been proposed to play a significant role in the development of addiction (5, 6). However, direct DA-dependent reinforcement learning does not conclusively treat the issue of persistence of addictive behaviors to extinction (5), and the opponent process or allostasis (6) theory does not provide a specific computational account of how addiction is acquired. To unravel two such aspects, the acquisition and persistence, we model the interplay between the phasic and the persistent effects of nicotine on the DA pathway and its assumed control of the plasticity in corticostriatal action-selection (A-S) circuits (7). We propose a hypothetical minimal computational circuit of neuronal and pharmacological processes (see Fig. 1) and apply it to a specific animal model of smoking in humans: self-administration of nicotine. Our framework implements the dynamics of interaction between the effect of nicotinic acetylcholine receptors (nAChRs) on the DA neuronal population and learning in a model of A-S. We specify the differential roles of nicotine in the positive (direct) DA response on the acquisition of nicotine-taking and the slow opponent process in its persistence.

At the molecular level, nicotine actions are mediated by persistent changes in brain nAChRs involving, among others, the β2-subunit (8–11). These nAChRs modulate the excitability of the DA neurons in the ventral tegmental area (VTA) that project to striatal structures (e.g., nucleus accumbens and striatum) (2, 12–15). Under chronic nicotine, nAChRs cycle through sensitization/desensitization (16, 17), leading to long-lasting up-regulation (18, 19). We propose that such cycles progressively recruit an opponent process: nAChR depression or down-regulation (10, 20). Hence, nicotine would provoke both transient and persistent changes in the DA signaling on different time scales.

At the circuit level, nicotine boosts VTA DA neuron activity and their response to glutamatergic afferents (13) in part by differentially increasing the phasic responses (21). Phasic nicotine injections lead to transient increases in DA levels in the nucleus accumbens and striatum, contingent on nAChR activation (8). Nicotine also affects VTA activity through the GABAergic interneurons (12) that further modify DA signals and may also subserve complementary reward signaling (22).

At the behavioral level, long-term smoking has been for a long time considered to be addictive and habit-forming (23–25). In animal models of smoking, nicotine injection has both motor-activational effects (through the dorsal nigrostriatal pathway) and rewarding effects (through the ventral mesolimbic pathway) (14). Specifically, animals learn to self-inject nicotine (see refs. 26 and 27) in two choice paradigms through nose pokes (8), lever pressing (28), and/or entering specific maze locations (such as Y-maze).

Deletion of the β2-subunit in mice abolishes nicotine administration by direct intra-VTA injections in a Y-maze task (29). Reexpressing β2*nAChRs in VTA (but not more dorsally) reestablishes self-administration in $\beta2^{-/-}$ animals (29). This finding demonstrates the role of VTA nAChRs at the behavioral level and justifies the assumptions of the model.

The model addresses the following experimental data: (*i*) The transient increases and long-term renormalization of the striatal DA levels (30); (*ii*) the development of behavioral sensitization (31, 32); (*iii*) nicotine self-administration both systemically and intracranially (27), where the drug delivery latency decreases to reach stable rates and the choice of the neutral action declines until the animal performs almost exclusively the armed action; and (*iv*) the development of withdrawal symptoms by removal of nicotine, i.e., plunging DA levels and a state of decreased reward (e.g., decrease in brain stimulation reward described in ref. 33).

Main Hypothesis

Nicotine effects on the VTA DA signaling initiate a cascade of molecular changes that, in turn, bias glutamatergic learning processes in the dorsal striatal structures responsible for behavioral choice, leading to the onset of stable self-administration. Nicotine, by activating and up-regulating nAChRs, dynamically changes the gain of the DA signaling: nicotine both potentiates the phasic DA signal to rewarding stimuli and evokes such a signal by itself. Phasic DA, in turn, directs learning at the excitatory (corticostriatal) connections in the A-S circuit. Tonic DA gates this learning process (3, 34). Persistent nicotine-

Conflict of interest statement: No conflicts declared.

Abbreviations: A-S, action-selection; DA, dopaminergic; nAChR, nicotinic acetylcholine receptor; VTA, ventral tegmental area.

[§]To whom correspondence should be addressed. E-mail: changeux@pasteur.fr.

www.pnas.org/cgi/doi/10.1073/pnas.0510220103

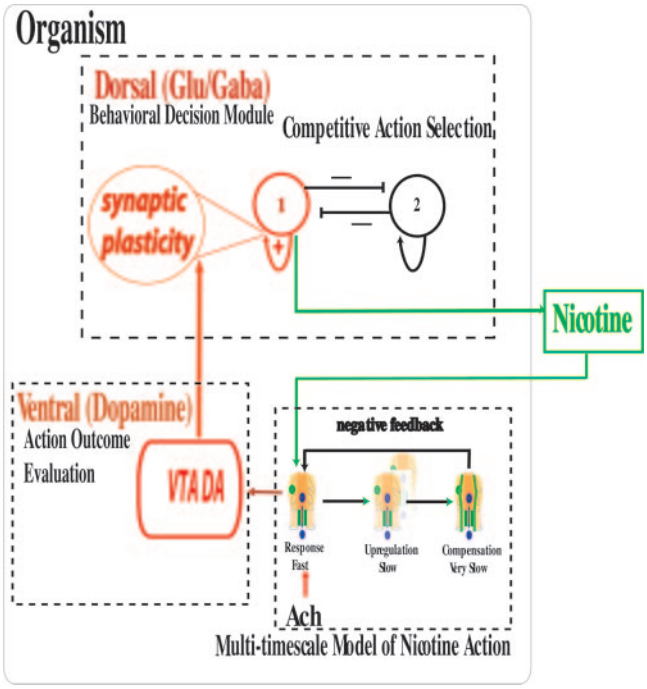

Fig. 1. Minimal functional circuit for nicotine addiction revealing main components of the model. The circuit illustrates that the DA signals from the ventral pathway influence learning in the dorsal action selector. The DA signal is modulated by the nicotine binding to the nAChRs.

dependent renormalization in tonic DA causes the learned behavioral bias to become rigid, and nicotine self-administration progressively escapes from the DA control. Because DA adaptations lead (in time) the learning in the A-S machinery, our model implements and gives a functional meaning to the ventral–dorsal progression of long-term drug addiction (12).

Model Framework

The model consists of two modules, both affected by nicotine: a module that simulates action selection and a DA signal module governing outcome-driven learning of action choices.

DA Signaling Module. This module consists of a single neuronal population that is under the control of acetylcholine (or nicotine) binding to the nAChRs. Phasic DA activity signals action outcomes: the burst responses of the VTA DA neurons (15). In behavioral terms, the DA response relates to the reward-prediction error (35), which in some versions of reinforcement learning theory modifies the "value function"/"learned incentive" of the stimulus/action choice (36, 37). In our specific context the VTA DA signal is affected directly by nicotine binding to the nAChRs; thus, the notion of "outcome" of the action becomes implicit. For the neutral (unrewarded) action (no nicotine injection), we hypothesize no phasic activation of the DA population.

We model nicotine-injection-dependent changes in the DA module as a cascade occurring on three distinct time scales: (*i*) phasic nicotine effect, a short time scale activation (seconds) of nAChRs by nicotine action [nicotine activates nAChRs on the DA neurons in the VTA, thereby directly increasing their excitability (gain effect) and potentiating the phasic activity of the DA population]; (*ii*) a slower positive-feedback process acts on a time scale of minutes and reflects the up-regulation of nAChRs by a repeated exposure to nicotine that persists after a long-time drug exposure; (*iii*) and long-term homeostatic opponency, a renormalization of the nAChR response after prolonged hyperactivation by nicotine (scale of several weeks) and settling of nAChRs into an inactivated but sensitized state/phenotype. This effect results in lower tonic DA response while preserving the phasic signals in response to nicotine.

A-S Module. We implement A-S by a competitive activation in a "winner-take-all" network representing activity in the dorsal nigrostriatocortical pathway (7, 38, 39). We identify each of the units, with its recurrent excitatory connections, as a separate corticostriatocortical loop coding for an action plan. Competition between the action plans (or loops) is implemented through cross-inhibition between the units. In physiological terms, this effect can be subserved either by the GABAergic collaterals whose physiological function has been identified in the striatum (40) or as an outcome of the negative feedback between direct and indirect corticostriatal pathways. Recurrent excitation ensures that A-S occurs even in response to transient stimuli. Action choice depends on integrating the stimulus strength, the internal dynamics (self-activation and cross-inhibition), and the random inputs (see supporting information, which is published on the PNAS web site). This model generates realistic response-time distributions (41), which change as the reward history modifies the A-S connections.

DA-Governed A-S Learning. DA signal modifies responses of the action selector by gating learning in the recurrent excitatory connections (17) through a DA-dependent rule shown in Scheme 1 (34). This scheme represents activity-dependent plasticity of the glutamatergic synapses in the striatocorticostriatal loops. The weight increases for each action-plan unit are contingent on both the unit activation (pre/postsynaptic) and a phasic DA signal, whereas weight decreases are contingent on tonic DA and unit activity (or possibly the absence of such with phasic DA) (see ref. 42 for similar schemas).

The learning in the A-S recurrent excitatory weights is governed by a DA-gated and informed learning rule (34) shown in Scheme 1. The differential equations are given in *Methods*.

Under control conditions, a choice yielding an unexpected reward (signaled by phasic DA firing) is potentiated (weights increased) and other actions are depressed (weights decreased). Nicotine injections act at three time scales to (*i*) provoke, (*ii*) potentiate, and (*iii*) depress different aspects of the DA signal. Initially, this boosts the drug-armed action weight, installing the positive-feedback cycle of self-administration. In the long-term, the opponent process reducing the tonic DA would freeze the plasticity and make the previously potentiated behavior robust to changes in the rewarding quality of nicotine.

Results

Nicotine Injections Elicit Neuroadaptation in the DA Module. The model simulates the following aspects of nicotine actions. Phasic nicotine injection provokes an acute increase in nAChR activation. Under tonic nicotine injection, this increase is followed by a slower up-regulation of the nAChRs (see Fig. 2). A third, much slower

$$\text{Tonic DA and } \begin{cases} + \text{ Phasic DA increase above tonic and pre/postactivity increase} \Rightarrow \text{potentiate weights} \\ + \text{ Phasic DA increase above tonic and no pre/postactivity increase} \Rightarrow \text{decrease weights} \\ \text{no Phasic DA and pre/postactivity increase} \Rightarrow \text{decrease weights} \end{cases}$$

Scheme 1

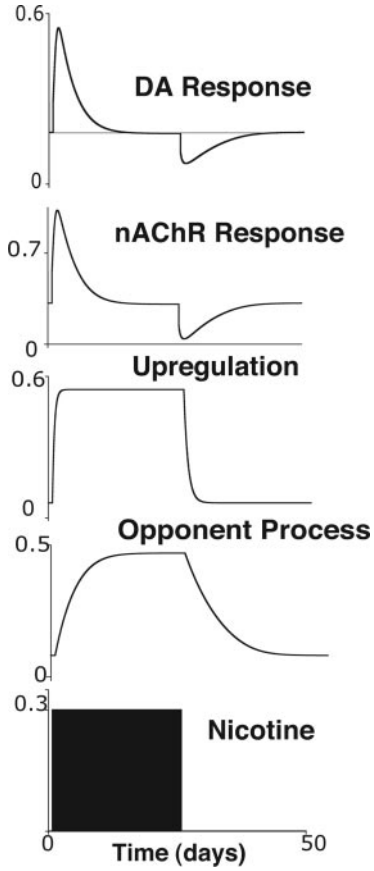

Fig. 2. Simulated long-lasting nicotine effects on DA neurons. Graphs (from top to bottom) show the DA response, the nAChR response, the up-regulation, and the long-term opponency. The bottom graph marks the presence of nicotine through either injection or infusion. Note the phasic nature of the nAChR response, which causes a transient change in the DA activity. After nicotine delivery stops, DA response decreases, which is a putative signature of withdrawal.

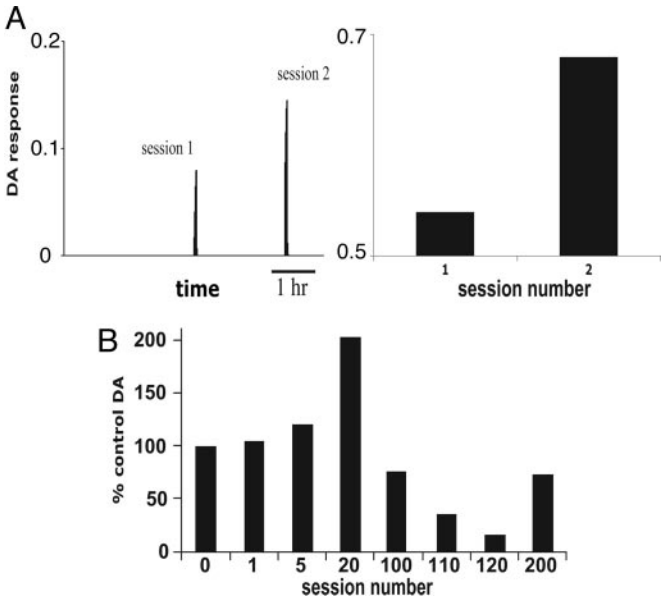

Fig. 3. Phasic nicotine sensitization and withdrawal after prolonged use. (*A*) Simulation of phasic nicotine effect and sensitization of action choice. (*Left*) DA response to two short-lasting administrations of nicotine. Nicotine activates and sensitizes the DA signal to successive doses through nAChR up-regulation. Behavioral outcome of the sensitization is shown. (*Right*) The choice probability of the drug-reinforced action. The choice for the reinforced action is differentially amplified because of the A-S mechanism. (*B*) Simulated consequence of nicotine withdrawal after long-term administration. The time-averaged response of the DA module to a transient excitation is shown as percentage of the control response: the response increases during the up-regulation of the nAChR activity, followed by a renormalization. Upon nicotine removal, DA signal drops dramatically and recovers on the slow opponent process time scale.

opponency acts to renormalize the average nAChR sensitivity to control levels. When nicotine is removed, this slow opponency leads to a long-term depression of the nAChR activity. Because nAChR activity controls the gain of the DA population, nicotine modifies the DA signal that may be elicited by behaviorally relevant environmental stimuli and actions. The transient or medium-term nicotine administration leads to activation, followed by up-regulation of the nAChR population, in turn increasing the gain of the DA population. The long-term opponency results in a decrease in the amount of nAChR activation. When nicotine is injected (only phasically in the model) the DA signal shows, as expected, a rapid increase due to the activation of the nAChRs (see Fig. 3 for an example of such response and ref. 43 for an experimental analog); however, in the simulation, the slow opponent renormalization of the DA response to nicotine is not evident (see below for discussion).

A chronic administration of nicotine leads to rapid DA signal increase (Fig. 2, first graph) followed by a sensitized DA signal and a slower renormalization (to control levels when nicotine is present). This cascade directly corresponds to DA signaling in animals chronically administered with nicotine (44) and hints at what would happen in animals that chronically self-administer nicotine (15). In the long term, the spontaneous DA activity returns to normal levels, the DA signal for weaker stimuli is depressed, and nicotine-boosted stimuli are potentiated. Under removal of nicotine the DA level drops dramatically (see below) and recovers much later, on the time scale of the slow opponent process.

Nicotine Self-Administration Is Driven by the DA-Gated Action Learning. Injections of nicotine potentiate the DA signal to gate plasticity of the excitatory self-connections in the A-S module. Provided that the injection is contingent on a specific action choice, the excitatory weights of the corresponding neural population are increased and a bias in A-S is established, leading to self-administration of nicotine (Fig. 4; open symbols represent the neutral action choice, and filled symbols represent nicotine-armed choice).

During the initial stages of self-administration (Fig. 4*A*, sessions 2–4), the nicotine intake leads to a general increase in motor activity: both armed and neutral action latencies decrease. In this regime the A-S is similar to control conditions: random inputs break the symmetry in the dynamics and lead to self-initiated choices (50/50 between the two) (see supporting information for noise-driven A-S). Subsequently, the behavior becomes selective: the recurrent weight for the nicotine-administration action plan increases beyond the threshold necessary to overcome the random inputs. This action is progressively robustly selected, and the neutral choice declines. The initial latency decrease corresponds to the motor activational effects of nicotine observed experimentally (45, 46), and the differential responding for the armed vs. the neutral action choice is in direct concordance with experimental results on nicotine self-administration (27).

For transient self-administration, responding extinguishes once nicotine is withdrawn but is reacquired by the model once nicotine is reestablished (Fig. 4*A*), indicative of behavioral sensitization (31,

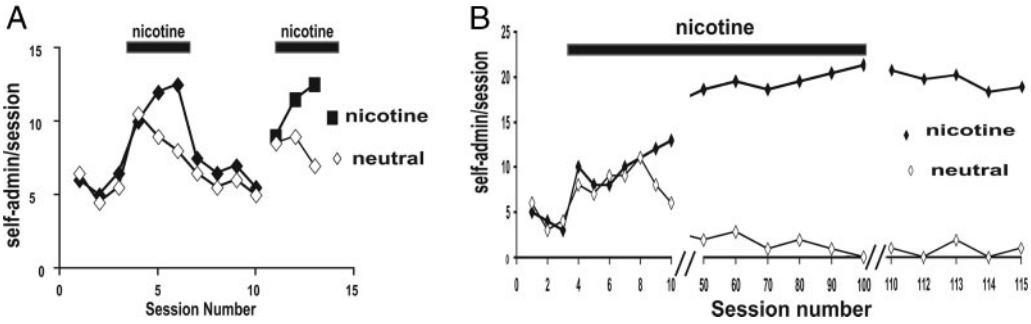

Fig. 4. Simulated nicotine self-administration: short and long-term effects. (*A*) Acquisition of self-administration, extinction, and reinstatement with a brief nicotine delivery (four sessions): a transient response increase for both choices, followed by selective responding; increase in the drug-reinforced choice; and decrease in the neutral choice. Note the transient differential responding (declines after session 6 and reinstates with session 11). Bars show the availability of nicotine. (*B*) Long-term differential responding becomes persistent with learning frozen because of the opponent process. Nicotine is available in self-administration sessions 3–100.

32). This occurs because the tonic DA levels during the short-term self-administration allow for transient plasticity in the A-S along with the direct activating effects of nicotine.

With prolonged self-administration (Fig. 4*B*), the influence of the DA signal and the plasticity progressively renormalize because of the opponent process: the behavioral bias for the action plan leading to nicotine becomes "stamped-in." In the long-term, the self-administration behavior becomes "routinized," independent of the hedonic or motivational value of nicotine: it continues to be self-administered when nicotine is withdrawn and even when paired with an aversive stimulus (simulations not shown). This progression forms a key experimental prediction of the model.

Nicotine Self-Administration Leads to Behavioral Sensitization and Withdrawal. The model shows that transient injections of sufficient nicotine doses potentiate the DA activity to the levels necessary to initiate plasticity of the self-connections in the A-S module. The priming dose leads to a small bias in the behavior, which is boosted to observable levels by a subsequent nicotine dose and results in behavioral sensitization (Fig. 3) (31, 32). We note here that the sensitization of the DA responses to nicotine injections does not necessarily lead to persistent behavioral sensitization directly: this effect is crucially dependent on DA-gated learning (potentiation of weights in the model). In the model, DA may, depending on the dose, lead to transient behavioral sensitization.

Furthermore, the model shows that persistent self-administration of nicotine leads to up-regulation of nAChRs and the recruitment of the opponent process (simulations not shown). Subsequently, if in simulations of long-lasting self-administration (as in Fig. 4*B*) nicotine is withdrawn, the DA signal shows a dramatic decline that persists for a long period (Fig. 3*B*). This decline is due to the long-term removal of the up-regulated nAChR signal by the slow opponent process. Such decline has been reported during nicotine withdrawal after a prolonged self-administration (15). In our framework the nicotine-evoked opponent process has a double effect: on learning and on the hedonic state. The drop in the tonic DA leads to a pathology of learning, and the behavior becomes stamped in. Because the outputs of the DA neurons from the VTA terminate in the brain structures associated with an hedonic state as well as behavioral choice, a decrease of DA signal would putatively result in anhedonia and possibly other somatic withdrawal signals.

Discussion

We have shown that a minimal neurocomputational model, positing nicotine-elicited neuroadaptations in the VTA and DA-modulated learning in the dorsal striatocortical A-S, suffices to account for a number of basic experimental findings associated with nicotine addiction. The model gives a key role to the long-lasting up-regulation of the nAChRs by nicotine with the subsequent effects on the DA signals. Short-term nicotine effects lead to behavioral motor activation. Provided that nicotine is administered in a behaviorally contingent manner (such as self-administration), this effect provides the initial behavioral bias that leads to a preferential selection of actions resulting in addiction. The long-term up-regulation of the receptors causes a potentiation of the DA signal carrying information about the incentive salience of actions (action-outcome contingency), in turn leading to differential gating of synaptic plasticity in neuronal populations coding for the specific action choice.

Our model also shows that a slow opponent process plays a key role in cementing the drug-associated behavior by reducing DA responses. When nicotine is removed, the DA levels drop below those required for efficient learning. During the withdrawal the animal does not unlearn drug-taking choices despite possible negative consequences. Hence, once learned, self-administration behavior escapes from motivational control and becomes robust to extinction, i.e., habitual. The reduced DA responses to all stimuli/actions lead to a general hypohedonic state. The model predicts reduced striatal plasticity in animals chronically administered and then withdrawn from nicotine (the reduced or withheld dose), which would then show deficits in readjusting their behavior under new conditions. This prediction is testable in both humans and laboratory animals and perhaps in analogy with behavior observed in nAChR knockout mice (46).

The above results imply that distinct dose/duration thresholds for stable acquisition and persistence of the nicotine-associated behavior are defined by the drug ability to activate nAChRs, to recruit nAChR up-regulation and the opponent process. The lowest threshold is for behavioral sensitization, which does not require learning in the A-S and is only transient. Below this threshold, desensitization rather than sensitization would occur and nicotine would have an aversive behavioral valence. Next is the threshold for differential self-administration. It occurs when the action choice weights are learned beyond the threshold necessary to overcome the internal noise: the up-regulated nAChRs and persistently increased DA gain ensure learning. Further threshold for the persistence of nicotine-taking is defined by recruitment of the opponent process (withdrawal, lack of unlearning, and escape from DA control of behavior).

Our computational framework further implies that, in the course of addiction, the persistent changes should progress from the mesolimbic circuit to the A-S nigrostriatofrontal loops. Initial mesolimbic changes are necessary for initiation of addictive behav-

ior, caused by the motivational, activational, and or hedonic effects of nicotine. The subsequent A-S changes are largely independent of the motivation valence of the drug as long as the behavioral bias has been acquired and reinforced for a sufficiently long time scale (that of the slowest opponent process).

We have made a number of mechanistic choices and assumptions to subserve the functional effects, such as the nicotine→DA pathway and the identity of the opponent process. Hence, experimental data are required to validate such choices. For instance, the nAChR up-regulation needs to be observed in self-administration paradigms; a subsequent opponent process, at the level of the receptors, needs to be identified as, for example, a down-regulation in the number of receptors or a change in affinity/cooperativity mediated by covalent modifications or changes in subunit composition (20). The crucial role of the VTA DA signaling in the onset of self-administration implies that wild-type animals should self-administer nicotine most robustly into the VTA directly (see ref. 29). Recent work (21) indicates indirectly that nicotine may be self-administered into the striatum, provided that the action choice is already sufficiently salient to evoke bursting DA activity. This requirement is different from other addictive drugs (like cocaine) that control DA at the targets of the VTA afferents.

Our model is inspired by abstract reinforcement learning models (47) of instrumental conditioning. It is in line with a recent model for cocaine addiction as a disorder of DA-signaled reward learning (5). We also assign a central role to DA signal and, in particular, to the phasic signaling mode. However, unlike the previous efforts, we provide a specific biological model of drug action at the receptor and DA-circuit level and of the tonic DA function in the onset and progression of addictive behavior, and we further pinpoint the role of the opponent process. We do so without an explicit representation of value (or "incentive value") of the various actions, but we rely on action choice and learning. In our model, drug-taking (self-administration) does not result in an infinitely increasing value for the drug-seeking, nor does it imply context-independent value encoding.

As with any modeling study, we should sound a note of caution. Possibly, alternative neuronal mechanisms may yield the equivalent effects. For example, the complex nicotine effect on multiple neuromodulatory systems may perturb the delicate balance of signals in the VTA and explain both positive and negative immediate hedonic consequence as reported in ref. 22. Furthermore, nicotine-modulated DA may have different roles in the nucleus accumbens shell vs. the nucleus accumbens core (25), an issue we leave aside for the moment. In our abstract model of the VTA we focus explicitly on medium- and long-term nicotine effects mediated through DA signals, while the details of acute nicotine effects [possibly mediated by GABAergic non-DA reward signals (22)] remain an important and complex issue for future investigation. Although we may discover that some of the particular mechanisms proposed herein may require modification, the basic computational framework is expected to remain true.

Our hypothesis unites the major classes of previously discussed theories. By modeling the computational consequences of neuromodulatory effects of nicotine on the DA reward-related pathways (2) and synaptic plasticity (48, 49), the model combines the DA-dependent sensitization of behavior that is necessary for the initial acquisition of drug-taking (37) with a drug-evoked opponent process (6, 50) that leads to a persistent drug-related addictive "habit." Hence, the framework is compatible with the ventral–dorsal progression to habits proposed by Everitt *et al.* (51) and recently reported in ref. 52. Here opponency acts not to initiate the drug taking (as suggested by ref. 6) but to cement down the persistence and hedonic independence of the addictive habit. Furthermore, the "positive DA/opponent process" hypothesis of our model may equally apply to other drugs, such as cocaine, that after prolonged exposure provoke compulsive addictive behavior that is independent of its "value" and robust to aversive conditioning (53, 54).

Finally, how nicotine-taking escapes from voluntary control of behavior remains a key issue to investigate. A cognitive control deficit may be the principle cause of the apparent compulsivity and the long-term relapses to smoking. Such top-down control may be mediated by a reciprocal prefrontal cortex–DA–striatal link, a brain circuit postulated to contribute to the conscious "neuronal workspace" (42, 55). One may envision that a long-term selective depression by nicotine of this loop (13) disconnects A-S from cognitive control and uncovers the compulsive nonconscious aspects of nicotine addiction.

Methods

Neuronal activity in the modules is described by the firing rate U_i on [0:1]. Synaptic coupling is given by weights W_i. Plasticity is modeled as changes in W_i. By convention time constants, reaction rates and random noise are marked by Greek letters. The thresholds are marked by θ.

The DA module is a single neuronal population:

$$\frac{dU_{DA}}{dt} = -U_{DA} + S_{DA}\left\{\sum_i r_i; N(t)\right\},$$

where S is a sigmoid input–output function:

$$S_{DA} = \frac{1}{2}\left(1 + \tanh\left(N(t)\sum_i r_i(t) - \theta_{DA}\right)\right).$$

The r_i is the effect of an action i on the DA signal, from -1 (aversive) to 1 (appetitive). $N(t)$ is the nAChR activation with a gain-modulatory effect on U_{DA}. θ_{DA} is the threshold setting the minimum tonic DA. For neutral actions $r_i \neq 0$ and $< \theta_{DA}$.

In the A-S module, U_{Ai} is the unit activity for each of the action plans. We consider a circuit of two units (two-choice task):

$$\tau_A \frac{dU_1^A}{dx} = -U_1^A + S_A\{w^e_{11}U_1^A - w^i_{12}U_2^A - \theta_A\} + \sigma\xi \text{ and}$$

$$\tau_A \frac{dU_2^A}{dx} = -U_2^A + S_A\{w^e_{22}U_2^A - w^i_{21}U_1^A - \theta_A\} + \sigma\xi.$$

Here, w^i_{12};w^i_{21} cross-plan competition (inhibition) and self-excitation weights w^e_{11};w^e_{22} form closed excitatory loops. S_A is a sigmoid function, identical in form to S_{DA} but with $U_{DA}(t)$ taking the gain-modulatory role. Random input ξ with strength σ plays a crucial role: it leads to random symmetry breaking (network chooses an action plan "at will"). τ_A is the A-S time constant. We use a threshold near 1 to mark the selection of a action and reset units to U_{reset} to start the next trial.

The learning in the action selection excitatory weights is governed by

$$\tau_w \frac{dw^e_{ii}}{dt} = L(\langle U_{DA}\rangle, N(t))[\{U_{DA} - \theta_{wDA}\}H\{U_{Aj} - \theta_{wA}\}].$$

The factors are the phasic DA activity U_{DA}, the activation of an action plan U_{Aj} (postsynaptic signal), tonic DA activity (running average $\langle U_{DA}\rangle$) and nAChR activation $L[\langle U_{DA}\rangle, N(t)] = \kappa[\langle U_{DA}\rangle + N(t)]$. $H()$ is 0 when the argument is negative and equal to the argument otherwise. The weights increase when both U_{DA} and U_{Aj} are above their respective thresholds: θ_{wDA}; θ_{wA}. We take $\theta_{wDA} = \langle U_{DA}\rangle + thr$ and $\theta_{wA} = \langle U_{Aj}\rangle$, ensuring that neural population activity and no phasic DA lead to a decrease in the weights.

We model nAChR signal with three dynamical variables on [0;1]: n is the activation of the nAChRs; s is the up-regulation of the nAChRs; and c is an opponent process. Key to the model are

the time scales for the different processes: n is the fastest and c is the slowest. The kinetic equations for dynamics are

$$\tau_n \frac{dn}{dt} = -\beta_n(c)n + \alpha_n(\text{drug})(1-n),$$

$$\tau_s \frac{ds}{dt} = -\beta s + \alpha_s(n, \text{drug})(1-s), \text{ and}$$

$$\tau_c \frac{dc}{dt} = -\beta_c c + \alpha_c(s, n)(1-s).$$

The forward (α) and backward (β) rates are

$$\alpha_n(\text{drug}) = \frac{1}{2}\{1 - \tanh(\text{drug} - \theta_n)\},$$

$$\alpha_s(n, \text{drug}) = \frac{1}{2}\{1 - \tanh(n + \text{drug} - \theta_s)\},$$

$$\alpha_c(s, n) = \frac{1}{2}\{1 - \tanh(s + n - \theta_c)\}, \text{ and}$$

$$\beta_n(c) = \frac{1}{2}\{1 - \tanh(c - \theta_c)\}.$$

The thresholds θ_i are free parameters set such that in the absence of the drug n is at a small stable value and β_n is non-zero. The time constants are separated by orders of magnitude $\tau_n \ll \tau_s \ll \tau_c$.

The up-regulation increases nAChR number; in the model the total nAChR signal is $N(t) = n(t)s(t)$. Chronic nicotine injection is modeled as $\text{drug}(t) = K$. Phasic nicotine is a double exponential:

$$\text{drug}(t) = \text{dose} \times \sum_{t_i'} e^{a(t-t_i')} e^{-b(t-t_i')},$$

where t_i are the armed-choice times (U_{A1} above the threshold). The doses and the time course of the nicotine are scaled to agree qualitatively with available data (e.g., the nicotine onset on the order of 1 min and the offset of several minutes).

For the chronic nicotine the model ran continuously and no A-S was simulated. Simulated self-administration sessions lasted ≈60 min with timeout periods (≈20 h of simulated time; no A-S was simulated). To calculate response probabilities, response latencies, and response number per session we averaged 20–40 runs.

We thank Peter Dayan and Nathaniel Daw for comments on the manuscript and Philippe Faure, Sylvie Granon, and Uwe Maskos for discussions of this project. Support was provided by College de France, the Gatsby Foundation, the Association pour la Recherche Contre Cancer, and the European Community.

1. Volkow, N. D., Fowler, J. S., Wang, G. J. & Goldstein, R. Z. (2002) *Neurobiol. Learn. Mem.* **78,** 610–624.
2. Di Chiara, G. (2000) *Eur. J. Pharmacol.* **393,** 295–314.
3. Balfour, D. J. (2002) *Curr. Drug Targets CNS Neurol. Disord.* **1,** 413–421.
4. Changeux, J.-P. & Edelstein, S. E. (2005) *Nicotinic Receptors: From Molecular Biology to Cognition* (Odile Jacob, New York).
5. Redish, D. (2004) *Science* **306,** 1944–1947.
6. Koob, G. F. & LeMoal, M. (2001) *Neuropsychopharmacology* **24,** 97–129.
7. Dehaene, S. & Changeux, J.-P. (1991) *Cereb. Cortex* **1,** 62–79.
8. Picciotto, M. R., Zoli, M., Rimondini, R., Lena, C., Marubio, L. M., Pich, E. M., Fuxe, K. & Changeux, J.-P. (1998) *Nature* **391,** 173–177.
9. Champtiaux, N., Gotti, C., Cordero-Erausquin, M., David, D. J., Przybylski, C., Lena, C., Clementi, F., Moretti, M., Rossi, F. M., Le Novere, N., *et al*. (2003) *J. Neurosci.* **23,** 7820–7829.
10. Dani, J. & Heinemann, S. (1996) *Neuron* **16,** 905–908.
11. Changeux, J.-P., Bertrand, D., Corringer, P. J., Dehaene, S., Edelstein, S., Lena, C., Le Novere, N., Marubio, L., Picciotto, M. & Zoli, M. (1998) *Brain Res. Brain Res. Rev*. **26,** 198–216.
12. Di Chiara, G. (1999) *Eur. J. Pharmacol.* **375,** 13–30.
13. Mansvelder, H. D., Keath, J. R. & McGehee, D. S. (2002) *Neuron* **33,** 905–919.
14. Picciotto, M. R. & Corrigall, W. A. (2002) *J. Neurosci.* **22,** 3338–3341.
15. Rahman, S., Zhang, J. & Corrigall, W. A. (2003) *Neurosci. Lett.* **348,** 1–4.
16. Fenster, C. P., Rains, M. F., Noerager, B., Quick, M. W. & Lester, R. A. (1997) *J. Neurosci.* **17,** 5747–5749.
17. Pidoplichko, V. I., DeBiasi, M., Williams, J. T. & Dani, J. A. (1997) *Nature* **390,** 401–404.
18. Buisson, B. & Bertrand, D. (2002) *Trends Pharmacol. Sci.* **23,** 130–136.
19. Sallette, J., Bohler, S., Benoit, P., Soudant, M., Pons, S., Le Novere, N., Changeux, J.-P. & Corringer, P. J. (2004) *J. Biol. Chem.* **279,** 18767–18775.
20. Marks, M. J., Grady, S. R. & Collins, A. C. (1993) *Pharmacology* **266,** 1268–1276.
21. Rice, E. M. & Cragg, S. J. (2004) *Nat. Neurosci.* **7,** 583–584.
22. Laviolette, S. R. & van der Kooy, D. (2004) *Nat. Neurosci.* **19,** 3033–3041.
23. Armitage, A. K., Hall, G. H. & Morrison, C. F. (1968) *Nature* **217,** 331–334.
24. Benowitz, N. L. (1999) *Prim. Care* **26,** 611–631.
25. Balfour, D. J. (2004) *Nicotine Tob. Res.* **7,** 899–912.
26. Goldberg, S. R., Spealman, R. D. & Goldberg, D. M. (1981) *Science* **214,** 573–575.
27. Corrigall, W. A. & Coen, K. M. (1989) *Psychopharmacology* **99,** 473–478.
28. Donny, E. C., Caggiula, A. R., Mielke, M. M., Booth, S., Gharib, M. A., Hoffman, A., Maldovan, V., Shupenko, C. & McCallum, S. E. (1999) *Psychopharmacology* **147,** 135–142.
29. Maskos, U., Molles, B. E., Pons, S., Besson, M., Guiard, B. P., Guilloux, J. P., Evrard, A., Cazala, P., Cormier, A., Mameli-Engvall, M., *et al*. (2005) *Nature* **436,** 103–107.
30. Pontieri, F. E., Tanda, G., Orzi, F. & Di Chiara, G. (1996) *Nature* **382,** 255–257.
31. Cadoni, C. & Di Chiara, G. (2000) *Eur. J. Pharmacol.* **387,** 23–25.
32. Nisell, M., Nomikos, G. G., Hertel, P., Panagis, G. & Svensson, T. H. (1986) *Synapse* **22,** 369–381.
33. Epping-Jordan, M., Watkins, S. S., Koob, G. F. & Markou, A. (1998) *Nature* **393,** 76–79.
34. Reynolds, J. N. & Wickens, J. R. (2002) *Neural Networks* **15,** 507–521.
35. Montague, P. R., Dayan, P. & Sejnowski, T. J. (1996) *J. Neurosci.* **16,** 1936–1947.
36. McClure, S. M., Daw, N. D. & Montague, P. R. (2003) *Trends Neurosci.* **26,** 423–428.
37. Robinson, T. E. & Berridge, K. C. (2003) *Annu. Rev. Psychol.* **54,** 25–53.
38. Beiser, D. G., Hua, S. E. & Houk, J. C. (1997) *Curr. Opin. Neurobiol.* **7,** 185–190.
39. Dehaene, S. & Changeux, J.-P. (2000) *Prog. Brain Res.* **126,** 217–229.
40. Tunstall, M., Oorschot, D. E., Kean, A. & Wickens, J. R. (2002) *J. Neurophysiol.* **88,** 1263–1269.
41. Usher, M. & McClelland, J. L. (2001) *Psychol. Rev.* **108,** 550–592.
42. Dehaene, S., Kerszberg, M. & Changeux, J.-P. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14529–14534.
43. Ferrari, R., Le Noverre, N., Picciotto, M. R., Changeux, J.-P. & Zoli, M. (2001) *Eur. J. Neurosci.* **15,** 1810–1818.
44. Rada, P., Jensen, K. & Hoebel, B. G. (2001) *Psychopharmacology* **157,** 105–110.
45. King, S. L., Caldarone, B. J. & Picciotto, M. R. (2004) *Neuropharmacology* **47,** 132–139.
46. Granon, S., Faure, P. & Changeux, J.-P. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 9596–9601.
47. Sutton, R. S. & Barto, A. G. (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
48. Dani, J. A., Ji, D. & Zhou, F. M. (2001) *Neuron* **31,** 349–352.
49. Berke, J. D. & Hyman, S. E. (2000) *Neuron* **25,** 515–532.
50. Solomon, R. & Corbitt, J. (1973) *J. Abnorm. Psychol.* **36,** 158–171.
51. Everitt, B., Dickinson, A. & Robbins, T. W. (2001) *Brain Res. Rev.* **36,** 129–138.
52. Porrino, L. J., Lyons, D., Smith, H. R., Daunais, J. B. & Nader, M. A. (2004) *J. Neurosci.* **24,** 3554–3562.
53. Vanderschuren, L. J. & Everitt, B. J. (2004) *Science* **13,** 951–953.
54. Deroche-Gamonet, V., Belin, D. & Piazza, P. V. (2004) *Science* **13,** 1014–1017.
55. Dehaene, S. & Changeux, J.-P. (2005) *PLOS Biol.* **3,** e141.