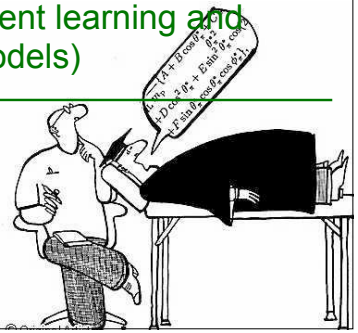
Computational Psychiatry Rising

(on the shoulder of reinforcement learning and decision making models)



What are the big problems that neuroscience could solve?

- mood disorder (Depression, Bipolar ..) : ~ 10% of the population (at some point in life) in US
- anxiety disorder (Panic, OCD, PTSD): ~18% of the population
- addiction: alcohol ~ 10% of the population (at some point in life)
- eating disorder (Anorexia, Bulimia): ~ 4 %
- ADHD: ~4 % (adults)

[NIMH]

- drugs often work poorly
- precise mechanisms of action unknown
- computational neuroscience very poorly represented in psychiatry in the past (often not at all)

-- partly due to nomenclature of psychiatric diseases based on qualitative concepts, incompletely tied to neuroscientific foundations

New hopes

- but this is changing.

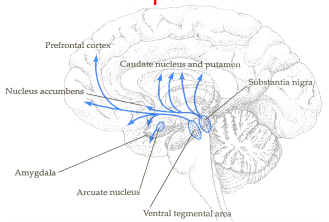
- a new approach: seek a firmer foundation of the science of **decision making**

- **pioneers :** P. Dayan, Q. Huys, T. Braver., J. Cohen, M. Frank, S. Kapur, R. Montague, D. Pizzagali, K. Stephan, D. Steele, J. Williams, D. Redish and others ...

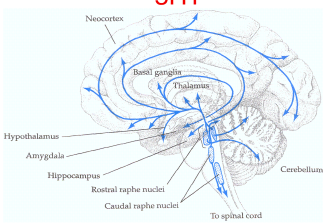
- **"**hope of a specific and quantitative anatomy of normal and abnormal function along with the prospect of rigorous tests for each underlying defect".

The 4 Main Neuromodulators

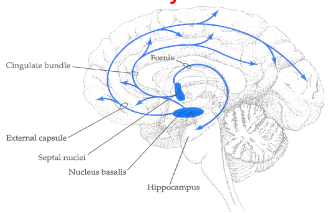
dopamine



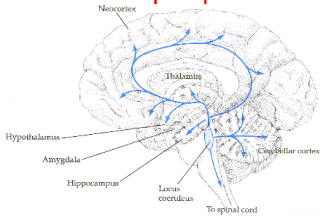
5HT



acetylcholine



norepinephrine



The 4 Main Neuromodulators: critically involved in Major Psychiatric Disease

- Dopamine (DA) involved in Parkinsons', Schizophrenia, Addiction,
- Serotonin (5HT) involved in Depression, OCD, Eating disorders
- Acetylcholine (ACh) involved in Alzheimer's Disease
- Norepinephrine (NA) involved in ADHD, Depression

Yet How Neuromodulation influences Neural Activity is very poorly understood.

Drug Addiction as abnormal decision making

Addiction

A chronically relapsing disorder that is characterised by :
(i) compulsive drug seeking and taking
(ii) inability to limit the intake of drugs
(iii) emergence of a withdrawal syndrome during cessation of drug taking

Goal of neuroscience: understand the cellular & molecular mechanisms that mediate transition between occasional controlled drug use and loss of behavioural control over drug seeking and taking

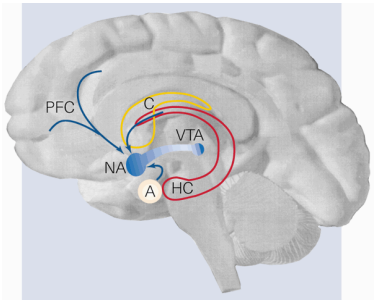
a promising field for modeling, building on models of decision making and reinforcement learning.

Systems involved: the reward system

- mesolimbic dopaminergic system - increase of dopamine release
- mesolimbic DA system: originates in the ventral tegmental area (VTA) of the midbrain, and projects to the nucleus accumbens (NA - ventral striatum). The amygdala (A), hippocampus (HC) and medial prefrontal cortex (PFC) send excitatory projections to the nucleus accumbens.
- drug seeking behaviour induced by glutamatergic projections from the prefrontal cortex to the NAc.

Table 1. Neurobiological Substrates for the Acute Reinforcing Effects of Drugs of Abuse

Drug of Abuse	Neurotransmitter	Sites
Cocaine and amphetamines	Dopamine Serotonin	Nucleus accumbens Amygdala
Opiates	Dopamine Opioid peptides	Ventral tegmental area Nucleus accumbens
Nicotine	Dopamine Opioid peptides?	Ventral tegmental area Nucleus accumbens Amygdala?
THC	Dopamine Opioid peptides?	Ventral tegmental area
Ethanol	Dopamine Opioid peptides Serotonin GABA Glutamate	Ventral tegmental area Nucleus accumbens Amygdala



Why making a maladaptive choice over and over again? Theories of addiction

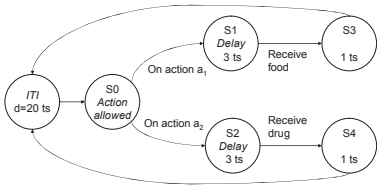
- In the past 30 years, lots of theories
- e.g.

- compulsion zone : self administration is automatically induced when brain cocaine levels within a specific range.
- set point model (or allostasis): goal = adjust sensitivity of brain reward system to set level, by increasing tonic dopamine
- opponent process theory: drug addiction = result of emotional pairing between pleasure and symptoms of withdrawal. Motivation is first related to pleasure, and then to relief from withdrawal.
- impulsivity.

- recently, addiction as a vulnerability in the decision process -- inspiration from reinforcement learning

TD learning -- 101

- world is made of states, actions and rewards; actions are selected so as to maximize future rewards.
- states are associated with value functions defined as expected future reward



$$V(t) = \int_t^{\infty} \gamma^{\tau - t} E[R(\tau)] d\tau \qquad (1)$$

- Goal of TD learning : correctly learn the values. To do this, iteratively use the difference between expected and observed change in value -- the prediction error:

$$\delta(t) = \gamma^d [R(S_l) + V(S_l)] - V(S_k) \qquad (2)$$

- Value is then updated using:

$$V(S_k) \leftarrow V(S_k) + \eta_V \delta$$

- Once the value correctly predicts the reward, learning stops.
- a powerful learning algorithm in machine learning

Phasic dopamine signals prediction error

- the "largest success of computational neuroscience" [Niv]

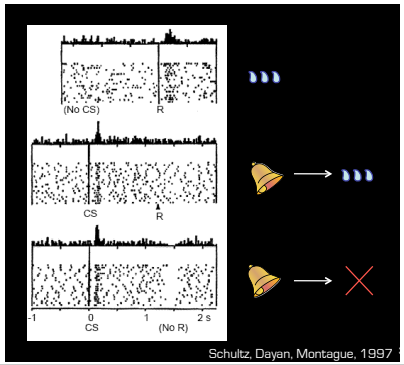
- Monkeys underwent simple instrumental or pavlovian conditioning
- disappearance of dopaminergic response at reward delivery after learning, in VTA and SN.
- if reward is not presented, response depression below basal firing at expected time of reward.

http://www.sciencemag.org • SCIENCE • VOL. 275 • 14 MARCH 1997

A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague*

The capacity to predict future events permits a creature to detect, model, and manipulate the causal structure of its interactions with its environment. Behavioral experiments suggest that learning is driven by changes in the expectations about future salient events such as rewards and punishments. Physiological work has recently complemented these studies by identifying dopaminergic neurons in the primate whose fluctuating output apparently signals changes or errors in the predictions of future salient and rewarding events. Taken together, these findings can be understood through quantitative theories of adaptive optimizing control.



Schultz, Dayan, Montague, 1997

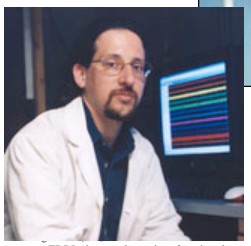
Redish's (*Science,* 2004) model

REPORTS

Addiction as a Computational Process Gone Awry

A. David Redish

Addictive drugs have been hypothesized to access the same neurophysiological mechanisms as natural learning systems. These natural learning systems can be modeled through temporal-difference reinforcement learning (TDRL), which requires a reward-error signal that has been hypothesized to be carried by dopamine. TDRL learns to predict reward by driving that reward-error signal to zero. By adding a noncompensable drug-induced dopamine increase to a TDRL model, a computational model of addiction is constructed that over-selects actions leading to drug receipt. The model provides an explanation for important aspects of the addiction literature and provides a theoretic viewpoint with which to address other aspects.



- cocaine and other drugs produce a transient increase in dopamine
- idea: this dopamine surge induce an increase in prediction error δ that can't be compensated by changes in values.

$$\delta = \max\{\gamma^d [R(S_l) + V(S_l)] - V(S_k) + D(S_l),\ D(S_l)\}$$

where $D(S_l)$ indicates a dopamine surge occurring on entry into S_l.

Consequence: values of states leading to the drug increase without bound.

Redish's (2004) model

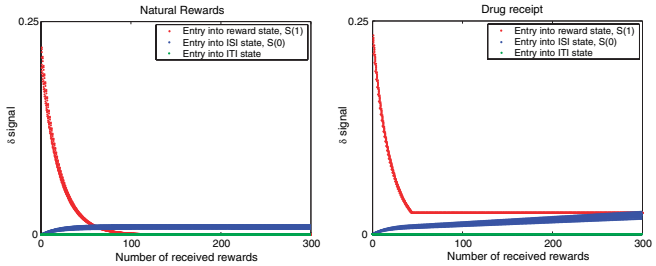


Fig. 3. Dopamine signals. **(Left)** With natural rewards, dopamine initially occurs primarily at reward receipt (on entry into reward state S_1) and shifts to the conditioned stimulus [on entry into interstimulus-interval (ISI) state S_0] with experience. (State space is shown in fig. S7.) **(Right)** With drugs that produce a dopamine signal neuropharmacologically, dopamine continues to occur at the drug receipt (on entry into reward state S_1) even after experience, as well as shifting to the conditioned stimulus (on entry into ISI state S_0), thus producing a double dopamine signal.

- drug is hijacking the learning pathways, creating a prediction error where there should be none.

Redish's (2004) model: predictions

- With repeated experience, drug choice become 1) less sensitive to alternative non drug reinforcers [some evidence]; 2) more inelastic to costs [confirmed]

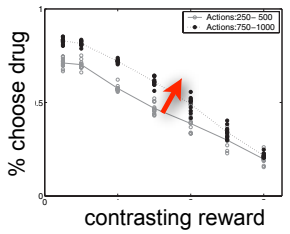


Fig. 1. Probability of selecting a drug-receipt pathway depends on an interaction between drug level, experience, and contrasting reward. Each line shows the average probability of selecting the drug-receipt pathway, $S_0 \xrightarrow{a_2} S_2$, over the contrasting reward pathway, $S_0 \xrightarrow{a_1} S_1$, as a function of the size of the contrasting reward $R(S_3)$. (State space is shown in fig. S1.) Drug receipt on entering state S_4 was $R(S_4) = 1.0$ and $D(S_4) = 0.025$. Individual simulations are shown by dots. Additional details provided in (*14*).

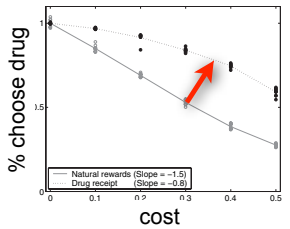


Fig. 2. Elasticity of drug receipt and natural rewards. Both drug receipt and natural rewards are sensitive to costs, but natural rewards are more elastic. Each dot indicates the number of choices made within a session. Sessions were limited by simulated time. The curves have been normalized to the mean number of choices made at zero cost.

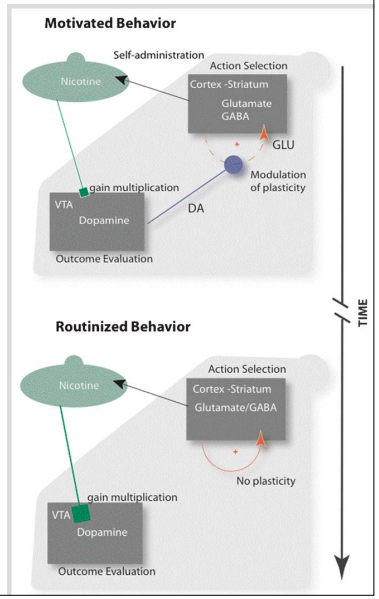
Gutkin, Dehaene & Changeux (*PNAS*, 2006) model of nicotine addiction

- a circuit model, 3 time scales

- Nicotine, through action on nACHRs in VTA, evokes phasic DA signal and changes the gain of DA signaling: potentiates DA transmission.

- The phasic DA instructs learning of action selection. Tonic DA gates this process.

- Slow onset opponent process decrease tonic DA neurotransmission to the point that extinction learning and response unlearning is impaired: routinized/ rigid behavior.



Models of Addiction: Conclusions

- Redish's model, extensions and RL framework
  --> a new generation of models and model-driven experiments.

Lots of remaining challenges:

- addiction to ordinary rewards such as fatty foods, which unlike cocaine produce a dopamine signal that can be accommodated
- addiction to non-stimulant substances which depend less on mesolimbic dopamine (e.g. alcohol)
- describing withdrawal symptoms -- opponent mechanisms
- why do people want to get sober?
- why do people relapse?; accounting for effect of stress.
- vulnerability: only a minority of people become addicted -- while other people can enjoy casual use, why?

Serotonin, Inhibition and Negative Mood

P. Dayan & Q. Huys (2008)

Serotonin - 5-HT

- role in normal and abnormal function still mysterious
- involved in prediction of aversive events (opponent of dopamine which would be related to prediction of reward)
- involved in behavioral inhibition
- involved in models of depression and anxiety:

i) depleting 5-HT by dietary depletion of precursor tryptophan can re-instate depression

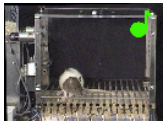
ii) selective serotonin re-uptake inhibitors(SSRIs) = antidepressant

iii) but constitutive decreases in efficiency of 5HT re-uptake is a risk factor for depression.

Idea

- idea: inhibition is directly associated with aversive predictions.
- Prediction of a sufficiently distant threat leads to inhibition, in the form of withdrawal and disengagement (as in conditioned suppression)

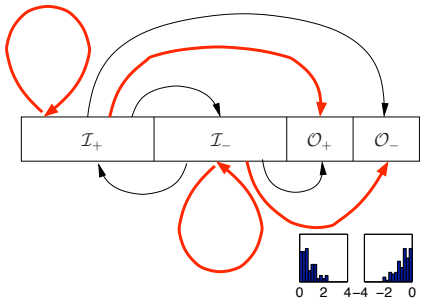
http://go.owu.edu/~deswartz/procedures/conditioned_suppression.html



- 5-HT terminates trains of thought that have a negative value

Model

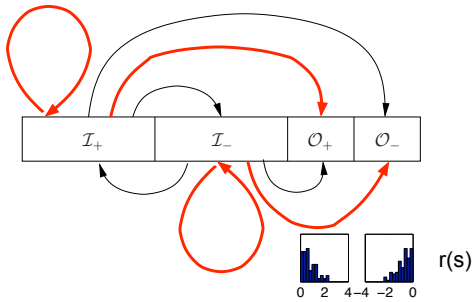
- a model of trains of thoughts
- belief= state
- thought = change of belief = action
- thoughts gain value through their connections with a group of terminal states O+/O- that are assigned + or - affective values



Model

- O+ and O- (each with 100 elements) are associated with value r(s)
- I+ and I- (400 elements) are internal states
- sparse connections between states

- A fixed policy π^0 defined the transition probabilities from one state to the next.
- Internal states will acquire value through (TD reinforcement) learning.



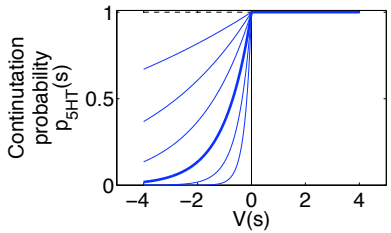
Idea

- 5-HT terminates trains of thought that have a negative value

- Probability of continuing a train of thought depends on V(s)

$$p_{5\mathrm{HT}}(s) = \min\left(1, \exp\left(\alpha_{5\mathrm{HT}} V(s)\right)\right)$$

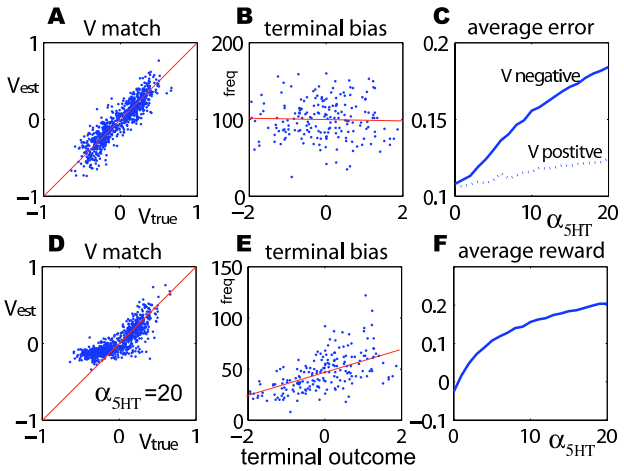
- When thoughts are terminated, they stop and restart randomly in I+ or I-.
- Consequence: the more the 5HT the less the 'negative' states are explored -- sampling bias



Values after learning

$\alpha_{5HT} = 0$

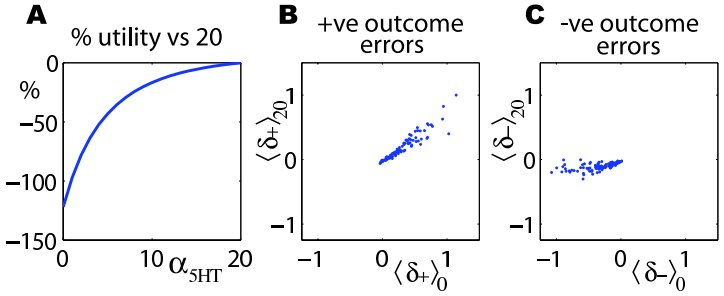
$\alpha_{5HT} = 20$



- 5HT is favorable - enhanced average rewards
- but values are overly optimistic and errors for aversive chains (overvalued)

Serotonin (via Tryptophan) depletion after learning

- after learning, switching α_{5HT}=20 to <20
- suddenly more negative states become explored
- -> more negative average affective outcome
- surprises (errors) associated with transitions that were previously inhibited



Discussion

- 5HT is favorable - enhanced average rewards
- but values are overly optimistic and errors for aversive chains

- consistent with the fact that 5-HT suppression leads to impulsivity (choosing states that would not be selected otherwise)
- consistent with the idea that 5-HT is related to prediction of aversive outcomes

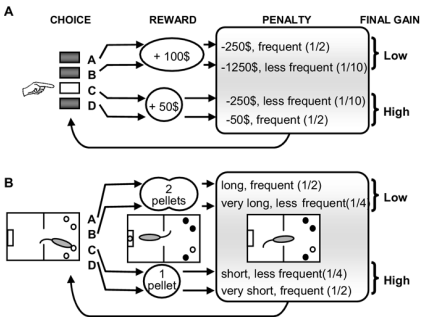
- consistent with the fact that 5-HT depletion after learning leads to depressive symptoms.

- predictions: 5-HT levels during learning would control the extent to which negative states are explored / learned.

- dopamine and serotonin: mutual opposition model. serotonin proposed to report negative prediction errors

(Vincent Valton in Bordeaux) Rat Gambling Task

- decision making in rats using adapted version of Iowa Gambling task.
- large inter-individual differences
- poor decision making results from hypersensitivity to reward and higher risk taking
- TD modeling



[Rivalan et al, *Biol Psych*, 2009]

Decision-making Priors

Optimism: a prior on the likelihood of future reward ?

- "Optimism : the extent to which people hold generalised favourable expectancies for the future"
- the LOT-R questionnaire.

1) In uncertain times, I usually expect the best
2) ~~It's easy for me to relax~~
3) If something can go wrong for me it will
4) I'm always optimistic about my future
5) ~~I enjoy my friends a lot~~
6) ~~It's important for me to keep busy~~
7) I hardly ever expect things to go my way
8) ~~I don't get upset too easily~~

0 = strongly disagree
1 = disagree
2 = neutral
3 = agree
4 = strongly agree

Questions

- Are people usually biased in estimating probability of future reward?
- is this bias correlated with the LOT-R score?
- Can this bias be described as a Bayesian Prior?

[Stankevicius, Kalra, Huys, Seriès, Plos Comp Biol. in press]

Bayesian Model

- Assume subjects are Bayesian Optimal. Based on observed data D and their prior belief, they form the posterior p(c|D)

$$p(c_i|\mathcal{D}_i) = \frac{p(\mathcal{D}_i|c_i)p(c_i)}{p(\mathcal{D}_i)}$$

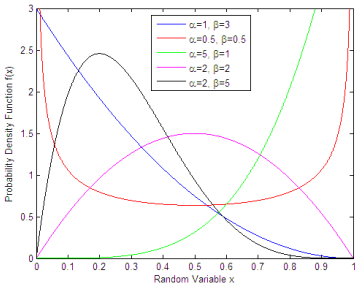
- Parametrisation of prior p(c): Beta distribution, parameters α and β.
- Subjects form estimate of c using the mean of the posterior

$$\hat{c}_i = \int_0^1 c_i p(c_i|\mathcal{D}_i) dc_i \qquad \hat{c}_i = \frac{n_i + \alpha}{N_i + \alpha + \beta}$$

- Subjects make decision based on comparing c with b = probability of reward of square.

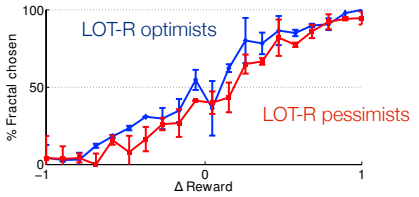
$$p(\text{choose fractal}) = \frac{\exp(\gamma \hat{c})}{\exp(\gamma \hat{c}) + \exp(\gamma b_t)}$$

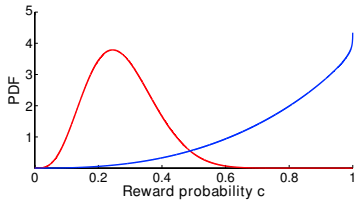
- Each subject is described by 3 parameters (α, β,γ).

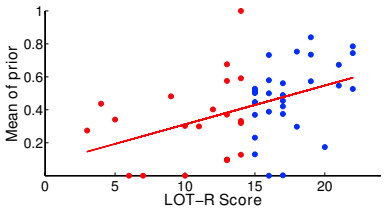


Optimists overestimate probability of future reward

- 51 subjects
- People show significant bias in estimation of probability of future reward.
- ML estimation --> estimation of the prior for each participant.
- LOT-R scores correlate with mean of the prior (r=0.438, p<0.001).

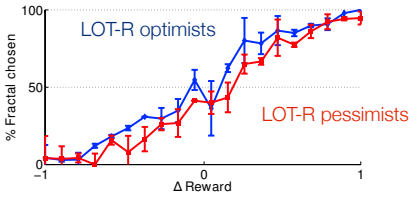


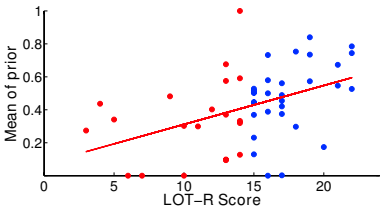




Optimists overestimate probability of future reward

- a new quantitative & behavioural measure of some aspects of optimism
- applications in Depression
- applicable to other personality traits





Conclusion : Computational Psychiatry - new hopes

- new hopes

- a new approach: seek a firmer foundation of the science of **decision making**

- **pioneers :** P. Dayan, Q. Huys, T. Braver., J. Cohen, M. Frank, S. Kapur, R. Montague, D. Pizzagali, K. Stephan, D. Steele, J. Williams, D. Redish and others ...

- "hope of a specific and quantitative anatomy of normal and abnormal function along with the prospect of rigorous tests for each underlying defect".

- interesting times.

Thanks !

This is the end of CCN lectures

The idea of a continuum between health and disease

- decision making in rats
- There is a variability of performance comparable to that in humans -- extreme behaviour could correspond to disorder
- no need for dedicated animal model
- electrophysiology

Dimensional Analysis of ADHD Subtypes in Rats

Candice Blondeau and Françoise Dellu-Hagedorn

Background: Attention-deficit/hyperactivity disorder is a heterogeneous disorder that is classified into three subtypes in which the main symptoms, inattention, hyperactivity, and impulsivity, are expressed with various degrees of severity. The nature of the biological dysfunction sustaining each subtype (common or distinct) is unknown, and animal models encompassing different subtypes are needed.

Methods: A cluster analysis separated subgroups of rats on the basis of similarities in both impulsivity and attentional scores in the five-choice serial reaction time task. These subgroups were characterized behaviorally and were compared for several aspects of spontaneous hyperactivity in different environmental contexts. The dose effects of two agents used clinically (methylphenidate and atomoxetine) were tested on attention and impulsivity.

Results: Four distinct subgroups were demonstrated: efficient, middle, inattentive, and inattentive–impulsive. Hyperactivity expressed in a cage, characterized the last subgroup. Subgroups were differentially sensitive to environmental and pharmacologic challenges. Methylphenidate increased impulsivity mainly in the combined subgroup, whereas atomoxetine decreased impulsivity, neither with any effect on the efficient subgroup and on accuracy.

Conclusions: This new approach is the first to demonstrate behavioral subtypes in rats that parallel those observed in human beings and is a promising tool to clarify the biological bases of these behavioral subtypes and to explain therapeutic effects.

Key Words: Atomoxetine, attention, cluster analysis, hyperactivity, impulsivity, methylphenidate

Attention-deficit/hyperactivity disorder (ADHD), the most common behavioral disorder of childhood, is heterogeneous: symptoms of impulsivity, hyperactivity, and inattention are expressed with various degrees of severity (American ADHD addressing the main symptoms and reflecting distinct subtypes of the disorder could be of great interest.

Several animal models of ADHD have been developed, mainly on the basis of selected strains or experimentally modified animals, each reflecting either separate or combined symptoms of impulsivity, inattention, and hyperactivity (Davids *et al.* 2003; Sagvolden *et al.* 2005). However, these models fail to