



**HAL**  
open science

# Nicotine-induced modifications in value-based decision-making

Marie Louise Dongelmans

► **To cite this version:**

Marie Louise Dongelmans. Nicotine-induced modifications in value-based decision-making. Neurobiology. Université Pierre et Marie Curie - Paris VI, 2017. English. NNT : 2017PA066594 . tel-01924871

**HAL Id: tel-01924871**

**<https://theses.hal.science/tel-01924871>**

Submitted on 16 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT DE  
L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité Neurosciences

Ecole doctorale Cerveau–Cognition–Comportement

Présentée par

Mlle. Marie Louise DONGELMANS

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de thèse:

**Nicotine-induced modifications in value-based decision-making**

Soutenue le 3 Octobre 2017

devant le jury composé de :

Dr. Catherine Le Moine	Rapporteur
Dr. Stéphanie Caillé-Garnier	Rapporteur
Dr. Karim Benchenane	Examineur
Dr. Magalie Lenoir	Examineur
Dr. Bruno Giros	Président du jury
Dr. Philippe Faure	Directeur de thèse

---

Université Pierre & Marie Curie - Paris 6  
Bureau d'accueil, inscription des doctorants  
et base de données  
Esc. G, 2<sup>e</sup> étage  
15 rue de l'école de médecine  
75270 - PARIS CEDEX 06

Tél. Secrétariat : 01 44 27 23 95  
Fax : 01 42 34 68 40  
Tél. pour les étudiants de A - EL : 01 44 27 28 05  
Tél. pour les étudiants de EM - MON : 01 44 27 28 05  
Tél. pour les étudiants de MOO - Z : 01 44 27 28 02  
E-mail : [scolarite.doctorat@upmc.fr](mailto:scolarite.doctorat@upmc.fr)



# Acknowledgements

I am very grateful to Catherine Le Moine and Stéphanie Caillé-Garnier for accepting to evaluate my manuscript and for the helpful comments, as well as Bruno Giros, Karim Benchenane and Magalie Lenoir for accepting to take part in my thesis jury.

I would also like to thank Hervé Chneiweiss for hosting me in the Neuroscience Paris Seine laboratory the last few years.

A very big thank you to Philippe Faure, my supervisor, for giving me the opportunity to start a project in your team. Thank you for caring relentlessly, breaking the insurmountable into pieces, but above all for your frankness and your critical and insightful approach to science that is truly appreciated.

To all my other colleagues, thanks Steve for teaching me *Argo* and spirit, and the very beginnings of in vivo electrophysiology in freely moving mice. Samir, my old neighbour, thanks for always being incredible helpful and cheerful. Nicolas for your true kindness, patience, impressive database and coding knowledge. Jeremie, for trying to teach me computational neuroscience, for harvesting from your intellect and for being witty in all aspects of life. Fabio, for the fabulous unexpected stand up performances and your incredible patience and passion to share knowledge. Alex, for recommending obscure music and having one foot outside France. Stefania for introducing me to the world of nicotine solutions and for all the wonderful delicatessen. Damien, for your limitless resourcefulness and your will to help others with it.

To all the PhD students in the team, I would like to wish you the best of luck with the continuation of your theses and thanks for helping me throughout the last months of writing: Romain, thanks for your tremendous care, passionate views on life and moreover teaching me the basics of optogenetics; Sarah, for your big heart, behind your joyful appearance; Claire for your stimulating dedication to make everything better. Elise, thank you for your honest sweetness. Carole and Sebastien thanks for setting me up in

the lab, for the first hand Batiment B rumours and helping me through the marvels of French administration. Also a big thank you to the interns I had the pleasure to work with: Etienne thank you for bringing my project to another level, including the person that was leading it. Raphael and Amelie thanks for your contribution to my projects and your excitement about in vivo electrophysiology. Nicolas Debray, thanks for your never-ending optimism and building the 3D bandit task.

Bernadette, merci d'avoir surveillée mes souris quand je n'étais pas la et pour ton immense dévouement en general. Sylvie, je te remercie pour tous ce que tu fais pour ce laboratoire, sans toi, tout serait plus compliqués, et merci aussi pour les coups de mains que tu m'as donné.

And I would like to thank all other great minds on the several floors of *la barre Cassan* for the great times during lunch, coffee breaks and late hours. Thank you Salah El Mestikawy, Vincent Vialou, François Tronche, Layal, Ana Clara, Camille, Marine, Soumee, Oriane, Laura, Quentin, Tom, Benjamin, Juliette, Sophie, Gabrielle, Michael, Alexandre, Julie, Tom Watson, Mehdi, Arturo and Marina.

My PhD project would also not have been possible without my acceptance into École des Neurosciences de Paris. Thank you Yvette, Laura, Laetitia, Laure, Patricia and Laurent for all the opportunities which we have been given during our PhDs and for the warm welcome to Paris at the very beginning, as well a big thank you to the ENP crowd, Jaime, Darinka, Valerie and Andrea with whom I started the PhD adventure. Similarly, I would like to acknowledge Labex Biopsy and FRM for financially supporting me during the past 4 years.

Finally I would like to say thank you to my dear Parisian friends Suresh, Heike, Samah and Gabo for their incredible take on life. For being a constant reminder that the best is yet to come. My fabulous "colocs" at Evangile over the years: Margaux, Elise, Nicolas, Marlène, Maïte and Rune for the true introduction to La Vie Française; De lieverds thuis, Micha, Linde, Lea, Sophie. Ik draag jullie overal mee naar toe. Mijn lieve ouders die altijd voor mij klaar staan als ik nog niet eens weet dat het nodig is. Mijn reislustige zus, Tessa, die ik al veel te lang niet heb gezien, maar liefde kent geen afstand. And finally, Jamie for keeping me awake with your comical late night sketches; every time I had to catch a deadline, for introducing me to the devil called cinema, but most of all for your unlimited care and love.

# Abstract

Nicotine addiction is a major societal dilemma: despite the well-known adverse consequences to health, more than one billion persist in the habit. Much addiction research focuses on the motivation to obtain and take drugs and the difficulties of abstinence due to withdrawal effects and learned behaviour. My research focuses on the interactions between the dopaminergic and nicotinic systems in value-based decision-making behaviour, as well as the effects of passively administered nicotine within this framework. Understanding the alterations in the dopamine (DA) system and choice behaviour occurring with drug consumption provides another angle on the addictive state, the general predispositions and vulnerabilities for comorbid disorders. This knowledge could unveil new perspectives towards addiction treatment.

Theories of decision-making suggest that individuals analyse potential benefits and costs to guide their actions based on experience. A multitude of studies have drawn links between DA cell activity and such choice behaviour. Dopamine signals attribute to the evaluation of available options to select future actions. The addictive nature of nicotine is well established, but can a much more alarming link be made between its effect on the DA system and fundamental perception of our environment? Could basic computations underlying choices be significantly modified by drug exposure and/or by the manipulation of nicotinic receptors?

We designed a behavioural paradigm that allows us to assess value-based decision-making in different mouse models and with different manipulations of the cholinergic-dopaminergic circuits. We investigated the role of nicotinic transmission in the exploration/exploitation tradeoff using a  $\beta 2$  subunit-containing nicotinic acetylcholine receptor knockout mouse model. Using a spatial version of a three-armed bandit task and intracranial self-stimulation as reward, we demonstrate that  $\beta 2$ KO mice explore less than the WT. This finding suggests the role of specific nAChRs in the translation of expected uncertainty into motivational value and exploratory decision-making.

Secondly, we analysed the effect of chronic nicotine exposure in this decision process. It has been proposed that phasic dopamine release could provide a teaching signal necessary for learning, while tonic dopamine levels could influence the implementation of previously learned behaviours. Here we show that long-term nicotine exposure can alter the spontaneous activity of VTA DA cells and therefore change phasic and tonic DA signaling. These modifications translate into an enhanced value-sensitivity; the ability to discriminate between different values. We replicated this effect through optical stimulation of VTA DA neurons on a high tonic frequency. This illustrates that organisms under chronic nicotine focus on higher rewards. This data sheds a new light on the mechanisms underlying drug-induced changes in decision-making.



## Résumé

La dépendance à la nicotine est un problème sociétal majeur. De nombreuses personnes continuent à fumer, malgré des conséquences négatives pour la santé bien connues. La recherche sur la toxicomanie se concentre généralement sur la motivation à obtenir cette drogue, le sevrage, rendu difficile par les symptômes de manque et la persistance des comportements appris. Au cours de cette thèse, je me suis concentrée sur les interactions entre les systèmes dopaminergique et nicotinique dans les comportements de prise de décision. Je me suis particulièrement intéressée aux effets de la nicotine administrée passivement de manière chronique. Mettre en évidence les modifications dans la signalisation dopaminergique, et les altérations des comportements de choix qu'implique la consommation de drogue, permet un autre éclairage sur l'état de dépendance, mais ouvre aussi de nouvelles perspectives sur le traitement des addictions et sur les notions de vulnérabilité et de comorbidité avec d'autres pathologies.

Les théories de la prise de décision suggèrent que les individus analysent les bénéfices et les coûts potentiels pour guider leurs actions. Ainsi, faire des choix appropriés nécessite d'apprendre la valeur des options disponibles à partir de l'expérience. Ce processus reposerait principalement sur l'activité des neurones dopaminergiques. Une question cruciale est de savoir si les processus computationnels sous-jacents aux choix peuvent être modifiés de manière durable par l'exposition à la nicotine et/ou par la manipulation des récepteurs nicotiques. Nous avons développé un paradigme comportemental, de type bandit manchot avec des stimulations intra-crânielles comme récompenses, qui permet d'évaluer la prise de décision basée sur la valeur et l'exploration. Nous avons étudié le rôle de la transmission nicotinique dans l'équilibre comportemental entre exploration et exploitation à l'aide de souris knock-out pour la sous-unité  $\beta 2$  du récepteur nicotinique. Nous démontrons que les souris  $\beta 2$ KO explorent moins que les souris de type sauvage suggérant un rôle des nAChR dans la traduction de l'incertitude attendue en motivation à explorer.

Dans une deuxième étape, nous analysons l'effet d'une exposition chronique à la nicotine sur ces processus de prise de décision. Il a été proposé que la libération phasique de dopamine serait cruciale pour l'apprentissage, alors que l'activité tonique serait plus impliquée dans l'expression d'un comportement précédemment acquis. Nous montrons ici que la nicotine, sur le long terme, peut altérer l'activité spontanée des cellules dopaminergiques et ainsi modifier la libération phasique et tonique de dopamine. Cette modification se traduit par une augmentation de la "sensibilité à la valeur", et donc une altération des choix. Un individu sous nicotine chronique se concentre davantage sur les récompenses plus importantes. Nous avons mimé cet effet par une activation optogénétique tonique des neurones dopaminergiques.

Ces travaux mettent en lumière les mécanismes qui sous-tendent les changements dans la prise de décision lors d'une exposition aux drogues.





# List of Abbreviations

AADC :	Aromatic amino acid decarboxylase
ACh :	Achetylcholine
AP :	Action potential
BNST :	Stria terminalis
CPP:	Conditioned place preference
DA :	Dopamine
DA :	Dopamine neuron
DAT :	Dopamine transporter
DS :	Deterministic setting
GABA :	$\gamma$ -Aminobutyric acid
GAD :	Glutamic acid decarboxylase
GP :	Globus Pallidus
ICSS:	Intra-cranial self-stimulation
$I_h$ :	Hyperpolarisation-activated inward current
IPN :	Interpeduncular nucleus
IPSPs :	Inhibitory postsynaptic potentials
LDTg :	Laterodorsal tegemental nucleus
LH :	Lateral hypothalamus
LHb :	Lateral Habenula
MSNs :	Medium spiny neurons
NAc :	Nucleus accumbens
nAChR :	Nicotinic acetylcholinergic receptor
mAChRs :	Muscarinic acetylcholinergic receptor
OFC:	Orbitofrontal cortex
PET:	Positron Emission Tomography

PFC: Prefrontal cortex  
PPTg : Pedunculopontine tegmental nucleus  
PS : Probabilistic setting  
RL: Reinforcement Learning  
RMTg : Rostromedial tegmental nucleus  
RPE : Reward Prediction Error  
SN : Substantia Nigra  
TANs : Tonically active neurons  
TD: Temporal difference  
TH : Tyrosine hydroxylase  
VMAT2 : Vesicular monoamine transporters type 2  
VTA : Ventral Tegmental Area

# List of Publications

1. Naude, J., Tolu, S., Dongelmans, M., Torquet, N., Valverde, S., Rodriguez, G., Pons, S., Maskos, U., Mourot, A., Mart, F. & Faure, Ph. (2016). Nicotine receptors in the ventral tegmental area promote uncertainty-seeking, *Nature Neuroscience*, 4223, p1-12.
2. Naude, J., Dongelmans, M. & Faure, Ph (2015) Nicotinic alteration of decision-making, *Neuropharmacology*, 96B, p244-254.
3. Parnaudeau, S., Dongelmans, M.L, Turiault, M., Ambroggi, F., Delbes, A-S., Cansell, C., Luquet, S., Piazza, P-V., Tronche, F. & Barik, J. (2014). Glucocorticoid receptor gene inactivation in dopamine-innervated areas selectively decreases behavioral responses to amphetamine. *Frontiers in behavioral neuroscience*, 8:35.
4. Eagle, D.M., Noschang, C., d'Angelo, L-S. C., Noble, C.A., Day, J.O., Dongelmans, M.L., Theobald, D.E., Mar, A.C., Urcelay, G.P., Morein-Zamir, S. & Robbins, T.W. (2014). The dopamine D2/D3 receptor agonist quinpirole increases checking-like behaviour in an operant observing response task with uncertain reinforcement: A novel possible model of OCD. *Behavioural brain research*. 264:207-29.



# Contents

<b>Introduction</b>	<b>3</b>
<b>1 From tobacco use to altered decision-making</b>	<b>5</b>
1.1 The prevalence and risks of tobacco use . . . . .	5
1.2 Nicotine dependence and tobacco use disorder . . . . .	5
1.2.1 Diagnosis of tobacco use disorder . . . . .	6
1.3 The reinforcing cycle of tobacco use conditioning . . . . .	6
1.3.1 Conditioned stimuli . . . . .	7
1.3.2 Self-administration experiments . . . . .	8
1.3.3 Non-contingent nicotine administration . . . . .	8
1.3.4 Cycle of nicotinic receptor availability . . . . .	9
1.4 Reward, reinforcement learning and dopamine . . . . .	9
1.4.1 The reward prediction error . . . . .	11
1.4.2 Time-scaled roles of dopamine . . . . .	11
1.5 Scope of the presented research . . . . .	12
<b>2 The dopaminergic system</b>	<b>15</b>
2.1 Dopaminergic projections . . . . .	16
2.2 Dopamine synthesis and receptors . . . . .	18
2.2.1 Dopamine receptors . . . . .	18
2.3 VTA dopamine function . . . . .	19
2.3.1 Electrophysiological features of VTA DA neurons . . . . .	19
2.3.2 Spontaneous activity of VTA DA neurons . . . . .	20
2.3.3 Origins of burst firing . . . . .	21
2.4 Heterogeneity of dopaminergic neurons in the VTA . . . . .	22
2.5 Optogenetic manipulation of VTA neurons . . . . .	23

<b>3</b>	<b>The nicotinic cholinergic system</b>	<b>25</b>
3.1	Nicotinic acetylcholine receptors . . . . .	25
3.2	nAChRs in the dopaminergic system . . . . .	27
3.3	Nicotinic control of the dopaminergic system . . . . .	28
3.3.1	Acute effects of nicotine . . . . .	28
3.3.2	Chronic exposure to nicotine: a smoker's brain . . . . .	31
3.4	Behavioural implications of nicotine-use . . . . .	32
<b>4</b>	<b>Nicotinic alteration of decision-making: a review</b>	<b>35</b>
<b>5</b>	<b>Objectives</b>	<b>51</b>
	<b>Methods</b>	<b>53</b>
<b>6</b>	<b>Methods and materials</b>	<b>55</b>
6.1	Behavioural paradigms . . . . .	55
6.1.1	Multi-armed ICSS Bandit task . . . . .	55
6.1.2	Closed economy for mice . . . . .	57
6.2	Nicotine and minipump implantation . . . . .	58
6.3	Electrophysiological recordings of dopaminergic neurons in the VTA . . . . .	58
6.3.1	Juxtacellular recordings . . . . .	58
6.3.2	Development of the micro-system . . . . .	59
6.3.3	Surgery . . . . .	60
6.3.4	Data acquisition and treatment . . . . .	60
6.3.5	Spike extraction and sorting . . . . .	61
6.3.6	Dopamine neuron identification . . . . .	61
6.4	$\beta 2$ knock out mice and local re-expression . . . . .	63
6.4.1	$\beta 2$ knock out mice . . . . .	63
6.4.2	Stereotaxic injection of lentivirus . . . . .	63
6.5	Optogenetic methods . . . . .	63
6.5.1	Virus injection and optogenetic experiments . . . . .	63
6.6	Fluorescence immunohistochemistry . . . . .	65
6.7	Overview of performed experiments . . . . .	66

6.7.1	Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking . . . . .	66
6.7.2	Chronic nicotine enhances value sensitivity . . . . .	67
	<b>Results</b>	<b>69</b>
<b>7</b>	<b>Nicotine receptors in the ventral tegmental area promote uncertainty-seeking</b>	<b>71</b>
7.1	Expected uncertainty . . . . .	72
7.2	Exploration-exploitation trade-off . . . . .	73
7.2.1	Bandit-style problems . . . . .	73
7.3	Nicotine receptors in the VTA promote uncertainty-seeking . . . . .	74
<b>8</b>	<b>Chronic nicotine enhances value sensitivity in mice through altered VTA DA activity</b>	<b>87</b>
8.1	Reward sensitivity and nicotine . . . . .	87
	<b>Discussion</b>	<b>129</b>
<b>9</b>	<b>Perspectives</b>	<b>131</b>
9.1	Closed economy for mice . . . . .	131
9.2	Online 3D Human bandit task . . . . .	137
<b>10</b>	<b>Nicotine-induced alterations in value-based decision-making</b>	<b>141</b>





# List of Figures

1.1	DSM-5 criteria for tobacco use disorder . . . . .	7
1.2	A hypothetical cycle for perpetuating nicotine use . . . . .	10
1.3	A schematic representation of the reward prediction error (RPE) . . . . .	11
2.1	A simplified schematic of the major connections to and from the VTA in a mouse brain. . . . .	16
2.2	A simplified representation of the VTA topology and DA neuron identification	17
2.3	VTA DA cell firing . . . . .	20
3.1	Schematic representation of a nicotinic receptor . . . . .	26
3.2	Location of the nAChRs within the ventral tegmental area . . . . .	28
3.3	Nicotine evoked response in VTA DA cells . . . . .	29
4.1	Reinforcement learning principle . . . . .	36
6.1	Multi-armed ICSS bandit task . . . . .	56
6.2	Closed economy paradigm . . . . .	57
6.3	Overview of the implantable electrophysiology micro-drive system . . . . .	59
6.4	Spike time analyses . . . . .	62
6.5	Timeline of performed experiments in Chapter 7 . . . . .	66
6.6	Timeline of the performed experiments in Chapter 8 . . . . .	67
9.1	The effect of nicotine in a closed economy paradigm . . . . .	133
9.2	VTA DA Cells recording in close economy paradigm . . . . .	136
9.3	Online 3D Human Bandit Task . . . . .	138
9.4	Fagerström test . . . . .	140



# Introduction



# 1 | From tobacco use to altered decision-making

## 1.1 The prevalence and risks of tobacco use

The World Health Organization (WHO) reports the tobacco epidemic as one of the biggest public health threats the world has ever faced, killing around 7 million people a year. The WHO estimates that one-third of the global adult population smokes. Approximately half of the smoking population perishes as a direct consequence of tobacco use, another 890,000 lost lives are of non-smokers being exposed to second-hand smoke. Tobacco smoke contains more than 4000 chemicals. At least 250 of those chemicals are known to be harmful and around 70 constituents are known to cause cancer. The causes of death in order of incidence are cardiovascular diseases (43%), all forms of tobacco-caused cancer (36%), respiratory diseases (20%), and other smoking-caused deaths (1%), including those due to passive smoking [WHO, tobacco fact sheet, 2017]. The driving addictive, psychoactive component in tobacco is nicotine, although other tobacco and tobacco smoke constituents may contribute to dependency as well [1–3]. Nicotine's addictive power is reflected by the initiation and cessation numbers of smokers: a third of individuals who have ever tried smoking become a daily smoker and of those smokers who try to quit, less than 5 percent are successful, as abstinence is usually short-lived. In contrast to many other drugs of abuse, the pattern of cigarette smoking is rarely occasional. Whereas only 10-15% of alcohol consumers are considered problem drinkers, approximately 90% of smokers smoke at least 5 cigarettes a day [4].

## 1.2 Nicotine dependence and tobacco use disorder

The mean onset age of cigarette smoking is 13-14 years. Initiation to smoking is often mediated by a variety of social and environmental factors, such as close relatives and peers who smoke, a stressful environment, and being able to afford cigarettes. After a period of frequent use the reinforcing effects of nicotine strengthen. Eventually, the starting smoker

will slowly lose control over its use and develop tolerance and physiological dependence, due to physiological changes in the dopaminergic "reward" system (*see Chapter 2*) and the nicotinic receptors present in this system (*see Chapter 3*). The nicotine dependence syndrome is on average experienced within 23 months of tobacco use onset. Dependence can be illustrated by the experience of withdrawal symptoms after a non-smoking period. The nicotine withdrawal symptoms include craving for nicotine, irritability, frustration, anger, anxiety, difficulty concentrating, decreased heart rate and increased appetite or weight gain. Tolerance simply expresses itself in increased drug intake in order to obtain similar previous drug-induced effects. Tolerance to nicotine has been linked to a decreased responsiveness to the drug at the site of action: an increased number of nicotinic receptors and to some degree an increased metabolism of the drug [5].

### 1.2.1 Diagnosis of tobacco use disorder

A tobacco use disorder can be diagnosed according to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorder (DSM-V) criteria (Table 1., American Psychiatric Association, 2013). The previous DSM IV criteria for substance abuse and substance dependence have been combined into substance use disorders, specific to each substance. Each substance use disorder is divided into mild (2-3 symptoms), moderate (4-5 symptoms) and severe (6 or more symptoms) subtypes. Another dimensional scale to assess the severity of tobacco dependence is the Fagerström Test for Nicotine Dependence [6]. This questionnaire consists of 10 differently weighted questions measuring a smoker's daily intake and dependency (*see Discussion*).

## 1.3 The reinforcing cycle of tobacco use conditioning

Generally, physiological dependence and tolerance are not found to be sufficient to maintain nicotine dependence. Neither do these processes explain how nicotine dependence is established. Theories of nicotine dependence have placed increasing emphasis on the role of stimuli or "cues" associated with the pharmacological effect of nicotine and tobacco use [8,9]. Firstly, smoking as a route of administration creates a very effective basis for forming associations between drug intake and its effect. After inhaling smoke from a cigarette, nicotine is distilled from tobacco and carried in smoke particles into

*A problematic pattern of tobacco use leading to clinically significant impairment or distress, as manifested by at least two of the following, occurring within a 12-month period:*

1. Tobacco is often taken in larger amounts or over a longer period than was intended.
2. There is a persistent desire or unsuccessful efforts to cut down or control tobacco use.
3. A great deal of time is spent in activities necessary to obtain or use tobacco.
4. Craving, or a strong desire or urge to use tobacco.
5. Recurrent tobacco use resulting in a failure to fulfill major role obligations at work, school, or home.
6. Continued tobacco use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of tobacco (e.g., arguments with others about tobacco use).
7. Important social, occupational, or recreational activities are given up or reduced because of tobacco use.
8. Recurrent tobacco use in situations in which it is physically hazardous (e.g., smoking in bed).
9. Tobacco use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by tobacco.
10. Tolerance, as defined by either of the following:
  - A. A need for markedly increased amounts of tobacco to achieve the desired effect.
  - B. A markedly diminished effect with continued use of the same amount of tobacco.
11. Withdrawal, as manifested by either of the following:
  - A. The characteristic withdrawal syndrome for tobacco (refer to Criteria A and B of the criteria set for tobacco withdrawal).
  - B. Tobacco (or a closely related substance, such as nicotine) is taken to relieve or avoid withdrawal symptoms.

Figure 1.1 – **TABLE 1. DSM-5 criteria for tobacco use disorder, American Psychiatric Association (2013) [7]**

the lungs, where it is absorbed rapidly into the pulmonary venous circulation. Here, the specific behaviour, the act of smoking, is paired with the almost instant effects of nicotine being delivered to the brain. The faster the rate of absorption of a drug and entry into the brain, the greater the rush, and the more reinforcing and hence a greater sense of reinforcement of the drug is established. Secondly, smoking allows frequent self-administration which is necessary for its reinforcing effects to condition habitual use and sustain nicotine dependence.

#### 1.3.1 Conditioned stimuli

In parallel with repeated self-administration, the physiological effects of nicotine become associated with various non-nicotine stimuli, environmental factors, and even specific moods. People habitually smoke cigarettes in specific situations such as after a meal, with a cup of coffee or an alcoholic drink, or with friends who smoke. Likewise, aspects of the drug-taking ritual, such as rolling a cigarette become associated with the pleasurable effects of smoking. Furthermore, respiratory tract sensory cues associated with tobacco smoking represent a type of conditioned reinforcer that has been shown to play an important role in the regulation of smoke intake and the craving to smoke, in addition to



the rewarding effects of smoking. Over time these stimuli acquire conditional value or serve as cues for future nicotine delivery. Once dependence is reached, negative affects; stress, irritability and other emotional states that are experienced during withdrawal trigger smoking as well. Thus, smoking can be reinforced by both positive and negative states and associations. [10–12]. As a result, the conditional stimuli for tobacco drive smoking and may furthermore result in lapse/relapse after sustained abstinence. This is supported by the finding that such proximal stimuli - normally associated with smoking - can induce craving in smokers [13].

### **1.3.2 Self-administration experiments**

A large body of animal studies affirm the role of contextual stimuli associations during nicotine administration. When intravenously (I.V.) administered nicotine alone acts as the primary reinforcer in a classic instrumental self-administration paradigm the conditioned effects are modest. Rats lever press significantly less to self-administer nicotine than when the administration of nicotine was coupled with the presentation of a contextual visual stimulus. Interestingly, when the visual stimulus alone was the primary reinforcer the rats would press levers at a similar rate as for nicotine [11, 14–18]. However, different results have been found in mouse studies. Self-administration is more difficult to establish in mice and very dependent on pre-training for a food-reward. In addition, independent pressing for a cue-light has not been reported [19]. In general, self-administration of nicotine is highly dependent on the availability of dopamine [20] and nicotinic receptors [21, 22] and the presence of nicotinic receptors containing subunits with a high-affinity for nicotine [23–27].

### **1.3.3 Non-contingent nicotine administration**

The previous paragraph suggests that contextual stimuli are essential in the initiation and maintenance of nicotine dependence. Interestingly, when the I.V. administration of nicotine was non-contingent and the timing of the infusion was decided by the experimenter (yoked) the nicotine-treated rats still showed an increased response for reinforcing non-pharmacological stimuli compared to control animals. This suggests that nicotine would make one more susceptible to conditioning by enhancing the motivational valence of non-nicotine stimuli. Nicotine could therefore act as a primary reinforcer and as re-

inforcement enhancer [14, 16]. This suggests large behavioural implications for smokers, especially within the scope of neuroeconomics: when making a choice, different alternatives are being valued based on learned expectations through feedback processes. If these values or the motivational valence associated with stimuli are skewed due to drugs or drug-induced modifications in the brain decision-making behaviour will be altered and perhaps not optimal (*see 1.4.1. and Chapter 4*).

### 1.3.4 Cycle of nicotinic receptor availability

On a physiological level tobacco use conditioning is being reinforced by the cycle of nicotinic receptor availability. As mentioned before, nicotine absorption is rapid for cigarettes. Nicotine levels fall quickly because about half of the nicotine is redistributed throughout the body within 15-20 min of the last puff from a cigarette. Further decline is more gradual, with a terminal half-life averaging 2 hr, but highly variable across individuals, because of differences in metabolism. The high concentrations of nicotine in the brain after smoking a cigarette are comparable to those seen after intravenous administration. When nicotine diffuses into the brain tissue, it binds selectively to nicotinic cholinergic receptors (nAChRs), which are ligand-gated ion channels (*see Chapter 3*). These receptors are widely spread through the mammalian brain, but the ones present on the dopaminergic system play an important role in addiction and reward processing. Acute nicotine intake results in nicotinic receptor activation and transients of dopamine that are considered pleasurable. When the nicotine levels in the brain build up over a period of time the constant lower levels of nicotine subsequently desensitise, and in later stage inactivate the receptors. This inactivation and a decrease in nicotine leads to an up-regulation of nAChRs, paired with a hyper-excitability for nicotine and a craving to smoke (Figure 1.1, [28]). The pleasurable effect that is evoked after a cigarette, due to an increase in dopamine, reinforces the smoking behaviour and contextual stimuli. The underlying physiological processes will be discussed in more depth later on.

## 1.4 Reward, reinforcement learning and dopamine

The primary effects of nicotine and the process of reinforcement learning (conditioning) rely on overlapping brain networks with, at its core the dopaminergic system [29]. All

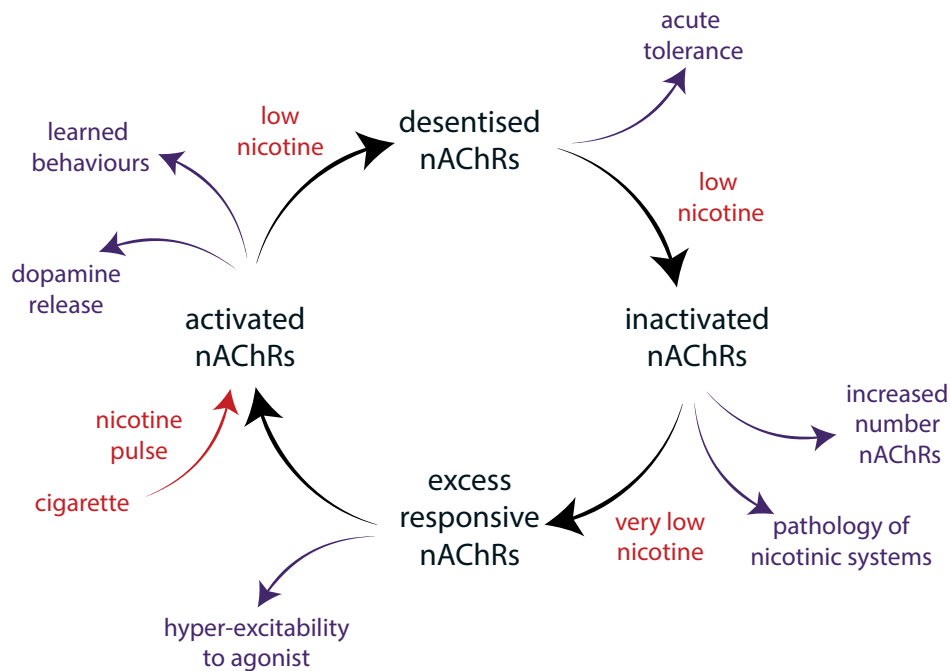


Figure 1.2 – A hypothetical cycle for perpetuating nicotine use proposed by Dani & Heinemann, 1996 [28]

addictive drugs increase the activity of dopamine (DA) neurons in the ventral tegmental area (VTA), which results in increased levels of extracellular dopamine concentrations in the nucleus accumbens (NAc), a projection site of the VTA (*see Chapter 2*, [30, 31]). Phillips and colleagues (2003) demonstrated rapid changes in extracellular dopamine concentrations in the NAc in different stages during a cocaine self-administration trail. After repeated training a DA increase was observed that coincided with the initiation of drug-seeking behaviour in approach to the lever. The concurrent presentation of the cocaine-related cues at the moment of the lever press evoked an even further increase, followed by a drug-induced DA increase. Interestingly, when these cues were presented without a cocaine infusion they also elicited similar rapid dopamine signalling, but only when they were previously associated with the cocaine delivery [32]. Dopaminergic activity is crucial in addiction research, since it is the main player in the first pleasurable effects of a drug up to the development of habitual use. However, it also plays an important role in more general learning and motivational processes, which implies that daily behaviours can be hijacked by drug use. A uniformly accepted theory that fully explains the function of the dopamine signal in behaviour is not yet reached, but controversies have mostly risen from

studies using different experimental techniques [33]. Two aspects of the DA function are discussed below.

### 1.4.1 The reward prediction error

A rich body of evidence obtained from classical conditioning experiments demonstrate that VTA DA activity increases at the presentation of a conditioned stimulus or inhibits in response to an unexpected omitted reward (*see Figure 1.2*). These dopamine fluctuations take place on a small time-scale (seconds). This has been demonstrated in monkeys [?], mice [34], rats [35], as well as in humans [36] and led to the formulation of the reward prediction error (RPE) theory [37]. The RPE is the difference between the expected reward, and the actual perceived reward and approaches zero in a stable learned environment. Computational temporal difference learning theories consider the RPE to be an important teaching signal, contributing to learning [38]. Such a signal would enable constant evaluation of previously presented choice options. However, a drug-induced dopamine increase at the time of evaluation (i.e. expected drug delivery) misrepresents the learning signal which could lead to an over-selection of drug-taking related actions (*see Chapter 4*, [39,40]).

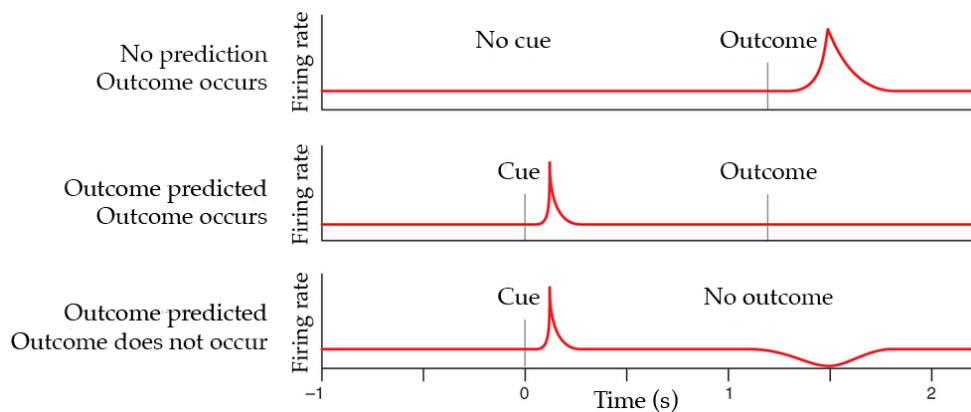


Figure 1.3 – A schematic representation of the reward prediction error (RPE) adapted from Schultz and colleagues, 1997 [41]

### 1.4.2 Time-scaled roles of dopamine

Pharmacological and lesion studies, that basically diminish all (local) dopamine signalling, mostly revealed that the dopamine signal is involved in motivation and action.

Berridge and Robinson have advocated that VTA DA neurons encode an incentive saliency signal. This signal assigns an attractiveness to environmental stimuli, invoking approach and interaction [42]. This concept is supported by a study that shows that animals approaching predictive cues show higher cue-induced NAc DA transients [43]. Other theories argue that the main role of the mesolimbic dopamine signal is assigned to behavioural activation [44] or flexible approach [45]. More recent studies, involving optogenetic, bidirectional manipulations, have revealed that indeed minute-by-minute dopamine levels reflected motivational vigor, while second-by-second dopamine release encoded value [46]. These findings are supported by NAc DA release ramping up towards proximity of an operant response for food [47], as well as spatial proximity to a food reward when navigating through a maze [48]. A better understanding of the specific dopamine-driven behaviours and their timescales will provide a more solid basis for addiction research.

## 1.5 Scope of the presented research

In this manuscript I will focus on the interaction between the dopaminergic and nicotinic systems in value-based decision-making behaviour. I am specifically interested in the effects of passively administered nicotine on value-based decision-making. Addiction research will focus generally on the motivation to obtain and take a drug and the difficulties of abstinence due to withdrawal effects and learned behaviour. However, differentiating the alterations in the dopamine system and choice behaviour from an addictive state sheds another light on addiction treatment, vulnerabilities for comorbid disorders and general predispositions. Non-contingent I.V. administration studies already demonstrated the synergistic effect of nicotine on sensory reinforcement, which suggests that repeated nicotine administration can change the incentive value of a stimulus [49]. However, these studies contribute to the understanding of the maintenance of addiction, but do not give insight on how nicotine use can change a brain state and what consequences this could have on non-addictive related behaviours, such as decision-making. Different neuroeconomical paradigms have already demonstrated a direct link between an altered mesocorticolimbic dopaminergic system and modified choice behaviour due to chronic nicotine [50–52], findings that support the work presented in this manuscript.

*In the following chapters I will first more thoroughly describe the dopaminergic and*

*cholinergic systems in relation to value-based decision-making. In the results section I will present a behavioural paradigm that our team has developed and which allows us to assess value-based decision-making in different mouse models and with different manipulations of the cholinergic-dopaminergic pathways. We demonstrate that indeed the nicotine-induced modifications of decision-making are modulated by the dopaminergic system. I will reflect on these findings in the discussion after presenting findings from ongoing projects.*



## 2 | The dopaminergic system

In 1957, Arvid Carlsson demonstrated that dopamine (DA) was a neurotransmitter in the brain and not just a precursor of norepinephrine. The highest concentration of dopamine was found in the basal ganglia. Antipsychotic drug studies linked a depletion in dopamine to a loss of control in movement, a phenotype that is clinically similar to symptoms seen in Parkinson's disease (PD). Furthermore, when PD patients were treated with L-dopa, a precursor of dopamine, the typical PD symptoms were relieved. However, in healthy subjects L-dopa treatment mimicked symptoms of paranoid schizophrenia. This finding led to the dopamine theory of schizophrenia and the role of dopamine in antipsychotic medications [53]. Around the same time, James Olds and Peter Milner started to experiment with Intracranial self-stimulation (ICSS) in the deeper brain structures and discovered what has been called "the brain's pleasure centers" [54,55]. Operant auto-stimulation of the basal forebrain was displayed to serve as an "operant reinforcer" [56]. Lesion, pharmacological, and anatomical studies eventually uncovered that the rewarding effects of medial forebrain bundle brain stimulation depend on the activation of the mesocorticolimbic dopamine (mesDA) system and that the potency of ICSS can indeed be altered by dopamine agonists and antagonists [57,58]. Dopamine agonists, like amphetamine, reduce the frequency of stimulation needed to provoke lever responding for ICSS [59]. This suggests a similar synergistic effect as observed in the non-contingency studies where a drug increases the motivational and incentive valence of stimuli, but moreover this indicates an overlap between drug effects and reward processing. In this chapter, I will first introduce the different components of the mesDA system with a focus on the ventral tegmental area. Secondly, I will review different behavioural implications, mostly related to reward processing and decision-making, of the mesDA system that have been discovered in the last decennia with a range of experimental techniques and recently confirmed with optogenetic manipulations.



## 2.1 Dopaminergic projections

In the mammalian brain, the mesocorticolimbic system is comprised out of two dopaminergic nuclei: the ventral tegmental area (VTA) and the substantia nigra compacta (SNc). The SNc is a relatively homogenous population of neurons, the majority of which are dopaminergic (90%) and project to the dorsal striatum [60]. In contrast, the VTA contains a mixture of dopaminergic (65-70%), GABAergic (30%) and glutamatergic neurons (2-3%), including co-transmitting neurons synthesising a combination of any two neurotransmitters (*see Figure 2.2b*) that project to the ventral part of the striatum, including the nucleus accumbens (NAc) and portions of the olfactory tubercle [61, 62].

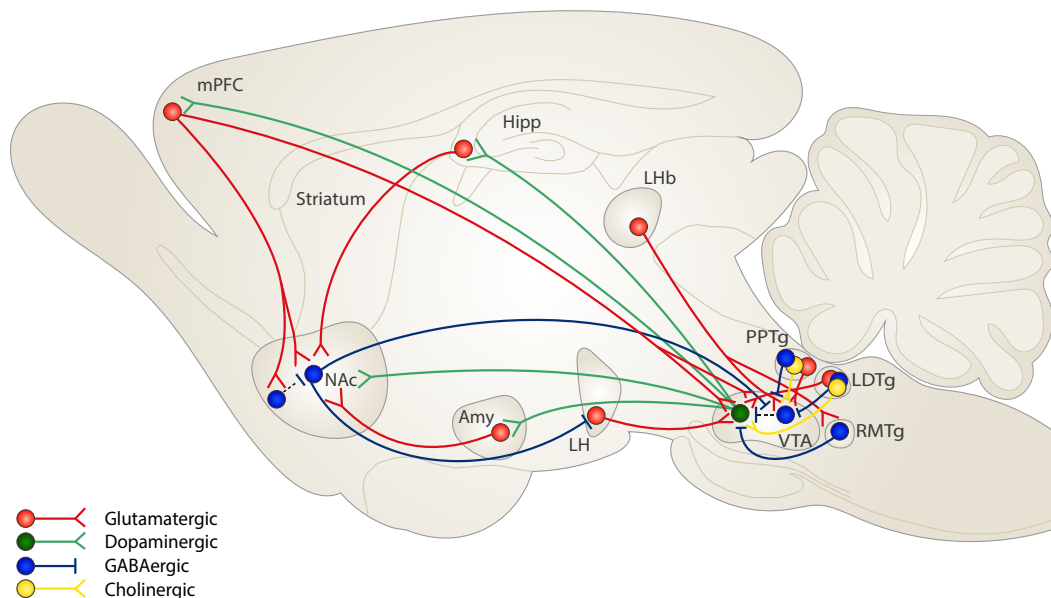


Figure 2.1 – **A simplified schematic of the major connections to and from the VTA in a mouse brain.** This image is adapted from Pistillo et al., 2015; Dautan et al., 2016 and Morales & Margolis, 2017 [63–65]. mPFC: medial Prefrontal Cortex; NAc: Nucleus Accumbens; Hipp: Hippocampus; Amy: Amygdala; LHb: Lateral habenula; LH: Lateral hypothalamus; PPTg: pedunculopontine nucleus; LDTg: Laterodorsal tegmental nucleus; RMTg: rostromedial tegmentum and VTA: ventral tegmental area. The synaptic circuits for the VTA are shown in greater detail in Fig. 2.2.b

The VTA additionally projects to the medial prefrontal cortex (mPFC), hippocampus (Hipp) and amygdala (Amy). Most projection sites also have afferents towards the VTA creating feedback loops, governing VTA’s activity (*see Figure 2.1 & 2.2b*, [62, 66]). Furthermore, the VTA receives inputs from the lateral habenula (LHb), lateral hypothal-

amus (LH), pedunculo-pontine tegmental nucleus (PPTg), laterodorsal tegmental nucleus (LDTg) and rostromedial tegmental area (RMTg), also depicted as the VTA's tail. The previously listed afferents all affect the VTA DA neuronal activity either directly by synapsing on dopaminergic neurons or indirectly by synapsing on GABAergic interneurons. The complexity of the interplay between the different subpopulation of dopamine neurons and their efferents will be discussed in 2.4.

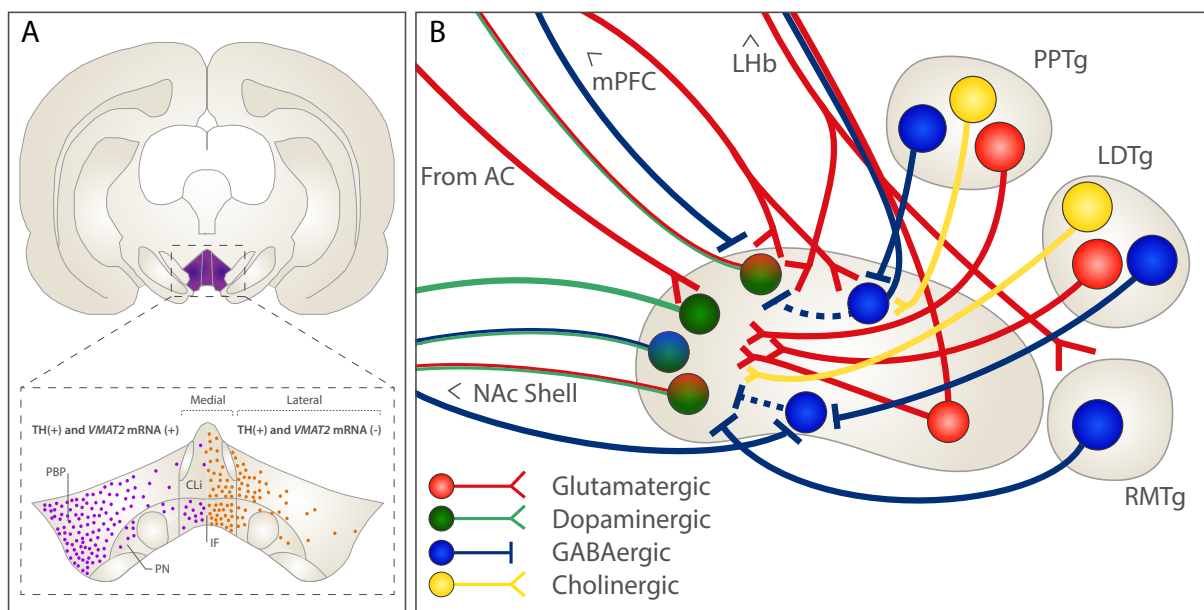


Figure 2.2 – **A. A simplified representation of the VTA topology and DA neuron identification based on TH and VMAT2 mRNA markers in a mouse brain.** This image is adapted from Morales & Margolis, 2017 [65]. TH: Tyrosine Hydroxylase; VMAT2: Vesicular monoamine transporter 2; PBP: parabrachial pigmented nucleus; PN: paranigral nucleus; Cli: caudal linear nucleus; IF: interfascicular nucleus and rostral linear nucleus of the raphe (RLi; not shown). **B. A simplified representation of the synaptic connections in the VTA in a mouse brain.** This image is adapted from Pistillo et al., 2015 and Morales & Margolis, 2017 [63, 65]. AC: Anterior cingulate; mPFC: medial Prefrontal Cortex; NAc: Nucleus Accumbens; LHb: Lateral habenula; LH: Lateral hypothalamus; PPTg: pedunculo-pontine nucleus; LDTg: Laterodorsal tegmental nucleus & RMTg: rostromedial tegmentum

## 2.2 Dopamine synthesis and receptors

Dopamine neurons synthesise and release dopamine and are classically defined as neurons that express tyrosine hydroxylase (TH). However, recent findings have shown that TH mRNA is not a guarantee for DA synthesis, nor DA transmission, since vesicular monoamine transporter 2 (VMAT2), was not detected in all TH-expressing VTA neurons in rodents (*see Figure 2.2a* [65]). Dopamine is considered to be a neuromodulator and belongs to the family of catecholamines. It is primarily synthesised by the enzyme aromatic L-amino acid decarboxylase from its precursor L-DOPA (L-3,4-dihydroxyphenylalanine), which is (in)directly synthesised from respectively the amino-acids L-phenylalanine and L-tyrosine. Dopamine is in turn converted into noradrenaline and adrenaline. Dopamine binds onto dopamine receptors.

### 2.2.1 Dopamine receptors

Dopamine receptors are G-protein-coupled receptors (GPCRs), with 7 trans-membrane regions. GPCRs are metabotropic receptors that functionally modulate other receptor systems and/or ion channels [67,68]. Activation of these receptors modulates, via intracellular cascades, a set of biophysical properties of the synapses and the postsynaptic neurons that changes their input-output relationship. There are at least five DA receptor subtypes in the CNS that are grouped into two major classes: The Gs-coupled D1 family (D1, D5 receptors) and the Gi/o D2 family (D2, D3, D4 receptors). D1-class and D2-class receptors have opposing effects on adenylyl cyclase activity and cAMP concentration, which results in respectively activation or inhibition of the neuron [69–71]. D1- and D2-like receptors are found on pyramidal and non-pyramidal neurons in the prefrontal cortex, with an up to 10-fold greater amount of D1-like receptors [72]. In the nucleus accumbens, D1- and D2-like receptors are located on medial spiny neurons (MSNs) and tonically active neurons (TANs). In the dorsal striatum the MSNs are clearly divided in two subpopulations and their pathways. The direct, striato-nigral, pathway originates from the D1R cells and sends excitatory output to the SN pars reticula (SNr) and the internal portion of the globus pallidus (GPi). The cells populated with D2Rs send inhibitory output to the external portion of the GP (GPe) and is called the indirect pathway. The output organisation of the ventral striatum (NAc) is less ordered [73]. The NAc core has similar connection

to the SNr through the dorsolateral portion of the ventral pallidum and the subthalamic nucleus. The projection to the VTA is part of an intra-basal ganglia projection loop as DA neurons contacted by shell MSNs project back to the core MSNs [74].

## 2.3 VTA dopamine function

DA function can be estimated from extracellular recordings of DA neurons' firing activity and from the extracellular concentration of DA by microdialysis, voltammetry, and positron emission tomography (PET). Each of these methods operates on different time scales: milliseconds for extracellular recording, seconds for voltammetry (with the exception of fast-scan voltammetry) and minutes for microdialysis and PET. Thus they do not necessarily estimate the same aspect of the DA function. As briefly mentioned before, it has been proposed that DA operates in different modalities depending upon the time-scale of its action [33, 46, 75, 76]. I will first discuss the characteristics of VTA DA firing and subsequently the consequences for DA release and behaviour.

### 2.3.1 Electrophysiological features of VTA DA neurons

Extracellular recordings in the VTA report two types of electrophysiologically distinct neuronal groups: broad-waveform slow-firing neurons ( $< 10$  Hz) and narrow-waveform fast-spiking neurons. VT DA cells belong to the first category of neurons. Overall, despite remaining difficulties [77, 78], different key electrophysiological features can identify DA cells. First of all, dopamine neurons have a slow, 2-10 Hz, firing rate, a tonic firing, which is driven by an intrinsic pacemaker potential and can be mixed with bursts. Burst firing is identified as series of 2-10 spikes. The first interspike interval (ISI) is less than 80 ms and the last larger than 160ms (*see Figure. 2.3*, [79]). Burst firing is mostly driven by afferent inputs [80]. Secondly, depending on the filtering, the action potential can have a biphasic (unfiltered and filtered) or a triphasic (high-pass filter,  $> 50$ Hz) action potential with a duration from 1.1 ms and higher, measured from the spike initiation to the maximal negative phase of the action potential [81]. Furthermore, mesolimbic (neurons projecting to the NAc) will inhibit by systemic administration of low doses of dopamine or dopamine receptor antagonists, but this seems not applicable for mesocortical (projecting to the PFC) neurons [82]. Interestingly enough, some non-dopamine neurons in the VTA

exhibit D2R-mediated inhibition [83]. In slice electrophysiology, dopamine neurons also display a hyperpolarisation-activated inward current ( $I_h$ ), although this has not been shown in mesocortical dopamine neurons [84]. Overall, despite recurrent debates, VTA DA cells match with the electrophysiological parameters [85,86]. The question remains if the selection parameters are not too restrictive and exclude VTA DA cell subpopulations due to atypical properties.

### A Regular Firing

### B Burst firing

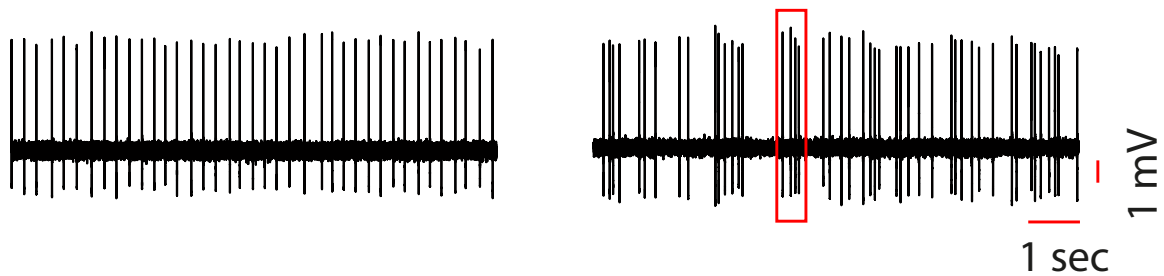


Figure 2.3 – **VTA DA cell firing** **A: Tonic firing.** Trace of a tonic firing neuron with a frequency of 4Hz. **B. Trace of burst firing.** A typical burst is marked with red in which the first ISI is less than 80ms and the final ISI more than 160 ms. Modified from Faure et al., 2014 [87]

### 2.3.2 Spontaneous activity of VTA DA neurons

Around 50% of the dopamine neurons are not spontaneously active [88,89]. GABA-mediated inhibitory postsynaptic potentials (IPSPs) originating from the ventral pallidum are thought to silence this population. Indeed, when the ventral pallidum was lesioned these neurons became active [90]. Active VTA DA cells exhibit firing patterns that range from regular spiking to bursting activity. At the level of the population, the basal activity of dopamine neurons can be affected on the long term and these changes can have extensive consequences on behaviour considering the modulatory role of dopamine [87,91,92] For example, stress [93,94] but also addictive drugs [95] can increase DA cell activity. Modifications of the basal activity can thus be used as an indicator of the effect of such external events. It can also be correlated with particular responses of an individual. The level of activity of DA neurons is a predisposing factor that could favour addiction [96]. Correspondingly, animals with an elevated reactivity to a novel environment display a

high-frequency and bursting DA activity compared to animals with less reactivity to the same environment [96]. Understanding and quantifying the long-term modulations of DA cell dynamics by external events is thus critical: it may drastically change the normal response of DA cells (e.g. to reward, aversion) and possibly lead to DA-associated pathological behaviours [87].

### 2.3.3 Origins of burst firing

Burst firing (synchronized transient increases in firing rate) appears to be regulated by several components that control firing rate at different time courses. Stimulation studies have shown that activation of the DA neuron axons in patterns resembling burst discharge will release two to three times more DA than is released by an equivalent number of evenly spaced stimuli [97, 98]. A set of evidences suggests that cholinergic afferents may play a key role in such synchronisation, which can be confirmed with the observation of nicotine-induced DA cell bursting and specific nicotinic knock out mice (*see Chapter 3*, [99]). The mesopontine tegmentum/nuclei, including the LDTg and the PPTg, sends cholinergic and glutamatergic projections to the VTA and have been shown to modulate burst firing in vivo [90, 100–102]. Burst firing is generally evoked by glutamatergic input [79, 90, 103] acting via N-methyl-D-aspartate receptors (NMDAR; [104, 105] and/or on GABAergic disinhibition [106–108]. However, the application of glutamate or NMDA agonists does not elicit bursting ex vivo, neither do these neurons exhibit bursting activity spontaneously, which implies the importance of other afferents [80]. Application of small conductance calcium-activated potassium (SK) channel blockers however does result in bursting [109, 110]. The glutamatergic and cholinergic activation of the LDTg [111] and the cholinergic innervation of the PPTg (since only the posterior part where cholinergic cells are concentrated projects to the VTA, [112, 113] both a pivotal role in this regulation. LDTg inactivation prevents burst firing of DA neurons [114], while PPTg activation directly induces bursting activity of the VTA [90, 115]. All these data suggest that the mesopontine tegmentum acts as a form of gain control [80] that allows DA neurons to burst in response to excitatory glutamatergic inputs from different parts of the brain.

## 2.4 Heterogeneity of dopaminergic neurons in the VTA

In the early 1960s, the presence of three major groups (A8, A9 and A10 neurons) of midbrain dopamine-releasing neurons was revealed. The A10 group, the VTA is distributed within several subregions, including the parabrachial pigmented nucleus (PBP), paranigral nucleus (PN), caudal linear nucleus (CLi), interfascicular nucleus (IF, *see Figure 2.2a*) and rostral linear nucleus of the raphe (RLi). Although there is general agreement that the PBP and the PN are subdivisions of the VTA, there has been a lack of consensus on whether the RLi, CLi and IF should also be included as part of the VTA [60, 65, 116]. Anatomical, biochemical and electrophysiological studies show heterogeneity among dopamine neurons within each of these subregions [117]. Properties are organized on different axes: medial-lateral, anterior-posterior or ventral-dorsal. Functional heterogeneity is frequently linked to axonal projections and input (*see Figure 2.2b*, [65, 118, 119]). The most apparent differentiation between VTA dopamine neurons was mentioned before and is related to the presence of  $I_h$ .

The DA neurons in the medial posterior VTA (PN and medial PBP) that selectively project to the NAc medial shell and core and basolateral amygdala (BLA), do not display an  $I_h$  current and neither respond to D2R-antagonists. DA neurons that project onto the NAc lateral shell are predominantly located in the lateral posterior and anterior VTA (lateral PBP) and do display these characteristics [78, 120]. VTA DA neurons projecting onto the mPFC that receive glutamatergic input from the LHb are merely involved in aversion [34, 118, 121, 122], as well as negative feedback processing through activation of the GABAergic interneurons. Furthermore, GABAergic input from the LHb neurons synapsing on dopamine neurons projecting to mPFC, have an inhibiting role on the mesocortical network [123, 124]. Indeed, electrical stimulation of the LHb as well as the RMTg resulted in complete inhibition of VTA DA neuron activity [125] and LHb cells fire phasically after reward omission or receipt of smaller rewards, whereas unexpected larger reward induce brief suppression in activity [33, 126, 127]. Glutamatergic input from the LDTg, VTA interneurons and anterior cingulate projecting onto the "mesolimbic" dopaminergic neurons all play a role in reward processing [65, 118, 122, 128–130]. Likewise, administration of cocaine selectively modified excitatory synapses on DA cells projecting to nucleus accumbens medial shell [122]. On the other hand non-rewarding, aversive stimuli like a tail pinch or a foot shock were shown to inhibit dorsal VTA DA neurons, while neurons

in the ventral VTA were phasically excited by these noxious stimuli [86, 131].

Identification of the specific cholinergic inputs, their synaptic VTA targets and projection sites indirectly reveals a lot of information about the potential effect nicotine treatment can have on these networks. Firstly, different subpopulations have been identified in the VTA based on excitatory and inhibitory responses to nicotine [86]. The inhibition in response to nicotine requires D2R activation of autoreceptors [132] or NAc/mPFC receptors of DA neurons that are located in the medial part of the VTA, while activated neurons are located in lateral parts [86]. This segregation can have a protective function in the regulation of drugs [133]. Nicotine self-administration, as well as ICSS were diminished after cholinergic (ChAT) neuron selective chemical lesions of the PPTg [134] confirming the balanced role of the mesopontine tegmentum in phasic activity and reward processing.

## 2.5 Optogenetic manipulation of VTA neurons

In the last decade, optogenetic tools have provided the possibility to control specific neurons with temporal precision in freely behaving mammals. Phasic stimulation of VTA DA neurons enhances dopamine transient magnitudes, compared to longer and lower frequency spiking [98, 135–137]. Phasic stimulation was also sufficient to drive behavioral conditioning in a conditioned place preference task<sup>1</sup> and [98], elicit vigorous self-photostimulation in rats [138], facilitate the learning of food-reinforced operant behavior in mice [135] and increase reward value to a stimulus in rhesus macaques [139]. However, when VTA DA neuron activity was attenuated through selective photostimulation of the inhibitory VTA GABA neurons mice developed an aversion in a conditioned place experiment [140]. In support, VTA GABA neuron activation also disrupted reward consummatory behavior, but not approach behavior. This behavioral effect was not observed when GABA projections to the NAc were activated [141], indicating a role for VTA interneurons in counterbalancing reward processing.

As mentioned before, aversive and rewarding properties were also linked to different subpopulations of dopamine neurons in the VTA. Indeed, when the "inhibitory" population was stimulated through activation of the lateral habenula input neurons place aversion was observed, while place preference was demonstrated when the "excitatory"

---

1. Conditioned place preference task: Pavlovian conditioning is used to measure the associated preference for objects, stimuli or drugs. This task can also be used to study aversion.



population was activated through stimulation of laterodorsal tegmentum neurons [124].

The photomanipulation of the VTA DA neurons has not only supported the role of VTA DA in reward processing and been a proof of concept for optogenetic tools, it can also assist in temporarily mimicking a phenotypic state by hijacking the baseline firing of VTA DA neurons. For instance, optogenetic induction of VTA DA phasic firing, but not tonic firing, promotes susceptibility to social defeat stress and increases social avoidance [142]. This effect is modulated by the mesolimbic [94, 143] and not the mesocortical pathway [144]. In contrast, temporary inhibition of VTA dopamine neurons induced depressive-like behaviors (sucrose-preference and tail-suspension task), which was reversed with temporary stimulation of the same population [136]. These results seems contradictory, however avoidance behaviour has been linked to an increase in VTA DA neuron activity [145]. Furthermore, it has been demonstrated that induced tonic patterns of dopamine transmission reduce ethanol [146] and sucrose self-administration [147], which can probably be explained through attenuation of the motivational drive [46]. Moreover, brief pauses in the firing of VTA DA neurons at the time of an expected reward was sufficient to mimic the effect of endogenous negative prediction errors and restore extinction learning [148]. Mimicking a phenotypic state allows studying different disorders, as well as the long term effects of nicotine addiction on the VTA DA cells. In combination, with time-controlled experiments like Chang and colleagues (2016) aberrant decision-making could be tested and manipulated under different baseline conditions.

## 3 | The nicotinic cholinergic system

The cholinergic complex exists out of multiple pathways and is also essential for attention, memory and motivation. The basal forebrain contains two groups of cholinergic neurons: (1) the medial septal group that project cholinergic axons to the hippocampus and parahippocampal gyrus and (2) the nucleus basalis group that project cholinergic axons to the neocortex, parts of limbic cortex and to the amygdala. The cholinergic mesopontine tegmentum neurons (LDTg and PPTg) project onto the VTA, hindbrain, thalamus, hypothalamus and basal forebrain. Cholinergic neurons secrete the neurotransmitter acetylcholine that binds onto presynaptic and postsynaptic muscarinic receptors (mAChRs; subtypes M1-M5) and onto nicotinic cholinergic receptors (nAChRs).

The nAChRs are named after their affinity for nicotine. Nicotine mimics the agonistic action of acetylcholine after binding onto nAChRs. nAChRs are present in various organs and tissues. They are also located throughout the brain, with the highest density seen in the thalamus, followed by the basal ganglia, frontal, cingulate, occipital, and insular cortices [?, 149].

### 3.1 Nicotinic acetylcholine receptors

nAChRs are prototypical members of the ligand-gated ion channel superfamily of neurotransmitter receptors (*see Figure 3.1*). They act as a neuromodulator and influence neuronal excitability and activity-dependent plasticity in most mammalian brain structures. They are expressed at presynaptic or -terminal sites, postsynaptically and on extra-synaptic locations [150]. nAChRs activation by Achetylcholine (ACh) or nicotine consists of the fast opening (microsecond to millisecond range) of a cationic channel. When a cholinergic agonist binds to the channel, the channel opens, allowing the entry of cations, including sodium and calcium. These cations further activate voltage-dependent calcium channels, promoting further calcium entry. nAChRs currents have relatively fast kinetics compared to the muscarinic acetylcholine receptor, the second acetylcholine receptor [151].

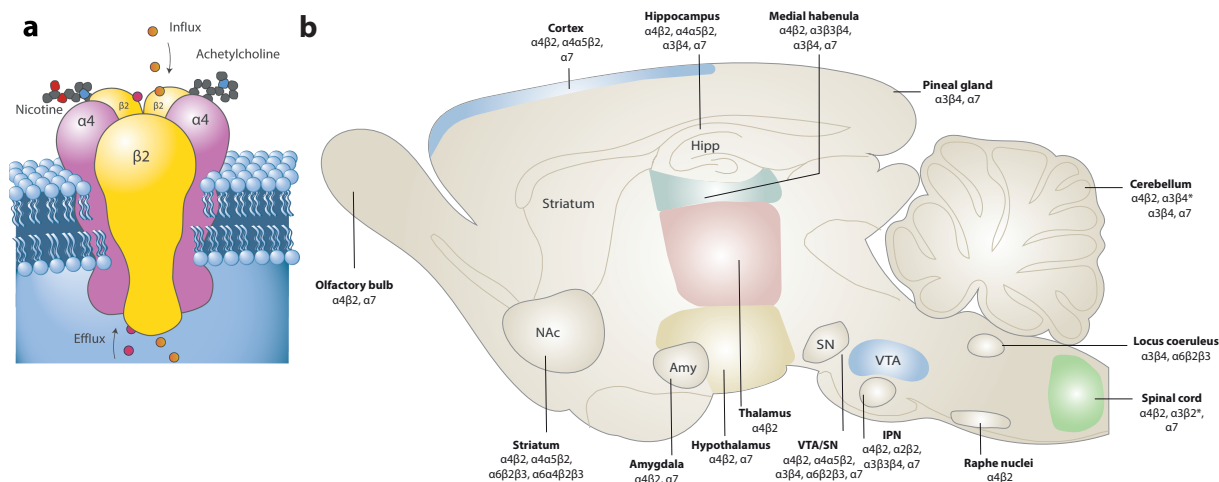


Figure 3.1 – **Nicotinic receptor.** **A** Schematic representation of a nicotinic receptor. **B:** **Representation of the nAChR distribution** The nicotinic receptors are widely distributed through the brain, and can therefore act through multiple mechanisms. Adapted from Taly et al, 2009 [152]

The nAChRs are transmembrane allosteric oligomers composed of five identical (homopentamer) or different (heteropentamer) subunits. There are as many as nine  $\alpha$  subunits ( $\alpha 2$  to  $\alpha 10$ ) and three  $\beta$  subunits ( $\beta 2$  to  $\beta 4$ ). These subunits can assemble in different combinations to construct receptors with distinct functional and pharmacological properties. Depending on the subunit composition, the receptors show very different affinity for ACh or nicotine. The most abundant receptor subtypes in the brains of mammals are  $\alpha 4\beta 2$ ,  $\alpha 3\beta 4$ , and  $\alpha 7$  (homomeric). The  $\alpha 7$  homopentamer has a low affinity for ACh and nicotine, a high calcium permeability and rapid activation and desensitization kinetics, while the  $\alpha 4\beta 2^*$  containing heteropentamers (\* indicates the possible presence of other subunits) are highly sensitive to agonists and strongly up-regulate in response to chronic nicotine exposure. This receptor is therefore believed to be the main receptor mediating nicotine dependence (see below, [153, 154]). The lack of specific agonists or antagonists capable to functionally isolate nAChR subtypes in vivo, initiated a fast developing field of subunit specific knock out (KO) mice [23, 155–158]. More recently, researchers have managed to change the sensitivity of specific subunits, but also to develop knock-in and transgenic (single point mutation or single nucleid polymorphism (SNP)) mice [159–161]. These mice models have elucidated a lot of knowledge on the reinforcing effects of nicotine through the dopaminergic system [158, 162, 163].

## 3.2 nAChRs in the dopaminergic system

As discussed in the previous chapter, the dopaminergic system receives cholinergic innervation from the PPTg and the LDTg [100,164]. In the midbrain, the nAChRs modulate DA cell activity and DA release according to their location on DA cell bodies, dendrites and terminals in the striatum and cortices, but also indirectly through glutamatergic, cholinergic and GABAergic afferents of the VTA DA neurons. As a consequence, the activation of nAChRs can have opposite modulatory effects on the same circuit depending on where they are expressed on excitatory or inhibitory neurons [27,63]. Almost all nAChRs subunits are expressed in midbrain DA neurons. However, two main populations, the  $\alpha 4\beta 2^*$  and  $\alpha 4\alpha 6\beta 2^*$  nAChRs are highly represented in the DA cell body/dendrite compartment of the VTA [160,165–167]. DA cell bodies/dendrites also express a high level of  $\alpha 5$  subunits that account for about half of the  $\alpha 4\beta 2^*$  nAChRs, and the lack of this subunit in  $\alpha 5$  KO mice halves the expression of  $\alpha 4^*$  nAChRs (*see Figure 3.1*, [168]).

It has been found that DAergic neuron firing patterns are hierarchically controlled by nAChRs: the activation of somatodendritic  $\beta 2^*$  nAChRs switches cells from a resting to an excited state, whereas the activation of  $\alpha 7^*$  nAChRs (those present in the cell body and/or those present in glutamatergic afferents) only fine tunes the latter state once  $\beta 2^*$  nAChRs have been activated [169]. GABAergic interneurons and terminals in the VTA express the  $\alpha 4\beta 2$  subtypes, whose stimulation promotes the release of GABA from synaptic terminals [158,165]. Furthermore, the release of GABA on DA neurons is blocked by the  $\alpha 6\beta 2$ -selective antagonist  $\alpha$ -conotoxin MII ( $\alpha$ Cntx-MII), thus suggesting that  $\alpha 6$  receptors may be present on GABAergic interneurons (*see Figure 3.1*).

In summary, the  $\alpha 4\beta 2$  subtype (with or without the  $\alpha 5$  subunit) is expressed on both DA and GABA neurons and terminals, whereas the  $\alpha 6\beta 2^*$  subtype is mainly expressed on DA neurons. Homomeric  $\alpha 7$  nAChRs are expressed by about half of DA neurons [165] and the glutamatergic afferents in the VTA [170]. ACh potentiates glutamatergic or GABAergic inputs to DA neurons respectively by acting on the presynaptic receptors expressing the  $\alpha 7$  or  $\alpha 4\beta 2$  on nerve terminals [171].

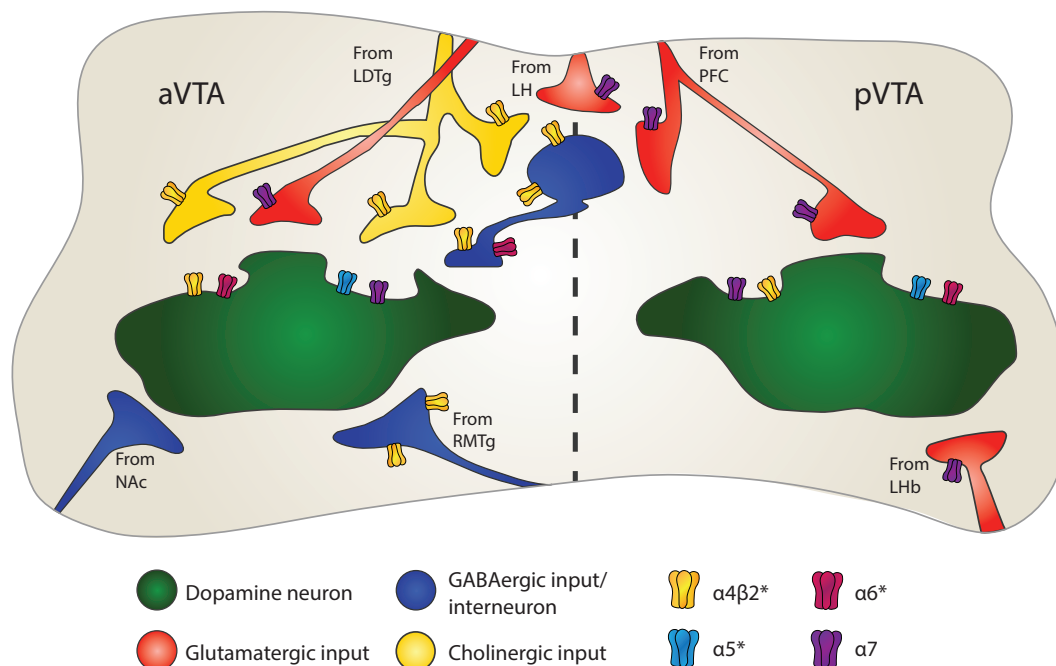


Figure 3.2 – **Location of the nAChRs within the ventral tegmental area.** Adapted from Pistillo et al., 2015 [63]

### 3.3 Nicotinic control of the dopaminergic system

#### 3.3.1 Acute effects of nicotine

Acute in vivo nicotine administration enhances synchrony among dopamine neurons and increases the general firing and bursting rates of DA neurons in the VTA [87, 172–174] with DA transients in the NAc, underlying the reinforcing properties of nicotine [30, 175]. However, nicotine increases the firing rates of DA neurons in the posterior VTA more than in the anterior VTA (*Figure 3.1*, [174], which can probably be explained by the reduced GABAergic input in the posterior VTA. Finally, as stated above, the laboratory has identified, in addition to the classically-described excited population, a VTA DA subpopulation located preferentially in the medial part of the VTA that is inhibited by an acute injection of nicotine [86]. These findings suggest unexplored roles for DA release in addiction and contrast with the classical views of reinforcement and motivation.

Systemic or intra VTA injection of nicotine elicits self-administration and conditioned place preference (CPP). The abundant  $\beta 2$  and  $\alpha 4$  containing nAChR subunits are essential in nicotine reinforcement, but play distinguished roles. Nicotine-induced bursting is completely absent in  $\beta 2$  KO mice. These mice also do not exhibit nicotine self-administration,

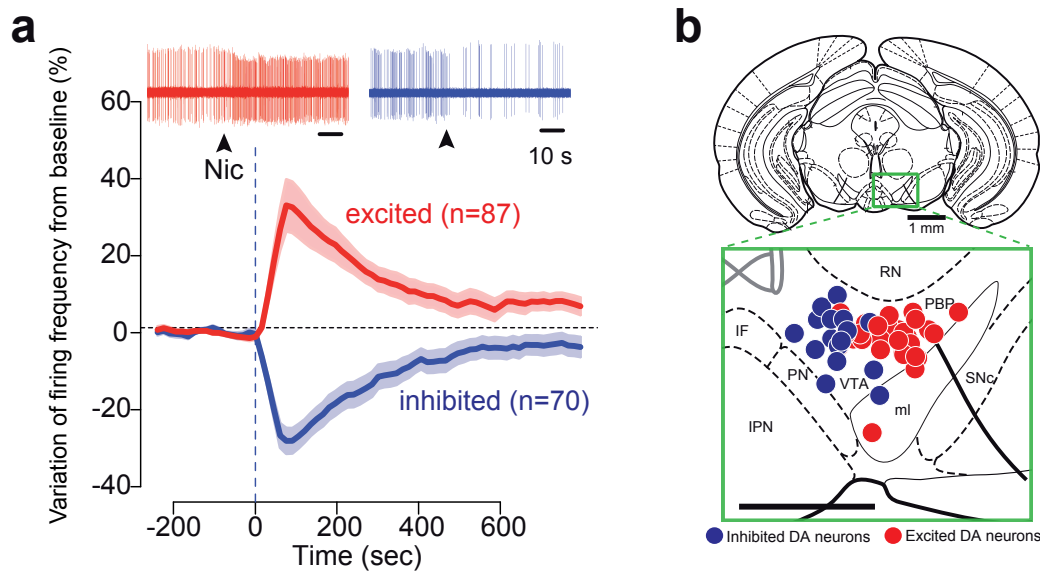


Figure 3.3 – Nicotine evoked response in VTA DA cells . **A** Mean variation of firing frequency of excited (red,  $n = 53$ ) or inhibited (blue,  $n = 23$ ) VTA DA neurons. Top : Raw single-cell extracellular recordings of an excited (red) and an inhibited VTA DA neuron (blue) at the time of nicotine ( $30\mu\text{g}/\text{kg}$ ) IV injection (indicated by arrowheads). **B**) Coronal diagram at  $-3.52$  mm from Bregma onto which are positioned the labeled cells (DA cells are labeled with TH immunolabeling). Modified from Eddine et al., 2015 [86].

neither nicotine-induced locomotor activity, nor a preference in the CCP. Both the behaviour and bursting can be restored by the targeted re-expression of the  $\beta 2$  subunit in the VTA [23, 25, 27]. However, re-expression of both the  $\beta 2$  subunit on the GABA, as well as DA neurons in the VTA was necessary to restore the reinforcing actions of nicotine through bursting of DA neurons [27].

The  $\alpha 4$  subunit is reported to be more involved in the mediation of nicotine-induced reward, tolerance, and sensitization [176]. The  $\alpha 4$  subunit appears to be an important determinant of sensitivity to nicotine. In mice, a single nucleotide point mutation in the pore-forming region results in a receptor that is hypersensitive to the effects of nicotine. This mutation makes mice much more sensitive to nicotine-induced reward behaviors, as well as to effects on tolerance and sensitization. On the other hand, since most  $\alpha 4$  nAChRs also contain  $\beta 2$  subunits, the  $\alpha 4$  KO mice recapitulate many phenotypes of  $\beta 2$  KO mice.  $\alpha 4$  KO mice also have blunted and delayed nicotine-elicited increases in firing rate, bursting and DA release, altered nicotine-stimulated locomotion, and abolished

nicotine self-administration [85, 177, 178].

Fast-scan cyclic voltammetry recordings have revealed a dual action of nicotine at the DA axon terminals in the NAc: nicotine enhances phasic release of dopamine, while it suppresses tonic release. This increase in the contrast is induced by a transition from tonic frequencies to high frequencies and enhances the sensitivity of DA release to presynaptic depolarisation [179]. Since DA release was neither further increased by the administration of nicotine to  $\alpha 6$ ,  $\alpha 4$ , or  $\beta 3$  subunit KO mice [180], the effect of nicotine is most probably accounted for by the  $\alpha 4\alpha 6\beta 2\beta 3^*$  nAChRs. Indeed, rat studies have also shown that the injection of  $\alpha$ Cntx-MII in the NAc blocks the DA release elicited by local nicotine perfusion [167] and that injections of  $\alpha$ Cntx-MII in the NAc shell dose-dependently decreases the number of break points and nicotine infusions in intravenous SA procedure, thus indicating that  $\alpha 6$  plays a role in regulating the motivational aspects of nicotine reinforcement [181].  $\alpha 6$  and  $\beta 3$  KO mice also have reduced nicotine-stimulated DA release in striatum [166, 182].  $\beta 3$  nAChR subunits are important for nAChRs  $\alpha 6^*$  biogenesis in the DA system, because  $\alpha 6^*$  nAChR binding and functional expression is dramatically reduced in  $\beta 3$  KO mice [182]. This amplification might provide a mechanism for nicotine facilitation of reward-related dopamine signals, including responses to other primary reinforcers, like discussed in *Chapter 1*, that governs nicotine dependence in smokers (*see next paragraph*, [26, 85, 183–185]. Together with the "gating" mechanism through activation of the somatodendritic nAChRs on midbrain DA neurons by mesopontinergic input [169, 186, 187] the contrast between tonic and phasic activity is modulated through nAChRs at different stages in the mesolimbic pathway and this makes nicotine act as a powerful artificial reward [187].

On the other hand,  $\alpha 6$  and  $\alpha 5$  subunits seem to specifically play a role in DA transmission in the VTA. Acute intravenous nicotine self-administration is eliminated in  $\alpha 6$  KO mice and is restored when  $\alpha 6$  subunits are selectively re-expressed in the VTA [178]. On the contrary,  $\alpha 5$  KO mice remain sensitive to nicotine [85, 188]. However, higher doses of nicotine are required to evoke similar DA increases [19, 188].  $\alpha 5$  KOs, as well as specific  $\alpha 5$  human polymorphisms, SNPs, have been associated with both schizophrenic phenotypes [189], (*see next paragraph*) and increased nicotine tolerance, i.e. increased vulnerability to become a smoker and to smoke significantly larger amounts, as derived from decreased DA neuron responsiveness to lower concentrations of nicotine and SA data is

mice [188], *see Chapter 7*).

### 3.3.2 Chronic exposure to nicotine: a smoker's brain

Administration of nicotine at the concentrations present in smoker's brains (300-500 nM; [190]) has shown to desensitize the nAChRs expressed by DA terminals. Long term smoking leads to up-regulation of nAChRs in the human brain, including the common  $\alpha 4\beta 2^*$  nAChR [191], which is confirmed with animal studies [192–196]. Higher concentration of nicotine can induce up-regulation of other subtypes, such as  $\alpha 7$  [197]. Since, the number of  $\alpha 7$  receptors was increased by self-administration of nicotine [198], while a decrease was found after exposure to nicotine by an osmotic minipump or drinking water [199, 200]. Even human postmortem tissue studies show that chronic smokers have increased numbers of  $\alpha 4\beta 2^*$  nAChRs compared to non-smokers [201, 202]. In contrast, former smokers ( $> 1$  year abstinent) have nAChR densities similar to non-smokers [202]. Thus, this upregulation is reversible after an extended period of abstinence [203].

The up-regulation might be a consequence of the rapid desensitization of nAChRs [204]. The loss of receptor function would promote up-regulation in order to compensate for the reduced signaling [205]. Most likely, several mechanisms for nicotine-induced up-regulation of nAChRs are involved, including alteration in receptor assembly [206], trafficking [207], decreased turnover [208] and maturation [209]. Finally, chronic exposure to nicotine can also change the stoichiometry of the receptors (i.e. the relative proportions of subunits in heteromeric receptors) and thus impact the cholinergic activity. For example, although the  $\alpha 5$  subunit does not contribute to the nicotine-binding site [210], its deletion leads to a dramatic shift in several nicotine-elicited alterations. In particular, it has been suggested that in the VTA, surprisingly,  $\alpha 4\beta 2^*$  nAChRs that do not comprise  $\alpha 5$  have a minor role in the nicotine-evoked response of DA cells [188]. Apart from these effects, chronic nicotine exposure may also modify circuits through synaptic and intrinsic plasticity, through signal transduction mechanisms arising from  $\text{Ca}^{2+}$  influx through nAChRs. These adaptations can arise downstream from the receptors. It has been shown that nicotine induces long term potentiation (LTP) at VTA DA excitatory synapses, 24 hours after *in vivo* administration of the drug [211], an effect that required activation of pre-synaptic  $\alpha 7^*$ nAChRs. Modification of the AMPA/NMDA ratio has been demonstrated in different parts of the brain in response to nicotine administration [212, 213].



After a night of complete abstinence, nicotine concentrations in the brain are at their lowest level and thus nAChRs have recovered from desensitization resulting into an excess (due to the up-regulation) of excitability of the systems under the influence of the endogenous nicotinic neuromodulation (*see Figure.1.1*) In contrast, during the day and after a few cigarettes some subpopulations of nicotinic receptors may have entered the desensitized state, depriving the neurons or synapses from its normal endogenous modulation by ACh. Positron emission tomography (PET) and Single-photon emission computed tomography (SPECT) scans, using radiotracers with a high affinity and relative specificity to  $\alpha 4\beta 2^*$  nAChRs, have demonstrated that daily smokers smoke in amounts that lead to nearly complete occupancy (95%) of  $\alpha 4\beta 2^*$  nAChRs, such that tobacco-dependent smokers maintain  $\alpha 4\beta 2^*$  nAChR saturation throughout the day. The required number of cigarettes differ per smoker due to significant inter-individual differences in satiety and periods of abstinence [203].

### 3.4 Behavioural implications of nicotine-use

Since, the nAChRs are present everywhere in the brain, the effects of nicotine can not only be limited to the mesolimbic system (*see Figure 3.1*). The plasticity due to nicotine has two main effects: it promotes addiction [214] and it modulates cognition, including attention and memory [215], as the endogenous cholinergic system modulates both local and global aspects of information processing [216]. Mostly, specific SNPs or the inactivation of  $\alpha 5$  and  $\alpha 7$  subunits have been associated with cognitive difficulties [217–219]. Several neuroimaging studies have examined the impact of nicotine and tobacco on large-scale brain networks. These intrinsic networks can be detected at rest by assessing correlations in temporal fluctuations in the blood-oxygen-level dependent (BOLD) signal across brain regions [215, 220, 221]. However, nicotine’s acute reinforcing and cognition-enhancing effects both appear to be linked to nicotine-induced enhancement of DA signaling since nicotine and DA agonists present similar inverted-U dose-effect curves with respect to both intensity of drug self-administration and cognitive task performance [222]. In studies of nicotine’s effects on cognition, absence of nicotine in heavy smokers produces cognitive deficits, similar deficits are seen with nAChR antagonists in non-smokers. Moderate doses of nicotine typically produce cognitive enhancement in non-smokers, nicotine-

deprived smokers and non-deprived smokers [223], whereas very high doses of nicotine typically impair cognitive performance. Similarly, DA depletion impairs cognitive processing, moderate levels of DA produce optimal cognitive performance, whereas very high levels of dopamine impair cognition.

Acute nicotine-induced enhancement of cognitive task performance is associated with increases in activity in subcortical regions, such as basal ganglia and thalamus, as well as with both increases and decreases in activity in frontal and parietal cortices. In addition, the impact of acute nicotine on task-related brain responses may be further modulated by the dose and route of administration, length of abstinence in smokers, and pre-existing cognitive deficits such as those associated with schizophrenia or attention-deficit/hyperactivity disorder (ADHD) diagnosis, as well as by the type and difficulty of the task itself. Individuals with cognitive disorders such as schizophrenia, ADHD, or Alzheimer's disease (AD) may self-medicate with nicotine to enhance cognitive and attentional processes, which may partially account for the high prevalence of cigarette smoking among these individuals. Overall, nicotine seems to enhance executive function not via a single neuroanatomical site or mechanism, but via the modulation of multiple brain networks [224].

However, long term smoking has detrimental consequences as the smoker's brain undergoes brain plasticity. Firstly, the severity of nicotine dependence was specifically and negatively associated with a dorsal anterior cingulate (dACC)- ventral striatum (VST) circuit. Furthermore, Brody and colleagues (2004) demonstrated smaller gray matter volume and lower gray matter densities in bilateral PFC, along with smaller volume in the left dACC, in smokers compared to matched non-smoking controls. Moreover, the gray matter densities in the PFC negatively correlated with the magnitude of lifetime smoking exposure as indexed by pack-years smoked. Similarly, Gallinat and colleagues (2006) showed that, compared to never-smokers, smokers displayed smaller gray matter volume and lower gray matter density in the frontal lobe (ACC, PFC, and OFC), occipital lobe, and temporal lobe (including parahippocampal gyrus), as well as volume or density deficits in the thalamus and cerebellum. Again, the gray matter volume in frontal, temporal, and cerebellar cortices was inversely correlated with lifetime smoking exposure in smokers [225].

Overall, only the acute effects of nicotine on non-smokers and in specific subgroups are

rather positive. For this reason, nicotine and other nAChR ligands are being investigated as potential therapeutics for the treatment of cognitive deficits in schizophrenia, attention-deficit hyperactivity disorder, and Alzheimer's disease [226, 227]. On the other hand, chronic cigarette smoking modifies the brain to such an extent that it has been associated with decreased cognitive performance in middle age, which has been linked to an increased risk of cognitive decline and dementia later in life [228]. Positive effects in smokers are merely a relieve from withdrawal symptoms.

# 4 | Nicotinic alteration of decision-making: a review

The focus of my research is studying the behavioural consequences of long term passive administered nicotine. Understanding the implications of nicotine-induced brain plasticity will provide insight about the state of an addicted brain, dissected from neurobiological and psychological processes that drive addiction behaviours. This knowledge could strengthen a fundamental basis in addiction research.

A powerful framework to quantify shifts in behaviour induced by acute or chronic administration of pharmacological substances like nicotine is provided by neuroeconomics. The discipline of neuroeconomics investigates the neurobiological and computational basis of value-based decision-making. Value-based decision-making occurs whenever an organism makes a choice from several alternatives. To survive in a dynamic environment, an organism must be able to effectively learn, store, and recall the expected benefits and costs of potential actions. Different normative theories of choice originating from the multidisciplinary field of economy, psychology and neuroscience have outlined the theoretical structure of decision valuations. More recently, neuroscientists have begun to reveal how value is encoded in the activity of neurons and neural circuits, including the cholinergic and dopaminergic systems [179, 229–232] and how values are misrepresented through the actions of drugs [39, 233]. This field is extremely young and many hypotheses are yet to be verified. We have written a review to discuss the developments up to now and to outline the perspectives for the future. I will first briefly introduce the decision-making framework, before presenting the consequences of nicotine and the role of nAChRs in decision-making behaviour in the review.

The process of decision-making involves different steps and brain areas (*see Figure 4.1*). At first, a representation of the necessity and the chance at success of the decision needs to be computed, which depends on different states. For example, when one gets presented with a cookie jar, you will take under consideration your level of hunger at this very moment (internal state), as well as the meals that are yet to come (external state). Secondly, the different actions need to be assigned a value. Multiple valuation systems

have been proposed: a pavlovian, a habitual and a goal-directed system. These systems can subsequently be described and fitted with different computational models [230, 231].

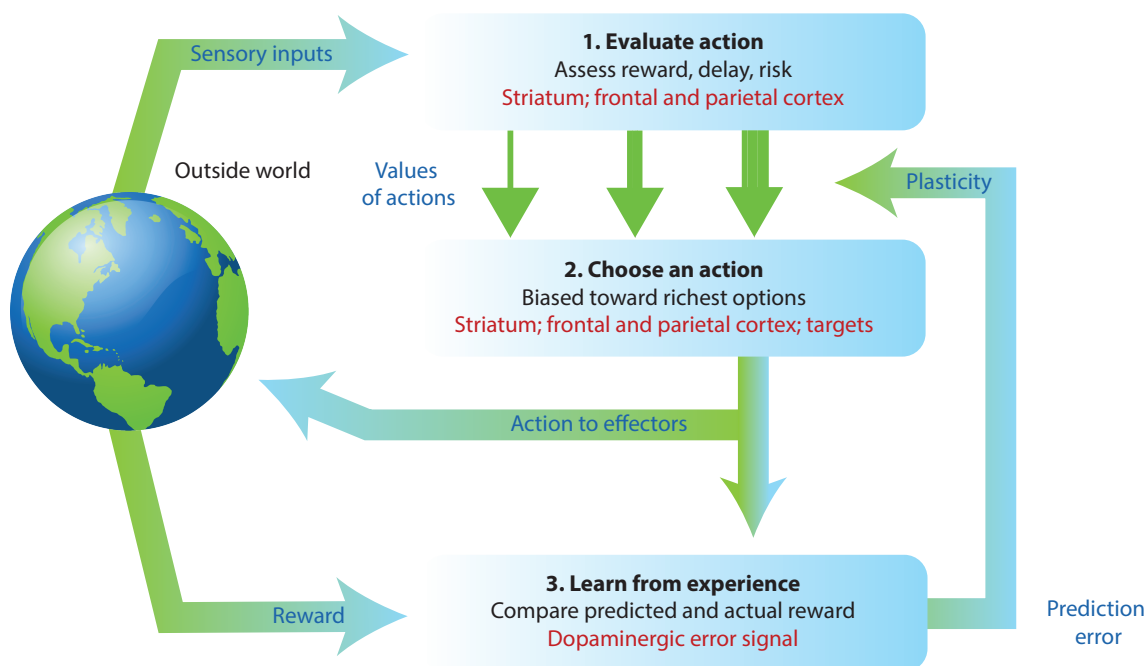


Figure 4.1 – **Principle of reinforcement learning adapted from Daw and Doya, 2006 [230]**. The three basic stages of reinforcement learning in learned decision-making. 1) Predict the rewards expected for candidate actions (the green arrows represent the different values of the available actions) in the current situation. 2) Choose and execute one by comparing the predicted rewards. 3) Finally, learn from the reward prediction error (RPE) to improve future decisions. When the reward is correctly predicted the RPE is low and leads to reinforced behaviour through plasticity.

Pavlovian systems assign values to a small set of behaviours that are innate responses to particular stimuli, like approaching cues that predict the delivery of food, and consummatory responses to a reward, such as the classical example of the salivating dog. Similarly, cues that predict a punishment or the presence of an aversive stimulus can lead to avoidance behaviours. Although pavlovian behaviours are innate, with training organisms can learn them in response to other stimuli. Computations of pavlovian systems rely on different controllers, like outcome-specific responses (eating/drinking), as described by the reward prediction error (*see Chapter 1*) or more general valence-dependent responses (approach motivation). The reward prediction error has been interpreted by different

---

methods or derivatives of these methods. Reinforcement learning (RL) is a computational theory that describes how organisms learn by trial and error to make decisions with the ultimate goal to obtain rewards and avoid punishments. A major learning signal used in RL is called the "temporal-difference error signal". Temporal difference (TD) learning allows to quantify the value or utility (magnitude of organism's reward expectation) assigned to an action [230]. In TD learning one adjusts predictions to match other, more accurate predictions about the future, by optimising the model between temporal measurements taking into account the error.

The habit/reflexive system can learn to assign values to a large number of actions with repeated training. At first, values are assigned to stimulus-response associations (f.e. learning to lever press after cue presentation) through trial-and-error-based processes. Secondly, the system learns to assign a value to actions that correspond with the expected reward that these actions generate in a stable environment. Reinforcement learning models differentiate conceptually between learning the value of each action, and transforming those actions into action selection [234]. Finally, since the trial-and-error-based updates are relatively slow, one has to rely on generalisation to assign action values in novel situations [231, 235].

Goal-directed/reflective systems assign values to actions by computing action-outcome associations (f.e. lever pressing results in food delivery) and then evaluating the rewards associated with the different outcomes. The important difference between habit and goal-directed systems has to do with how they respond to changes in the environment [231].

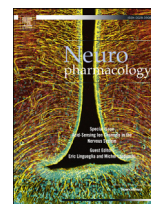
After a choice has been evaluated by the means of a single or multiple systems described above an action needs to be selected. Various modulators can play a role on the final valuation processes of an action, like risk and uncertainty (the likelihood of a specific outcome, *see Review and Chapter 7*) and time. Different delays in reward delivery can add another uncertain or discounting factor. Some models describe that the least uncertain outcomes are being preferred in action selection ("prospect theory", [236]), while others order the different choices on a scale of values (probability\*value) and expect the highest expected value as an outcome ("Expected utility theory", [237]). Finally, to learn to make good decisions the brain needs to compute a feedback signal, like the RPE, that signals the desirability of the outcomes generated by previous decisions and promotes

either exploiting good decisions or exploring new alternatives (*see Chapter 8*, [231]).

We have applied this decision-making framework in the following review to discuss the early findings of the consequences of nicotine and the role of nAChRs in decision-making behaviour.

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

# Neuropharmacology

journal homepage: [www.elsevier.com/locate/neuropharm](http://www.elsevier.com/locate/neuropharm)

Invited review

## Nicotinic alteration of decision-making



Jérémie Naudé, Malou Dongelmans, Philippe Faure\*

Université Pierre et Marie Curie, CNRS UMR 8246, INSERM U 1130, UPMC UM CR18, 75005 Paris, France

### ARTICLE INFO

Article history:  
Available online 9 December 2014

Keywords:  
Nicotine  
Decision-making  
Nicotinic acetylcholine receptors  
Impulsivity  
Risk-taking

### ABSTRACT

Addiction to nicotine is characterized by impulses, urges and lack of self-control towards cigarettes. A key element in the process of addiction is the development of habits oriented towards nicotine consumption that surpass flexible systems as a consequence of a gradual adaptation to chronic drug exposure. However, the long-term effects of nicotine on brain circuits also induce wide changes in decision-making processes, affecting behaviors unrelated to cigarettes. This review aims at providing an update on the implications of nicotine on general decision-making processes, with an emphasis on impulsivity and risk-taking. As impulsivity is a rather ambiguous behavioral trait, we build on economic and normative theories to better characterize these nicotine-induced alterations in decision-making. Nonetheless, experimental data are sparse and often contradictory. We will discuss how the latest findings on the neurobiological basis of choice behavior may help disentangling these issues. We focus on the role of nicotine acetylcholine receptors and their different subunits, and on the spatio-temporal dynamics (i.e. diversity of the neural circuits, short- and long-term effects) of both endogenous acetylcholine and nicotine action. Finally, we try to link these neurobiological results with neuro-computational models of attention, valuation and action, and of the role of acetylcholine in these decision processes.

This article is part of the Special Issue entitled 'The Nicotinic Acetylcholine Receptor: From Molecular Biology to Cognition'.

© 2014 Elsevier Ltd. All rights reserved.

### 1. Introduction

Nicotine contained in tobacco smoke is a widely consumed psychotropic agent, considered as the primary reinforcing component of tobacco, responsible for tobacco addiction (Fowler et al., 2008; Marti et al., 2011; Stolerman and Jarvis, 1995). Addiction to nicotine is characterized by impulses, urges and lack of self-control towards cigarettes (DSM-V, American Psychiatric Association, 2013; West and Brown, 2013). However, nicotine-induced modifications of decision-making processes may also be exhibited towards other rewards, such as food or money. Thus, disturbed decision-making is not only limited to drug-related activities. Instead, acute and chronic exposure to nicotine may change the global brain state of smokers, resulting in modified executive processes. This review aims at providing an update on these consequences of nicotine use with regard to specific decision-making processes.

Both human and rodent studies have demonstrated links between repeated drug use and alterations in brain pathways

involved in behaviors related to decision-making, such as impulsivity (impulsive choice and impulsive action), risk-taking, and cognitive flexibility. Decision-making results from the interaction of distinct executive functions. These processes are based on different computations and arise from various neural dynamics in which acetylcholine plays a role. In this review, we aim to focus on two constructs that involve choice behavior, namely impulsive choice and risk-taking (Evenden, 1999). First, we will briefly present these constructs and relate them to precisely defined economic and computational theories of decision-making. Subsequently, we will review studies of choice behavior that investigated the effect of nicotine on “non-drug” rewards, and the underlying neurobiology, in relation to these theories.

Nicotine acts on a range of nicotinic acetylcholine receptors (nAChRs) that differ in their sensitivity to nicotine, their pattern of expression, and thus in their impact on neural activity. These variations are subsequently linked to diverse complex behavioral effects (Picciotto, 2003). nAChRs are expressed throughout the neural circuitry of decision-making, in particular in the dopaminergic (DAergic) nuclei, as part of the mesocorticolimbic pathways (Dani et al., 2011; Faure et al., 2014; Klink et al., 2001). These structures play a pivotal role in reward seeking, reinforcement learning and decision-making (Bromberg-Martin et al., 2010; Salamone et al.,

\* Corresponding author. Boite 14, Université Pierre et Marie Curie, 9 quai St Bernard, 75005 Paris, France. Tel.: +33 1 44 27 39 40; fax: +33 1 44 27 25 84.  
E-mail address: [philippe.faure@snv.jussieu.fr](mailto:philippe.faure@snv.jussieu.fr) (P. Faure).



2009), but surprisingly, little is known about the contributions of nAChRs to value-based decision-making. Acetylcholine and its agonist nicotine act as neuromodulators that affect the way the brain processes information on different timescales. Smoking can thus have a direct, acute effect on decision-making. However, nicotine also induces desensitization and up-regulation of nAChRs, and modulation of plasticity at non-cholinergic synapses, resulting in the long-term alterations of decision-making processes. In this review, we will furthermore examine the possible role nAChRs have in circuits underlying decision-making processes.

## 2. Theories of decision-making

The process of how an individual makes a decision and selects an action has been investigated through very different paradigms. Neuro-economic and computational approaches have generated normative principles that constitute important steps towards quantitative measures describing cognitive functions related to decision-making. This systematization has been proposed to be useful for the diagnosis and understanding of brain dysfunctions, such as mental disorders (Kishida et al., 2010). In this part of the review, we will provide a brief overview of these principles, before adopting them to characterize the effects of nicotine. Nonetheless, although this framework provides powerful computational principles, their translation in terms of neural substrates is far from being understood (see the last part of this review) and identifying a model or one of its parameter or variable with a given structure may be misleading.

Theories that describe how the brain makes decisions are often based on economic theories of how optimal decisions should be made (Vlaev et al., 2011; Doya, 2008; Padoa-Schioppa, 2011; Kable and Glimcher, 2009). Theories of decision-making generally assume that choices can be compared using a common currency, their “values”, for example a real number, on an internal scale (Vlaev et al., 2011). They hold that the brain computes the value of each available option, represented by the activity of a population of neurons (Padoa-Schioppa, 2011). In optimal value-based decision-making, options with higher values are generally preferred. In the classical version of this theory, ‘the expected utility theory’ (see Schoemaker, 1982), or in modern variants such as “the prospect theory” (Kahneman and Tversky, 1979), each option is associated with a numerical value indicating its ‘utility’. The ‘utility function’ models the size of a reward (i.e. how rewarding is a reward), which may depend on internal and external context. For example, the value of food saturates depending on how much one can eat. The utility function for rewards may be seen as a homeostatic mechanism, where an organism obtains rewards to reduce a drive (eventually, preventively), in order to reach a homeostatic “set point” (Keramati and Gutkin, 2013).

It is widely accepted that in the brain, the value of outcomes at the time of the choices can be determined through at least two different processes: reflective (or goal-directed) control and reflexive (habitual) control (Balleine and Dickinson, 1998; Dolan and Dayan, 2013; Doll et al., 2012). In the habitual/reflexive system, each value of an association between the environment (e.g. a stimulus or a context) and an action is learned from past experiences with good and bad outcomes (Sutton and Barto, 1998). The habitual system is based on trial and error, therefore computationally simple, although inflexible in complex or dynamic environments. In contrast, the goal-directed/reflective system applies a forward-search - through the possible outcomes following an action - to guide action selection (Doll et al., 2012). Therefore, the reflective system supports cognitive flexibility more than habits. However, the exhaustive tree-search in large environments is highly demanding for the working memory and possibly comes

with errors. In general, the habitual system overrides the reflective system after repetition and training. However, one should note that these decision-making systems are not distinct entities (especially at the neural level, see below), but rather two computational frameworks for a range of neural dynamics implicated in value-based decisions (Daw et al., 2005).

The optimal decision maker would choose the option with the maximum utility (expected and/or discounted). However, humans and other animals do not exhibit optimal choice behavior, but rather display a distribution of choices that can be allocated according to, for example, the fraction of rewards expected in each alternative (“matching law”, Herrnstein, 1961). Different hypotheses can be advocated to explain why our choices are not optimal. One of them is that our daily choices require a trade-off between exploiting already known rewards and exploring alternative rewards. Exploratory decisions can be undirected (Daw et al., 2006). Various models of exploitation-exploration trade-off are frequently modeled as a variant of the matching law (called “softmax” or Boltzmann decision rule), where the proportion of choices depends on the difference in their expected values. In the softmax rule, the “inverse temperature” parameter represents the sensitivity to the difference of values: larger the inverse temperature, the more likely one chooses the expected optimal action. A low inverse temperature parameter can be considered as risk-taking (too much exploration), while a high inverse temperature reflects risk aversion (too much exploitation). Other models for exploration have been proposed, where exploratory decisions are stochastic or guided by the expected gain of information (Daw et al., 2006; Keramati et al., 2011).

Value-based decision-making critically involves learning, in order to transform information gained from past choices into a common currency – value. In habitual processes, values are learned by trial and error, based on a teaching signal called reward prediction error by learning theories (Schultz et al., 1997; Schultz, 2002). The prediction error,  $\delta$ , is defined as the difference between an expected reward and the reward actually obtained. An outcome better than expected (i.e. a positive prediction error) leads to an increase in the value of the preceding action, while negative errors decrease this value. This theory thus suggests that no further learning occurs when the reward is completely predicted by the cue, i.e. the prediction error is null. Different models of learning based on  $\delta$  have been proposed (variations of the Rescorla–Wagner rule, Rescorla and Wagner, 1972). Some of these models interpret the learning rate (how the prediction error affects the update of the value) as the salience or the associability of the stimulus with the reward (Pearce–Hall theory, Pearce and Hall, 1980). Modulation of this learning rate can lead to complex effects on decision-making processes, such as alterations in reward sensitivity, but also risk-sensitivity (Niv et al., 2012, see below).

Different factors can modulate the value associated with a given action. In particular the estimations of risk and the delay associated with a given choice represent important modulators that are associated with modifications of decision-making in smokers. “Temporal discounting”, a measure of impulsive choice, refers to inter-temporal choices where values decrease with delay. A very steep delay discounting indicates a situation where one prefers a small immediate reward over a big delayed one. “Probability discounting” represents the possible over- or underestimation (and thus valuation) of the probability to obtain variable rewards. Individuals are risk-neutral if they choose according to the expectation of average reward (e.g. they are indifferent between the options of 50% chance to receive two rewards versus 100% of receiving one). Risk-taking and risk-aversion correspond respectively to the over- and under-valuation of the uncertain choice.

Furthermore, these theories are mainly applicable to situations

with a positive reinforcing (appetitive) outcome. The positive or negative valence of a stimulus recruits specific decision-making brain areas. Stimuli predicting a potential reward activate systems dedicated to seeking, approach and consummation (Berridge and Robinson, 1998; Ikemoto and Panksepp, 1999). Predictions of punishment have more complex and wide-ranging effects on behavior, as they evoke mainly behavioral inhibition, but also fight or flight responses depending on the threat. Therefore, impulsive action may appear as an imbalance between appetitive approach and behavioral inhibition (Boureau and Dayan, 2010).

This decision-making framework provides perspective on the classification of behavioral modifications that can be observed in smokers, as well as in mental disorders (Montague et al., 2012; Kishida et al., 2010). On the one hand, normative theories of optimal choices provide a quantitative way to detect when subjects' choices match or deviate from an optimum, for example after smoking. On the other hand, the computational frameworks provide objective measures to evaluate dysfunctions that can be assigned to changes in the free parameters of the model (e.g. learning rate, inverse temperature, delay discounting factor). These dysfunctions can eventually be linked to altered neural activity. Deficits in decision-making systems can thus reflect alterations in i) utility, i.e. sensitivity to reward or punishment, ii) delay discounting, defined as the tendency to choose a small or inferior immediate reward over a larger delayed reward (Clark et al., 2004; Evenden, 1999; Peters and Büchel, 2011), iii) risk sensitivity, iv) learning rate, v) inverse temperature, i.e. sensitivity to the differences of rewards, v) balance between approach and inhibition, or vi) balance between goals and habits. This non-exhaustive list is intended as a characterization tool for this review, for a thorough description of computational frameworks applied to addiction, see for example Redish et al. (2008).

### 3. Effects of nicotine on decision-making

In humans, substance-use disorders, and particularly tobacco use disorder, are highly correlated with impulsivity (Moeller and Dougherty, 2002; de Wit and Richards, 2004; Perry and Carroll, 2008). Nicotine-dependent individuals score higher than non-dependent controls on self-report measures and questionnaires that assess impulsivity (Mitchell, 1999). Moreover, smoking, or self-administered nicotine (Fowler et al., 2008), is also associated with impaired performances on decision-making tasks that require cognitive control and flexibility (Lejuez et al., 2003; Nesic et al., 2011; Xiao et al., 2008). According to the decision-making framework, human smokers seem generally more exploitative (i.e. higher inverse temperature) than non-smokers. They show an increased sensitivity to the difference of value in a multi-choices task ("bandit task", Addicott et al., 2013), and thus focus more on the highest expected reward at the expense of exploring alternatives. In delay-discounting tasks, current cigarette smokers show a greater preference for immediate monetary and cigarette rewards over larger, delayed rewards than non-smokers and ex-smokers (Baker et al., 2003; Bickel et al., 1999; Businelle et al., 2010; Mitchell, 1999; Johnson et al., 2007). In contrast, data on nicotinic implications in probability discounting are limited and contradictory. Probability discounting seems to be altered in smokers compared to non-smokers, although the direction of these alterations (increased or decreased risky decision making) is unclear (Mitchell, 1999; Reynolds et al., 2004). Finally, Chiu et al. (2008) have directly based their analysis on reinforcement learning models in a monetary investment task. They distinguished between two classes of learning (error) signals, actual and fictive (or counterfactual, i.e. "what might have happened"). The fictive error signals in healthy individuals strongly predicted changes in investment behavior,

which correlated with fMRI signals measured in choice-related dopaminergic structures (see below). However, investment behavior of chronic smokers was not guided by these error signals, despite ongoing and robust neural correlates of these fictive errors (Chiu et al., 2008; Lohrenz et al., 2007). These examples from human studies suggest that nicotine exposure modify several decision-making processes. However, modified decision-making (e.g. delay discounting or risk-taking) may also act as a risk factor for substance abuse and could hypothetically even predict substance use. For example, delay discounting has been suggested to be a heritable trait, resulting from genetic and environmental factors (Anokhin et al., 2011), associated with the onset and maintenance of drug use. Therefore, impulsivity, as a global definition, is considered an endophenotype of substance use and addictive disorders (Ersche et al., 2010; Verdejo-García et al., 2008). Altered decision-making processing in smokers can thus derive from pre-existing vulnerabilities, present prior to substance use (Audrain-McGovern et al., 2009; Dalley et al., 2011) and from the active harmful effects of drug use that affect brain structures and functions. These findings suggest that the relationship between tobacco addiction and alterations in decision-making behavior is bidirectional, i.e. impulsivity is a risk factor and nicotine exposure increases impulsivity.

Animal studies are of utmost importance to dissect the causal mechanisms of this bidirectional relation, since they provide multiple possibilities to control pre- and post-testing components. First, in support to what has been observed in humans, rodents that score high on measures of impulsivity (impulsive action and choice) are typically more likely to develop addictive-like behaviors than those that score low on such measures (e.g., Belin et al., 2008; de Wit, 2009; Diergaarde et al., 2008; Perry and Carroll, 2008; Perry et al., 2005). This strengthens the notion that altered decision-making processes can pre-exist and favor the development of addiction. But the opposite is also true. In rodents, and independently of the state of the animal, acute nicotine dose-dependently increased the preference for an immediate reward over a larger, delayed reward (Dallery and Locey, 2005), an effect blocked by the nicotinic receptor antagonist mecamylamine (Kolokotroni et al., 2011). In contrast, Anderson and Diller (2010) found that acute injections of nicotine dose-dependently decreased the preference for a smaller, immediate reward in rats, while repeated injections tended to reduce this effect (Anderson and Diller, 2010). It should be noted however that nicotine effects on choices between a large, delayed reward and a small immediate one could be explained by differences in reward sensitivity rather than modifications of the delay-discounting factor, as both parameters vary in these paradigms. Indeed, it has been found that nicotine effects on delay discounting depend on the amount of food offered as an alternative (Locey and Dallery, 2012, 2009). In this study, the alteration of delay discounting was better explained by a decrease in the sensitivity to the large reward. Hence, nicotine might affect the reward utility function (sensitivity to the amount of food) and/or the inverse temperature (sensitivity to the difference of expected rewards). Similarly to human studies, data on probabilistic discounting showed contradictory results. In rats, acute nicotine administration has been found to increase (Mendez et al., 2012) or decrease (Mitchell et al., 2011) risky choices. However, Mendez et al. (2012) noted that depending on the order of presentation of reward probabilities, the effect of nicotine was different. This suggests that nicotine may impact behavioral flexibility by affecting the updating of values after acquisition, or how learned values impact futures decisions.

Effects in rodents do not mirror systematically what is observed in humans. These discrepancies do not contradict the effect of nicotine on decision-making, but are rather a consequence of

different experimental factors. Some of these factors can be explained by specific elements of the tasks used to measure choice or risk preferences in animals (Heilbronner and Hayden, 2013), e.g. the size of the rewards or the delays between trials. Besides, nicotine, as many psychoactive drugs, has an inverted U-shaped dose–effect in multiple cognitive processes. Conflicting behavioral effects elicited by nicotine can thus be the consequence of nicotine activation of “antagonist” brain pathways, resulting from the activation of different combinations of nAChRs subtypes. Furthermore, these effects could vary with small modification of the nicotine dose, the regimen of injection and the duration (acute vs. chronic) of the intake, but also the initial state of the animal (Picciotto, 2003). Indeed, routes and patterns of administration (I.V., I.P., continuous, discontinuous, Matta et al., 2007) differentially affect nAChRs subtypes and consequently the downstream cascades of events. Moreover, the mere control of the concentration of nicotine in rodents has been proven difficult, as evaluating the state of smokers (sated or not) at the time of the test. In the following sections, we review the studies delving further into the neurobiological mechanisms that likely underlie the diversity of nicotine's effects on decisions. Beyond the scope of this review, it should be noted that endogenous acetylcholine also acts on muscarinic receptors that are not targeted by nicotine. This might explain some of the discrepancies found in studies of the normal role of acetylcholine on decision-making and that of nicotine.

#### 4. Nicotinic acetylcholine receptors

The bidirectional relationship between altered decision-making and tobacco addiction likely arises from the overlap between neural structures (i.e. the mesocorticolimbic loop) that are implicated in value-based choices, and that express acetylcholine nicotinic receptors (nAChRs). nAChRs are ligand-gated ion channels widely expressed throughout the central nervous system (CNS), at pre-, post- and extra-synaptic locations. They act as a neuromodulator and influence neuronal excitability and activity-dependent plasticity in most, if not all, mammalian brain structures. nAChRs activation by ACh or nicotine consists of the fast opening (microsecond to millisecond range) of a cationic channel. nAChRs currents have relatively fast kinetics (albeit slower than glutamatergic fast neurotransmission) compared to the muscarinic acetylcholine receptor. nAChRs are transmembrane allosteric oligomers composed of five identical or different subunits (homo- or hetero-pentamers, respectively). Nine  $\alpha$  ( $\alpha 2$ – $\alpha 10$ ) and three  $\beta$  ( $\beta 2$ – $\beta 4$ ) subunits have been identified in the vertebrate brain. These subunits can assemble in different combinations to construct receptors with distinct functional and pharmacological properties. Depending on the subunit composition, the receptors show very different affinity for ACh or nicotine (Changeux and Edelstein, 2005). For example, the  $\alpha 7$  homopentamer has a low affinity for ACh and nicotine, a high calcium permeability and rapid activation and desensitization kinetics, while the  $\alpha 4\beta 2$  containing (\*) heteropentamers (asterisk indicates the possible presence of other subunits) are highly sensitive to agonists and strongly up-regulate in response to chronic nicotine exposure. Both  $\alpha 4\beta 2$  and  $\alpha 7^*$ nAChRs are heavily expressed in the dopaminergic nuclei (the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc), the amygdala, the medial habenula, the basal ganglia and the prefrontal cortex, i.e. the value-based decision systems.

In general, it is still not clear whether acetylcholine provides a phasic signal (restricted in time and space) directly activating neuronal targets, or a rather slow, tonic change in excitability to modulate the responsiveness of neurons. There is an ongoing debate between the existence of cholinergic synapses (in the

traditional sense) throughout the brain (Lamotte and Ascher, 2014; Sarter et al., 2009) or a more global “volume conduction” mechanism of action. Furthermore, the dynamics of acetylcholine release at short timescales are just beginning to be unveiled. For example, in the prefrontal cortex, attention (e.g. detection of relevant environmental cues) correlates with second-timescale variations of ACh levels (Parikh et al., 2007). However, it is not clear whether this is also true in other brain structures. For instance, the pedunculopontine tegmental nucleus (PPTg) is involved in reward, sensory and motor processing, (Pan and Hyland, 2005; Thompson and Felsen, 2013). The PPTg, together with the laterodorsal tegmental nucleus (LDTg), are important sources of cholinergic inputs to the VTA (see below). However, currently nothing is known about the dynamics of ACh release (presumably for these structures) in the VTA, and how nicotine could alter this input signal to the VTA. As the PPTg and LDTg also provide glutamatergic excitation to the VTA (Clement and Grant, 1990), the link between their neuronal activity and the dopaminergic one during decision-making tasks remain obscure. Overall, the standard view of ACh as a neuromodulator (i.e. a global signal affecting the neural networks' dynamics) provides an appealing explanation for the nicotine effects on decision-making unrelated to cigarettes. Rather than providing a specific (i.e. local in time and space) signal potentially related to drug taking, nAChRs activation by nicotine would control networks dynamics (i.e. the way neural networks process information) with general consequences beyond the discrete decision of having a cigarette.

Long-term dynamics of nAChRs are also of utmost importance. Chronic exposure to ACh or nicotinic drugs causes a gradual decrease of evoked currents, leading to a desensitized closed state of the receptor (Pidoplichko et al., 1997). Furthermore, chronic exposure to nicotine causes a drastic increase (up-regulation) in the total number of high-affinity nAChRs, in humans and rodents (Buisson and Bertrand, 2002; Marks et al., 1992; Wonnacott, 1990). Together, these two mechanisms underlie a common pattern of cigarette smoking (Dani et al., 2001) and suggest that a smoker can experience different forms of the overall “nicotinic state” (number and state of the receptors) during the day. After a night of complete abstinence, nicotine concentrations in the brain are at their lowest level and thus nAChRs have recovered from desensitization resulting into an excess (due to the up-regulation) of excitability of the systems under the influence of the endogenous nicotinic neuromodulation. In contrast, during the day and after a few cigarettes some subpopulations of nicotinic receptors may have entered the desensitized state, depriving the neurons or synapses from its normal endogenous modulation by ACh. Apart from these two effects, chronic nicotine exposure may also modify circuits through synaptic and intrinsic plasticity, through signal transduction mechanisms arising from  $\text{Ca}^{2+}$  influx through nAChRs. These adaptations can arise downstream from the receptors. It has been shown for example that nicotine induces LTP at VTA DA excitatory synapses, 24 h after in vivo administration of the drug (Mansvelder and McGehee, 2000), an effect that required activation of presynaptic  $\alpha 7^*$ nAChRs. However, nicotine may activate nAChRs at any given level in the neural circuitry (e.g. dopaminergic areas), while resulting in long-term plasticity or other modifications at another locus. For instance, nicotine-induced DA release produces a rapid and transient increase in ERK phosphorylation in the nucleus accumbens 15' after in-vivo administration (Valjent et al., 2004; Tolu et al., 2013), directly causes in vivo hippocampal synaptic potentiation throughout a local hippocampal dopamine signal arising from the midbrain (Tang and Dani, 2009). Nicotine may also act indirectly via the stress hormone response, as it activates the HPA axis to increase corticosterone, which subsequently affects ethanol-induced DA activity (Doyon et al., 2013). Finally, chronic

exposure to nicotine can also change the stoichiometry of the receptors (i.e. the relative proportions of subunits in heteromeric receptors) and thus impact the cholinergic activity. For example, although  $\alpha 5$  subunit does not contribute to the nicotine-binding site (Ramirez-Latorre et al., 1996), its deletion leads to a dramatic shift in several nicotine-elicited alterations. In particular, it has been suggested that in the VTA, surprisingly,  $\alpha 4\beta 2^*$  nAChRs that do not comprise  $\alpha 5$  have a minor role in the nicotine-evoked response of DA cells (Morel et al., 2014). The fact that  $\alpha 5$  subunits in nAChRs appear to reduce the magnitude of nicotine up-regulation (Mao et al., 2008) can thus change the proportion of  $\alpha 4\beta 2\alpha 5^*$  nAChRs in the VTA, or in downstream areas such as the PFC. Consequently, both the response to nicotine after long-term exposure, and other decision processes implicating  $\alpha 5$  (see below), may be altered. Hence, one should take into account all these parallel modifications induced by nicotine, which are not restrained to nAChRs. Dissecting one of these parallel processes is complicated and may require the development of specific approaches to overcome these challenges, e.g. fluorescently tagged subunits to assess the pattern of up-regulation (Nashmi et al., 2007).

In the two next parts we will describe known roles of nAChRs in specific brain regions involved in decision-making processes. The anatomical substrates for decision-making are very distributed (Doya, 2008), but in this non-exhaustive review we will focus on six of its major components, divided in i) the dopaminergic system and ii) its main downstream targets, mainly the prefrontal cortex and the striatum, and iii) we will also briefly discuss the implication of the hippocampus, the habenula and the amygdala (see below).

## 5. Nicotinic control of dopaminergic activity

The dopaminergic (DAergic) nuclei, with their projections to the striatum and prefrontal cortex (PFC), play important roles in the relation between acetylcholine, tobacco addiction and decision-making unrelated to nicotine intake. DA signals are thought to represent the core of value-based decisions. For example, dopaminergic neurons encode reward prediction errors (see above) as phasic bursts of action potentials. Reinforcement learning theories assign a teaching role to dopamine, as this prediction error may be used to learn the value of a stimulus or action. Other theories depict a more immediate role of dopamine in motivational processes, where the dopaminergic prediction of reward would affect the incentive salience of a stimulus to drive behavior towards a reward, promoting appetitive and seeking behaviors (Berridge, 2007; Ikemoto and Panksepp, 1999).

A peculiar relation between DA and nAChRs is revealed by observation that i) effects on value-based decisions induced by pharmacological manipulations of ACh (that target muscarinic but also nicotinic receptors) often mirror those of dopamine (Fobbs and Mizumori, 2014) ii) that DA cells are embedded in a complex network of nicotinic modulation at a somato-dendritic levels but also at a terminal levels (Maskos, 2010; Exley and Cragg, 2008). Dopaminergic nuclei are not only composed of dopaminergic neurons, but also of inhibitory (GABAergic) ones. Both types of cells express nAChRs, with mostly  $\beta 2$  and  $\alpha 7$  subunits, at their dendritic, somatic, and axonal (some of them situated in the striatum and in the cortex) compartments. Glutamatergic inputs from other parts of the brain also express  $\alpha 7$  subunits. Cholinergic innervation of dopaminergic nuclei arises from the mesopontine nuclei, the laterodorsal tegmental nucleus (LDTg) and the pedunculopontine tegmental nucleus (PPTg). The LDTg projects mainly to the VTA (Cornwall et al., 1990), while the PPTg innervates predominantly the SNc, with its posterior part also innervating the VTA (Oakman et al., 1995; Joel and Weiner, 2000). However, the PPTg has also been shown recently to innervate also the striatum (Dautan et al.,

2014, see below). PPTg is now seen as an important actor in the decision-making network. Activity of PPTg neurons reflects sensory cues, goal-directed actions and their outcomes in decision-making tasks (Thompson and Felsen, 2013; Pan and Hyland, 2005; Norton et al., 2011). PPTg inactivation strongly decreases reward prediction errors (after learning occurred) in freely moving rats (Pan and Hyland, 2005). A number of data thus suggest that some PPTg neurons could participate to the computation of reward predictions in the dopaminergic system (Humphries and Prescott, 2010). Nevertheless, PPTg lesions affect but do not suppress the learning of instrumental tasks (Wilson et al., 2009), hence it is unlikely that it constitutes the sole reward-related input to the DA nuclei. If the relation between the nature of PPTg cells (Glutamatergic, GABAergic or Cholinergic) and their activity in decision tasks is yet unknown, cholinergic inputs to dopaminergic neurons seem to play a key role in the generation of DA bursting activity. Mainly,  $\beta 2$ -containing nAChRs ( $\beta 2$ nAChRs) have an important role in DAergic bursting activity, since in the absence of the  $\beta 2$  nAChR subunits, DA neurons lack spontaneous bursting in anaesthetized animals (Mameli-Engvall et al., 2006; Tolu et al., 2013). Besides, nAChR are characterized by their property to switch to desensitized states. Prolonged exposure to low concentrations of nicotine stabilizes the receptor in a closed (desensitized) state, which blocks the responsiveness to agonists (i.e. both endogenous ACh and nicotine). Thus, chronic exposure to nicotine disconnects neurons, and particularly dopaminergic ones, from their physiological cholinergic drives. Especially, desensitization of  $\beta 2^*$ nAChRs can have an important effect on VTA DA cell dynamics (see below).

Nicotine exerts its reinforcing properties by acting on nAChRs of the VTA dopaminergic nucleus, inducing bursts of action potentials in dopaminergic cells (Mameli-Engvall et al., 2006; Maskos, 2010; Tolu et al., 2013). There is considerable evidence that environmental drug cues participate in addictive behavior. In both human and animals, cues reliably elicit craving and prompt the initiation of drug seeking/taking and relapse following cessation treatment. However, the mechanisms by which this occurs are not clear. Redish (2004) proposed a model of how drugs of abuse, by mimicking reward, may hijack normal value learning, which results in drug abuse. DA neurons could mimic the properties of the delta signal in temporal difference (TD) learning. Experimental data and theories predict that for natural reinforcers, reward-driven DA activity should diminish as conditioned stimuli is learned to be reward-predictive (Schultz, 2002), i.e. the prediction error is zero at the time of reward delivery, as the predicted reward equals the actual one. Redish (2004) proposed that this does not happen for addictive drugs. In contrast, after learning, the phasic DA signal might be observed for both the predictive CS and the drug delivery itself. Due to nicotine, the prediction error (induced pharmacologically at the time of the cigarette CS) is always positive and therefore the value of the cigarette keeps increasing, leading to overvaluation and addiction. However, it should be noted that at the level of the DA cell activity and/or the release of DA, there is a drastic difference between the timescale of reward prediction error (sub-second) and of nicotine-induced bursting (several minutes), so if these DA responses actually interact remains an open question. To date, the sequence of modifications of firing patterns in VTA DA cells during the transition to the nicotine-addicted state and withdrawal symptoms has never been studied. Recordings of the activity of VTA neurons in the same animals, during the different phases of addiction (i.e. before nicotine exposure, during repeated acute injection, and during craving) is a clear challenge, but would provide valuable information on modifications of DA systems during nicotine exposure (Redish, 2004). For instance, despite acting on DA re-uptake rather than DA firing, cocaine can induce a reward prediction error-like activity at the time of a cue with which cocaine

has been associated (Phillips et al., 2003).

The long-lasting increase in dopaminergic activity following acute nicotine delivery is likely a key mechanism in the effects of nicotine on decision-making. If nAChRs activation by nicotine controls the excitability levels of dopaminergic cells (rather than simply provide a drug-related reinforcement signal), it could explain the continuous effects of nicotine on DA and decision-making. The level of firing by dopaminergic cells affects the tonic levels of DA in target structures (Grace et al., 2007). Tonic DA levels have been proposed to control the exploration/exploitation trade-off (Beeler et al., 2010; Humphries et al., 2012). Increased tonic DA following acute nicotine would favor exploitation, as seen in smokers (Addicott et al., 2013). This provides an explanation for the lack of flexibility that is associated with nicotine. However, if phasic DA is also increased by nicotine, it is somewhat contradictory with findings of decreased reward sensitivity (Locey and Dallery, 2009, 2012). In addition, Niv et al. (2007) proposed that tonic DA levels would signal the average reward, so that alterations of these levels by nicotine would be likely to affect the perception of the average potential rewards, and potentially the vigor of actions (Niv et al., 2007).

In experiments assessing the sensitivity of the brain reward system using intracranial self-stimulation (ICSS), both acute injection of nicotine and self-administration decreased the threshold (minimal current to obtain ICSS behavior), i.e. increased the sensitivity of the reward system (Skjei and Markou, 2003). On the contrary, withdrawal from chronic nicotine increases the ICSS threshold, persisting at least four days, and thus a decrease in reward sensitivity (Epping-Jordan et al., 1998). Hence, under nicotine, rewarding outcomes may be overvalued and the behavior may be more exploitative (increased sensitivity to the difference in value) and less inhibited (by an imbalance between appetitive and aversive systems). The biological substrates underlying this effect is however unclear. Modifications at the level of the DA system can explain some aspects of these alterations in ICSS sensitivity. Acute nicotine increases the firing of the VTA DA cells. However, effect on chronic nicotine exposure on DA cell firing rate in the dopaminergic nuclei are ambiguous due to varied experimental conditions (route of administration; dose; duration of exposure; continuous exposure, see above) Nicotine self-administration for 2 months causes hyperactivity of dopaminergic neurons in the VTA, while no increase in firing rate or bursting was found after passive delivery of nicotine through mini-pumps (Besson et al., 2007; Caillé et al., 2009). However despite absence of apparent modification, homeostatic processes involving both  $\beta_2$ - and  $\alpha_7$ \*nAChRs occur in mice chronically exposed to nicotine (Besson et al., 2007). However, despite this absence of modification in the firing rate of DA neurons, it has been proposed that adaptations occurs at the levels of the nAChRs control of VTA DA cells (Besson et al., 2007). Finally, using higher concentration of nicotine in mini-pumps (2 mg/kg/h), Xiao and colleagues showed a decrease in the firing rate of SNC DA neurons (Xiao et al., 2009), one day after removing the mini-pumps, which may correspond to a withdrawal state (Epping-Jordan et al., 1998). At the terminal level, chronic nicotine decreases the tonic DA neuron activity through desensitization of the nAChRs (Zhang and Sulzer, 2004). This chronic altered state of the nAChRs has an overall diminishing effect on DA release in downstream targets (Koranda et al., 2014). Therefore, the hypothesis of Niv et al. (2007) that tonic DA signals the average reward could account for the alterations in reward sensitivity, since the baseline of tonic DA activity may be reduced.

## 6. Nicotinic control of the downstream DA targets

It is believed that the prefrontal cortices, due to their associative inputs and high encoding capacity (as recurrent networks) are

more involved in the reflective/goal-directed decision system (Dolan and Dayan, 2013). Populations of prefrontal neurons encode expectation of future value (average reward, i.e. reward probability times reward amount), reward amount and reward probability separately (see Padoa-Schioppa, 2011 for review). Frontal areas, such as the medial and lateral orbitofrontal cortex, the ventromedial prefrontal cortex, anterior cingulate cortex, the anterior lateral prefrontal cortex, but also more parietal areas (lateral inferior and posterior parietal cortex) contribute differentially to decision-making (Rushworth et al., 2011). Distinguishing between the nicotinic effects on these contributions is beyond the scope of this review. Globally, the PFC would learn an internal model of the environment and build a decision tree of possible future states and actions, providing cognitive flexibility (Daw et al., 2005, 2006; Dolan and Dayan, 2013). It should be noted however that parts of the prefrontal cortex (e.g. infralimbic cortex) have been demonstrated to take part in the habitual system, restricting goal-directed decision-making to the paralimbic cortex (Daw et al., 2005; Dolan and Dayan, 2013). The PFC also provides cognitive control over habitual actions (and is thus implicated in behavioral inhibition) and exploitive trends, hence promoting exploration (Daw et al., 2006). Anterior parts of the frontal cortices may be implicated in counterfactual decisions, based on what “might have been”, which may be altered in smokers (Chiu et al., 2008). The parietal cortex is also implicated in decision-making (albeit considered as more “perceptual”) and may be implicated in impulsivity, as it plays a role in setting the speed-accuracy tradeoff (Kim and Lee, 2011). nAChRs play a key role in the modulation of PFC activity and the frontal cortices receive acetylcholine inputs from nuclei in the basal forebrain (Hasselmo and Sarter, 2011). Nevertheless, the basal forebrain may not be the sole source of cholinergic modulation of cortical dynamics, as the PPTg also projects to the thalamus (Winn, 2006). From a functional point of view, overexpression of the  $\alpha_3/\alpha_5/\beta_4$  genomic cluster in mice, which increases the levels of expression of these subunits in the PFC and hippocampus (Gallego et al., 2012), augments behavioral inhibition (Viñals et al., 2012). The PFC also naturally highly expresses  $\beta_2$ \*nAChRs, implicated in selective attention (Guillem et al., 2011) and social interactions (Avale et al., 2008; dos Santos Coura and Granon, 2012), functions that likely require inferential or reflective processes (Yu and Dayan, 2002, 2005; Devaine et al., 2014). Moreover, nicotine robustly activates layer VI neurons, implicated in top-down contextual information and attentional control, via  $\alpha_4\alpha_5\beta_2$  receptors (Hay et al., 2014). It has been proposed that the  $\alpha_5$  subunit plays a key role in attention circuitry (Bailey et al., 2010), probably together with the  $\beta_2$  subunit (Guillem et al., 2011). As frontal cortices are also implicated in impulse control (i.e. inhibition of habitual processes)  $\alpha_5$  subunit, modifications of  $\alpha_4\alpha_5\beta_2$  expression after chronic exposure could impact the PFC processing, and thus impulsivity. Nevertheless, it seems unlikely, as  $\alpha_5$  subunits seem to confer a resistance to up-regulation by nicotine (Mao et al., 2008). Overall, these cholinergic implications in cognitive flexibility (the “reflective system”) have been formalized by a computational model (Yu and Dayan, 2002, 2005). This theory states that forebrain acetylcholine input might represent expected uncertainty (known unpredictability of an outcome following a cue) in cortical inference. This theory accounts for a few cognitive effects ascribed to nicotine (e.g. in the attentional Posner task) and for the implications of nAChRs in reflective decisions. However, if it correctly accounts for pro-attentional effects of acute nicotine, it is unclear how it can shed light on the imbalance between reflective and habitual systems following chronic nicotine and nAChRs up-regulation. Besides, the orbitofrontal cortex has been found to be implicated in delay discounting (Kable and Glimcher, 2007), a paradigm that indeed is affected by nicotine (Kelsey and Niraula, 2013). However, mice

lacking  $\beta 2^*nAChRs$  do not demonstrate altered delay discounting (Serreau et al., 2011). So whether the effects of chronic nicotine in delay discounting rely on modifications induced by, but not expressed by nAChRs, remains an open question.

The different parts of the basal ganglia are implicated in various forms of value-based decision-making and motor control. The structure of the basal ganglia, composed of two stages of inhibitory neurons (striatum and output nuclei) is ideally suited for action selection. One behavioral output can be dis-inhibited, while inhibiting others (Redgrave et al., 2011). Ventral parts of the basal ganglia (e.g. nucleus accumbens) are implicated in stimulus-outcome associations and Pavlovian appetitive approach towards rewards (Yin et al., 2008). The ventral striatum might constitute a locus of convergence of different decision systems, as it has been shown in humans to represent values and reward predictions, regardless of how (i.e. reflectively or reflexively) they may be computed (Daw et al., 2011; Simon and Daw, 2011). The dorsolateral and dorsomedial parts of the striatum seem respectively involved in habits (reflexive decision) and in goal-directed actions (reflective decision, together with the prefrontal cortex) (Yin et al., 2008). Neural correlates of action value have also been found in the striatum (Samejima et al., 2005; Lau and Glimcher, 2008; Kim et al., 2013; Cai et al., 2011). The synapses from the cortical neurons to the medium spiny neurons in the striatum are often hypothesized to be the primary locus for reinforcement learning, i.e. for updating and storing the values of states and actions (Reynolds et al., 2001; Hikosaka et al., 2006; Lo and Wang, 2006; Hong et al., 2011). Basal ganglia nuclei are not only composed of inhibitory neurons, but also of cholinergic ones, called tonically active neurons (TANs). The inhibitory cells express nAChRs and are thus sensible to this local source of ACh, but also to cholinergic inputs coming from the PPTg (Dautan et al., 2014). ACh also affects dopaminergic release due to the expression of nAChRs on dopaminergic terminals (Exley and Cragg, 2008; Koranda et al., 2014). Because of their correlated activity and widespread innervation throughout the striatum, TANs may play a major role in orchestrating the activity of the basal ganglia. A computational model (Stocco, 2012) proposed that striatal ACh might control the exploration-exploitation trade-off. In this model, the TANs provide a global inhibitory signal, controlling the number of actions that can be selected, and the level of competition between them. This is similar to modulating the inverse temperature parameter, as increased inhibition reduces the number of competing alternatives (i.e. those with higher value), which corresponds to increased exploitation. If nicotine indeed increases this global inhibitory signal, this might be consistent with the implication of nicotine in alterations of reward sensitivity. However, mice lacking  $\beta 2^*nAChRs$  display similar preferences for large versus small rewards as WT (Serreau et al., 2011), and are not impaired in motivation for food (Guillem et al., 2011). Again, future studies are needed to disentangle the different roles of endogenous ACh and over-activation by nicotine and to dissect neural modifications due to induction from those due to expression levels after chronic nicotine. The striatum is one of the principal loci displaying up-regulation of nAChRs (Govind et al., 2009) and likely plays a major role in the long-term alterations of decision-making induced by nicotine. Moreover, both chronic nicotine exposure and withdrawal durably affect dopamine release in the striatum (Rahman et al., 2004; Zhang et al., 2013), probably affecting a wide range of decision processes.

The cortex and striatum likely underlie different computations (respectively based on recurrent dynamics and selective disinhibition), rather than directly performing a decision, whether reflective or reflexive. The structures interact with each other, as the habitual and goal-directed “systems” do. Exposure to drugs can shift the normal balance between decision systems, promoting one

system over the other. Deficits in cognitive flexibility may arise from a shift towards more habitual than goal-directed control. Drugs that give more weight to the structures involved in the habit system or that impair structures involved in the planning system would favor the automation of behaviors (Redish et al., 2008), related to drug taking or not. In this framework, it has been proposed (Daw et al., 2005; Keramati et al., 2011) that the arbitration between goal-directed and habitual systems may depend on their relative accuracies. Arbitration might depend on the relative uncertainties of the systems, as tree-search decisions may be hard to compute, and habits are slow to learn (Daw et al., 2005). It provides an account of the gradual shift from goal-directed actions to habits with training. As ACh has been proposed to signal uncertainty (although in the different context of cortical inference, Yu and Dayan, 2005), nicotine might hypothetically bias the arbitration in favor of the habitual system, by altering the physiological signaling of uncertainty. Finally, nicotine-induced increases in DA would favor approach (Ikemoto and Panksepp, 1999; Berridge, 2007) over behavioral inhibition, resulting in more action impulsivity. This suggests that dopamine may mediate the effects of nicotine in the balance between goals and habits. However, DA agonists seem to favor goal-directed, reflective processes over habitual systems (Dolan and Dayan, 2013).

NAChRs are also expressed in the amygdala, hippocampus and the medial habenula, structures also known to play a role in value-based decision-making. The hippocampus may form a “cognitive map” of the environment, and thus play a role in the planning/reflective decision system. Together with the prefrontal cortex, the hippocampus is also involved in working memory. Hippocampal infusion with blockers of nAChRs impaired short- and long-term memory retrieval, while nicotine enhanced memory retrieval in rat behavioral tests (Martí Barros et al., 2004). The hippocampus plays a major role in the identification of “states” (i.e. recognition), a foundation upon which associations between states and values are built. In a variety of recognition tests (novel object, social, place),  $\alpha 7$  agonists diminish natural forgetfulness caused by long delays between trials (Hauser et al., 2010; Pichat et al., 2007; Prickaerts et al., 2012; Wallace and Porter, 2011).

The amygdala is believed to store learned values for positive and negative outcomes. Nicotine enhances synaptic transmission onto the amygdala (Barazangi and Role, 2001; Dani and Bertrand, 2007), which may be related to alterations of the utility function and to the balance between appetitive approach and behavioral inhibition. Neurons from the medial habenula (MH) are also implicated in aversive processing. They have been found to encode predictions of aversive outcomes (Matsumoto and Hikosaka, 2009) and inhibit dopaminergic areas (Lammel et al., 2011). Therefore, the MH may also be implicated in the aversive effects of nicotine (Fowler and Kenny, 2014). As an inhibitory input on the dopaminergic areas, the MH could be involved in the alterations of behavioral inhibition following nicotine exposure. However, there is a lack of studies implicating the MH in the chronic effects of nicotine on valued-base decision-making.

## 7. Discussion

In this review, we aimed at stressing some overlooked points regarding nicotine and decision-making unrelated to behaviors directed towards drug intake. The observed correlation between alterations of decision-making and tobacco addiction deserves further attention. Indeed, addiction is often seen as the result of maladaptive decision-making, while the causal relation is unclear and sometimes reversed. Acute or chronic nicotine can affect subsequent decisions, not restricted to addiction itself. A further understanding of these issues requires more formalization of

decision-making, and relevant experimental paradigms. Smokers have been described as impulsive, an ambiguous notion that may encompass various forms of value-based decision-making, ranging from probabilistic or delay discounting to the balance between goals and habits, or between appetitive approach and behavioral inhibition. Despite recent efforts (Redish et al., 2007; Kishida et al., 2010), the links with computational psychiatry (Montague et al., 2012) remain scarce. Several computational theories on the role of physiological acetylcholine in signaling uncertainty (Yu and Dayan, 2005) or entropy (Stocco, 2012), as well as on the arbitration between reflective and reflexive systems, provide testable working hypotheses. Moreover, animal studies are needed to delve into the neurobiological mechanisms underlying nicotine effects on decision-making. All the major structures implicated in decision-making highly express nicotinic receptors, but often receive endogenous acetylcholine from distinct nuclei, which are likely involved in different computational processes. This may explain the diversity of nicotine effects on behavior. In particular, the diversity in biophysical properties and location of nAChR subtypes offers a wide range of possible fine regulation of the DA system. The VTA itself seems more heterogeneous than previously thought (Bromberg-Martin et al., 2010; Lammel et al., 2011), and DA neurons respond differentially to reward and punishments depending on their preferred target structure. Therefore, it would be interesting to link, if existent, specific properties of nAChRs and specific pathways in the VTA (i.e. mesocortical, mesolimbic). Last, the interactions between nicotine and mental illnesses at the level of decision-making are not widely studied yet, but could provide a broader perspective on nAChRs and modified behavior. Tobacco addiction is more prevalent among people suffering from mental illnesses (schizophrenia, depression and anxiety, mood or personality disorders, addiction for other substances, pathological gambling) than in the general population (Dani and Harris, 2005). These mental illnesses can also be considered as pathologies of decision-making, and smoking seems to interact with at least some of their symptoms. Due to common neural pathways, addiction to nicotine and other mental illnesses may be synergistic. Factors like stress, anxiety and depression contribute to the vulnerability to nicotine addiction, and, like nicotine intake, have consequences for dopaminergic signaling (e.g. synaptic plasticity, Jones and Bonci, 2005; Saal et al., 2003). However, nicotine might counteract some symptoms of mental illnesses and thus be used as a self-medication. This has been proposed for schizophrenia (to regulate dopaminergic activity), attention-deficit hyperactivity disorder and Alzheimer's disease (to restore attentional deficits). Finally, tobacco addiction is also highly prevalent in pathological gambling (McGrath and Barrett, 2009), with a common implication of altered probabilistic discounting and dopaminergic alterations. This topic would deserve a complete review, but clearly further animal studies are needed to dissect the neurobiological mechanisms by which nicotine and mental illnesses interact to affect decision-making.

### Acknowledgments

This work was supported by the Centre National de la Recherche Scientifique and ATIP programme, the Agence Nationale pour la Recherche (ANR Neuroscience, Neurologie et Psychiatrie 2009, and ANR BLANC 2010 and 2012), the Neuropole de Recherche Francilien (NeRF) of Ile de France, the Bettencourt Schueller Foundation (prix Coup d'élan, 2009), Ecole des Neurosciences de Paris (ENP), the Fondation pour la recherche Médicale (Equipe FRM DEQ20130326488) and the Fondation pour le cerveau (FRC et les Rotariens de France, "Espoir en tête" 2012). Laboratories of P.F. are part of the École des Neurosciences de Paris Ile-de-France RTRA

network and P.F. is member of the Laboratory of Excellence, LabEx Bio-Psy.

### References

- American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed. Washington, DC: Author.
- Addicott, M.A., Pearson, J.M., Wilson, J., Platt, M.L., McClernon, F.J., 2013. Smoking and the bandit: a preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Exp. Clin. Psychopharmacol.* 21, 66–73.
- Anderson, K.G., Diller, J.W., 2010. Effects of acute and repeated nicotine administration on delay discounting in Lewis and Fisher 344 rats. *Behav. Pharmacol.* 21, 754–764.
- Anokhin, A.P., Golosheykin, S., Grant, J.D., Heath, A.C., 2011. Heritability of delay discounting in adolescence: a longitudinal twin study. *Behav. Genet.* 41, 175–183.
- Audrain-McGovern, J., Rodriguez, D., Epstein, L.H., Cuevas, J., Rodgers, K., Wileyto, E.P., 2009. Does delay discounting play an etiological role in smoking or is it a consequence of smoking? *Drug Alcohol Depend.* 103, 99–106.
- Avale, M.E., Faure, P., Pons, S., Robledo, P., Deltheil, T., David, D.J., Gardier, A.M., Maldonado, R., Granon, S., Changeux, J.-P., Maskos, U., 2008. Interplay of  $\beta 2^*$  nicotinic receptors and dopamine pathways in the control of spontaneous locomotion. *Proc. Natl. Acad. Sci. U. S. A.* 105, 15991–15996.
- Bailey, C.D.C., De Biasi, M., Fletcher, P.J., Lambe, E.K., 2010. The nicotinic acetylcholine receptor  $\alpha 5$  subunit plays a key role in attention circuitry and accuracy. *J. Neurosci.* 30, 9241–9252.
- Baker, F., Johnson, M.W., Bickel, W.K., 2003. Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. *J. Abnorm. Psychol.* 112, 382–392.
- Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.
- Barazangi, N., Role, L.W., 2001. Nicotine-induced enhancement of glutamatergic and GABAergic synaptic transmission in the mouse amygdala. *J. Neurophysiol.* 86, 463–474.
- Beeler, J. a, Daw, N., Frazier, C.R.M., Zhuang, X., 2010. Tonic dopamine modulates exploitation of reward learning. *Front. Behav. Neurosci.* 4, 170.
- Belin, D., Mar, A.C., Dalley, J.W., Robbins, T.W., Everitt, B.J., 2008. Europe PMC Funders group high impulsivity predicts the switch to compulsive cocaine taking. *Neurosci.* 320, 1352–1355.
- Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 191, 391–431.
- Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.* 28, 309–369.
- Besson, M., Granon, S., Mameli-Engvall, M., Cloëz-Tayarani, I., Maubourguet, N., Cormier, A., Cazala, P., David, V., Changeux, J.-P., Faure, P., 2007. Long-term effects of chronic nicotine exposure on brain nicotinic receptors. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8155–8160.
- Bickel, W.K., Odum, L., Madden, G.J., 1999. Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology (Berl)* 146, 447–454.
- Boureau, Y.-L., Dayan, P., 2010. Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36, 74–97.
- Bromberg-Martin, E.S., Matsumoto, M., Hikosaka, O., 2010. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834.
- Buisson, B., Bertrand, D., 2002. Nicotine addiction: the possible role of function upregulation. *Trends Pharmacol. Sci.* 23, 130–136.
- Businelle, M., McVay, M.A., Kendzor, D., Copeland, A., 2010. A comparison of delay discounting among smokers, substance users, and non-dependent controls. *Drug Alcohol Depend.* 112, 247–250.
- Cai, X., Kim, S., Lee, D., 2011. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* 69, 170–182.
- Caillé, S., Guillem, K., Cador, M., Manzoni, O., Georges, F., 2009. Voluntary nicotine consumption triggers in vivo potentiation of cortical excitatory drives to midbrain dopaminergic neurons. *J. Neurosci.* 29, 10410–10415.
- Changeux, J.-P., Edelman, S.J., 2005. Allosteric mechanisms of signal transduction. *Science* 308, 1424–1428.
- Chiu, P.H., Lohrenz, T.M., Montague, P.R., 2008. Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nat. Neurosci.* 11, 514–520.
- Clark, L., Cools, R., Robbins, T.W., 2004. The neuropsychology of ventral prefrontal cortex: decision-making and reversal learning. *Brain Cognit.* 55, 41–53.
- Clement, J.R., Grant, S., 1990. Glutamate-like immunoreactivity in neurons of the laterodorsal and pedunculopontine nucleus in the rat. *Neurosci. Lett.* 120, 70–73.
- Cornwall, J., Cooper, J.D., Phillipson, O.T., 1990. *Exp. Brain Res.* 9, 157–171.
- Dalley, J.W., Everitt, B.J., Robbins, T.W., 2011. Impulsivity, compulsivity, and top-down cognitive control. *Neuron* 69, 680–694.
- Dani, J.A., Ji, D., Zhou, F., 2001. Synaptic plasticity and nicotine addiction. *Neuron* 31, 349–352.

- Dani, J.A., Harris, R.A., 2005. Nicotine addiction and comorbidity with alcohol abuse and mental illness. *Nat. Neurosci.* 8, 1465–1470.
- Dani, J.A., Bertrand, D., 2007. Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annu. Rev. Pharmacol. Toxicol.* 47, 699–729.
- Dani, J.A., Jenson, D., Broussard, J.L., Biasi, M. De, 2011. Neurophysiology of nicotine addiction. *J. Addict. Res. Ther.* 1, 1–12.
- Dautan, D., Huerta-Ocampo, I., Witten, I.B., Deisseroth, K., Bolam, J.P., Gerdjikov, T., Mena-Segovia, J., 2014. A major external source of cholinergic innervation of the striatum and nucleus accumbens originates in the brainstem. *J. Neurosci.* 34, 4509–4518.
- Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.
- Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69 (6), 1204–1215.
- de Wit, H., Richards, J.B., 2004. Dual determinants of drug use in humans: reward and impulsivity. *Nebr. Symp. Motiv.* 50, 19–55.
- de Wit, H., 2009. Impulsivity as a determinant and consequence of drug use: a review of underlying processes. *Addict. Biol.* 14, 22–31.
- Diergaarde, L., Pattij, T., Poortvliet, I., Hogenboom, F., Vries, W. De, Schoffeleers, A.N.M., Vries, T.J. De, 2008. Impulsive choice and impulsive action predict vulnerability to distinct stages of nicotine seeking in rats. *Biol. Psychiatry* 63, 301–308.
- Dallery, J., Locy, M.L., 2005. Effects of acute and chronic nicotine on impulsive choice in rats. *Behav. Pharmacol.* 16 (1), 15–23.
- Devaine, M., Hollard, G., Daunizeau, J., 2014. Theory of mind: did evolution fool us? *PLoS ONE* 9, e87619.
- Dolan, R.J., Dayan, P., 2013. Goals and habits in the brain. *Neuron* 80, 312–325.
- Doll, B.B., Simon, D.A., Daw, N.D., 2012. The ubiquity of model-based reinforcement learning. *Curr. Opin. Neurobiol.* 22, 1075–1081.
- Dos Santos Coura, R., Granon, S., 2012. Prefrontal neuromodulation by nicotinic receptors for cognitive processes. *Psychopharmacology (Berl)* 221, 1–18.
- Doya, K., 2008. Modulators of decision making. *Nat. Neurosci.* 11, 410–416.
- Doyon, W.M., Dong, Y., Ostroumov, A., Thomas, A.M., Zhang, T.A., Dani, J.A., 2013. Nicotine decreases ethanol-induced dopamine signaling and increases self-administration via stress hormones. *Neuron* 1–11.
- Epping-Jordan, M.P., Watkins, S.S., Koob, G.F., Markou, A., 1998. Dramatic decreases in brain reward function during nicotine withdrawal. *Nature* 393, 76–79.
- Ersche, K.D., Turton, A.J., Pradhan, S., Bullmore, E.T., Robbins, T.W., 2010. Drug addiction endophenotypes: impulsive versus sensation-seeking personality traits. *Biol. Psychiatry* 68, 770–773.
- Evenden, J.L., 1999. Varieties of impulsivity. *Psychopharmacology (Berl)* 146, 348–361.
- Exley, R., Cragg, S.J., 2008. Presynaptic nicotinic receptors: a dynamic and diverse cholinergic filter of striatal dopamine neurotransmission. *Br. J. Pharmacol.* 153 (Suppl. 1), S283–S297.
- Faure, P., Tolu, S., Valverde, S., Naudé, J., 2014. Role of nicotinic acetylcholine receptors in regulating dopamine neuron activity. *Neuroscience* 282, 86–100.
- Fobbs, W.C., Mizumori, S.J.Y., 2014. Cost-benefit decision circuitry: proposed modulatory role for acetylcholine. *Prog. Mol. Biol. Transl. Sci.* 122, 233–261.
- Fowler, C.D., Arends, M.A., Kenny, P.J., 2008. Subtypes of nicotinic acetylcholine receptors in nicotine reward, dependence, and withdrawal: evidence from genetically modified mice. *Behav. Pharmacol.* 19, 461–484.
- Fowler, C.D., Kenny, P.J., 2014. Nicotine aversion: neurobiological mechanisms and relevance to tobacco dependence vulnerability. *Neuropharmacology* 76 (Pt B), 533–544.
- Gallego, X., Molas, S., Amador-Arjona, A., Marks, M.J., Robles, N., Murtra, P., Armengol, L., Fernández-Montes, R.D., Gratacós, M., Pumarola, M., Cabrera, R., Maldonado, R., Sabrià, J., Estivill, X., Dierssen, M., 2012. Overexpression of the CHRNA5/A3/B4 genomic cluster in mice increases the sensitivity to nicotine and modifies its reinforcing effects. *Amino Acids* 43, 897–909.
- Govind, A.P., Vezina, P., Green, W.N., 2009. Nicotine-induced upregulation of nicotinic receptors: underlying mechanisms and relevance to nicotine addiction. *Biochem. Pharmacol.* 78, 756–765.
- Grace, A. a., Floresco, S.B., Goto, Y., Lodge, D.J., 2007. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* 30, 220–227.
- Guillem, K., Bloem, B., Poorthuis, R.B., Loos, M., Smit, A.B., Maskos, U., Spijker, S., Mansvelder, H.D., 2011. Nicotinic acetylcholine receptor  $\beta 2$  subunits in the medial prefrontal cortex control attention. *Science* 333, 888–891.
- Hasselmo, M.E., Sarter, M., 2011. Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology* 36, 52–73.
- Hauser, T.A., Kucinski, A., Jordan, K.G., Gatto, G.J., Wersinger, S.R., Hesse, R.A., Stachowiak, E.K., Stachowiak, M., Papke, R.L., Lippello, P.M., Bencherif, M., 2010. TC-5619: an  $\alpha 7$  neuronal nicotinic receptor-selective agonist that demonstrates efficacy in animal models of the positive and negative symptoms and cognitive dysfunction of schizophrenia. *Biochem. Pharmacol.* 78, 803–812.
- Hay, Y.A., Andjelic, S., Badr, S., Lambolze, B., 2014. Orexin-dependent activation of layer Vb enhances cortical network activity and integration of non-specific thalamocortical inputs. *Brain Struct. Funct.* <http://dx.doi.org/10.1007/s00429-014-0869-7>.
- Heilbronner, S.R., Hayden, B.Y., 2013. Contextual factors explain risk-seeking preferences in rhesus monkeys. *Front. Neurosci.* 7, 7.
- Herrnstein, R.J., 1961. Relative and absolute strength of responses as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.
- Hikosaka, O., Nakamura, K., Nakahara, H., 2006. Basal ganglia orient eyes to reward. *J. Neurophysiol.* 95 (2), 567–584.
- Hong, S., Jhou, T.C., Smith, M., Saleem, K.S., Hikosaka, O., 2011. Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J. Neurosci.* 31, 11457–11471.
- Humphries, M.D., Khamassi, M., Gurney, K., 2012. Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Front. Neurosci.* 6, 9.
- Humphries, M.D., Prescott, T.J., 2010. The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Prog. Neurobiol.* 90, 385–417.
- Ikemoto, S., Panksepp, J., 1999. The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Brain Res. Rev.* 31, 6–41.
- Joel, D., Weiner, I., 2000. The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *NSC* 96, 451–474.
- Johnson, M.W., Bickel, W.K., Baker, F., 2007. Moderate drug use and delay discounting: a comparison of heavy, light, and never smokers. *Exp. Clin. Psychopharmacol.* 15 (2), 187–194.
- Jones, S., Bonci, A., 2005. Synaptic plasticity and drug addiction. *Curr. Opin. Pharmacol.* 5, 20–25.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47 (2), 263–291.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.
- Kable, J.W., Glimcher, P.W., 2009. The neurobiology of decision: consensus and controversy. *Neuron* 63, 733–745.
- Kelsey, J.E., Niraula, A., 2013. Effects of acute and sub-chronic nicotine on impulsive choice in rats in a probabilistic delay-discounting task. *Psychopharmacology (Berl)* 227, 385–392.
- Keramati, M., Dezfouli, A., Piray, P., 2011. Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7, e1002055.
- Keramati, M., Gutkin, B., 2013. Imbalanced decision hierarchy in addicts emerging from drug-hijacked dopamine spiraling circuit. *PLoS ONE* 8, e61489.
- Kim, S., Lee, D., Jun, J., 2011. Prefrontal cortex and impulsive decision making. *Biol. Psychiatry* 69 (12), 1140–1146.
- Kim, H., Lee, D., Jung, M.W., 2013. Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. *J. Neurosci.* 33 (1), 52–63.
- Kishida, K.T., King-Casas, B., Montague, P.R., 2010. Neuroeconomic approaches to mental disorders. *Neuron* 67, 543–554.
- Klink, R., de Kerchove d'Exaerde, A., Zoli, M., Changeux, J.-P., 2001. Molecular and physiological diversity of nicotinic acetylcholine receptors in the midbrain dopaminergic nuclei. *J. Neurosci.* 21, 1452–1463.
- Kolokotroni, K.Z., Rodgers, R.J., Harrison, A.A., 2011. Acute nicotine increases both impulsive choice and behavioural disinhibition in rats. *Psychopharmacology* 217 (4), 455–473.
- Koranda, J.L., Cone, J.J., McGehee, D.S., Roitman, M.F., Beeler, J. a., Zhuang, X., 2014. Nicotinic receptors regulate the dynamic range of dopamine release in vivo. *J. Neurophysiol.* 111, 103–111.
- Lammel, S., Ion, D.L., Roeper, J., Malenka, R.C., 2011. Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron* 70, 855–862.
- Lamotte, B., Ascher, P., 2014. High affinity and low affinity heteromeric nicotinic acetylcholine receptors at central synapses. *J. Physiol.* 819, 1–10.
- Lau, B., Glimcher, P.W., 2008. Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463.
- Lejuez, C.W., Aklin, W.M., Zvolensky, M.J., Pedulla, C.M., 2003. Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *J. Adolesc.* 26, 475–479.
- Lo, C.C., Wang, X.J., 2006. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *J. Neurosci.* 26, 956–963.
- Locy, M.L., Dallery, J., 2009. Isolating behavioral mechanisms of intertemporal choice: nicotine effects on delay discounting and amount sensitivity. *J. Exp. Anal. Behav.* 91, 213–223.
- Locy, M.L., Dallery, J., 2012. Nicotine and the behavioral mechanisms of intertemporal choice. *Behav. Process.* 87, 18–24.
- Lohrenz, T., McCabe, K., Camerer, C.F., Montague, P.R., 2007. Neural signature of fictive learning signals in a sequential investment task. *Proc. Natl. Acad. Sci. U. S. A.* 104, 9493–9498.
- McGrath, D.S., Barrett, S.P., 2009. The comorbidity of tobacco smoking and gambling: a review of the literature. *Drug Alcohol Rev.* 28, 676–681.
- Mameli-Engvall, M., Evrard, A., Pons, S., Maskos, U., Svensson, T.H., Changeux, J.-P., Faure, P., 2006. Hierarchical control of dopamine neuron-firing patterns by nicotinic receptors. *Neuron* 50, 911–921.
- Mansvelder, H.D., McGehee, D.S., 2000. Long-term potentiation of excitatory inputs to brain reward areas by nicotine. *Neuron* 27, 349–357.
- Mao, D., Perry, D.C., Yasuda, R.P., Wolfe, B.B., Kellar, K.J., 2008. The  $\alpha 4\beta 2\alpha 5$  nicotinic cholinergic receptor in rat brain is resistant to up-regulation by nicotine in vivo. *J. Neurochem.* 104, 446–456.
- Marks, M.J., Pauly, J.F.L., Heinemann, F., Collins, A.C., Deneris, E.S., Hermansborgmeyer, I., Gross, D., 1992. Nicotine binding and nicotinic nicotine receptor subunit RNA after chronic. *J. Neurosci.* 12, 2765–2784.



- Martí Barros, D., Ramirez, M.R., Dos Reis, E. a, Izquierdo, I., 2004. Participation of hippocampal nicotinic receptors in acquisition, consolidation and retrieval of memory for one trial inhibitory avoidance in rats. *Neuroscience* 126, 651–656.
- Martí, F., Arib, O., Morel, C., Dufresne, V., Maskos, U., Corringier, P.-J., de Beaurepaire, R., Faure, P., 2011. Smoke extracts and nicotine, but not tobacco extracts, potentiate firing and burst activity of ventral tegmental area dopaminergic neurons in mice. *Neuropsychopharmacology* 36, 2244–2257.
- Maskos, U., 2010. Role of endogenous acetylcholine in the control of the dopaminergic system via nicotinic receptors. *J. Neurochem.* 114, 641–646.
- Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841.
- Matta, S.G., Balfour, D.J., Benowitz, N.L., Boyd, R.T., Buccafusco, J.J., Caggiola, A.R., et al. Zirger, J.M., 2007. Guidelines on nicotine dose selection for in vivo research. *Psychopharmacology* 190, 269–319.
- Mendez, I.A., Gilbert, R.J., Bizon, J.L., Setlow, B., 2012. Effects of acute administration of nicotinic and muscarinic cholinergic agonists and antagonists on performance in different cost-benefit decision making tasks in rats. *Psychopharmacology (Berl)* 224, 489–499.
- Mitchell, M.R., Vokes, C.M., Blankenship, A.L., Simon, N.W., Setlow, B., 2011. Effects of acute administration of nicotine, amphetamine, diazepam, morphine, and ethanol on risky decision-making in rats. *Psychopharmacology (Berl)* 218, 703–712.
- Mitchell, S.H., 1999. Measures of impulsivity in cigarette smokers and non-smokers. *Psychopharmacology (Berl)* 146, 455–464.
- Moeller, F.G.M.D., Dougherty, D.M., 2002. Impulsivity and substance abuse: what is the connection? *Addict. Disord. Their Treat.* 1 (1), 3–10.
- Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. *Trends Cognit. Sci.* 16, 72–80.
- Morel, C., Fattore, L., Pons, S., Hay, Y.A., Martí, F., Lambomez, B., De Biasi, M., Lathrop, M., Fratta, W., Maskos, U., Faure, P., 2014. Nicotine consumption is regulated by a human polymorphism in dopamine neurons. *Mol. Psychiatry* 19, 930–936.
- Nashmi, R., Xiao, C., Deshpande, P., McKinney, S., Grady, S.R., Whiteaker, P., Huang, Q., McClure-Begley, T., Lindstrom, J.M., Labarca, C., Collins, A.C., Marks, M.J., Lester, H.A., 2007. Chronic nicotine cell specifically upregulates functional  $\alpha 4^*$  nicotinic receptors: basis for both tolerance in midbrain and enhanced long-term potentiation in perforant path. *J. Neurosci.* 27, 8202–8218.
- Nesic, J., Rusted, J., Duka, T., Jackson, a., 2011. Degree of dependence influences the effect of smoking on cognitive flexibility. *Pharmacol. Biochem. Behav.* 98, 376–384.
- Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191, 507–520.
- Niv, Y., Edlund, J. a, Dayan, P., O'Doherty, J.P., 2012. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J. Neurosci.* 32, 551–562.
- Norton, A.B.W., Jo, Y.S., Clark, E.W., Taylor, C.A., Mizumori, S.J.Y., 2011. Independent neural coding of reward and movement by pedunculopontine tegmental nucleus neurons in freely navigating rats. *Eur. J. Neurosci.* 33, 1885–1896.
- Oakman, S.A., Faris, P.L., Kerr, P.E., Cozzari, C., Hartman, B.K., 1995. Distribution of pontomesencephalic cholinergic neurons projecting to substantia nigra differs significantly from those projecting to ventral tegmental area. *J. Neurosci.* 15, 5859–5869.
- Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.
- Pan, W.-X., Hyland, B.I., 2005. Pedunculopontine tegmental nucleus controls conditioned responses of midbrain dopamine neurons in behaving rats. *J. Neurosci.* 25, 4725–4732.
- Parikh, V., Kozak, R., Martinez, V., Sarter, M., 2007. Prefrontal acetylcholine release controls cue detection on multiple timescales. *Neuron* 56, 141–154.
- Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87 (6), 532.
- Perry, J.L., Carroll, M.E., 2008. The role of impulsive behavior in drug abuse. *Psychopharmacology (Berl)* 200, 1–26.
- Perry, J.L., Larson, E.B., German, J.P., Madden, G.J., Carroll, M.E., 2005. Impulsivity (delay discounting) as a predictor of acquisition of IV cocaine self-administration in female rats. *Psychopharmacology (Berl)* 178, 193–201.
- Peters, J., Büchel, C., 2011. The neural mechanisms of inter-temporal decision-making: understanding variability. *Trends Cognit. Sci.* 15, 227–239.
- Phillips, P.E., Stuber, G.D., Heien, M.L., Wightman, R.M., Carelli, R.M., 2003. Sub-second dopamine release promotes cocaine seeking. *Nature* 422 (6932), 614–618.
- Picciotto, M.R., 2003. Nicotine as a modulator of behavior: beyond the inverted U. *Trends Pharmacol. Sci.* 24, 493–499.
- Pichat, P., Bergis, O.E., Terranova, J.-P., Urani, A., Duarte, C., Santucci, V., Guedet, C., Voltz, C., Steinberg, R., Stemmelin, J., Oury-Donat, F., Avenet, P., Griebel, G., Scatton, B., 2007. SSR180711, a novel selective  $\alpha 7$  nicotinic receptor partial agonist: (II) efficacy in experimental models predictive of activity against cognitive symptoms of schizophrenia. *Neuropsychopharmacology* 32, 17–34.
- Pidoplichko, V.I., DeBiasi, M., Williams, J.T., Dani, J.A., 1997. Nicotine activates and desensitizes midbrain dopamine neurons. *Nature* 390, 401–404.
- Prickaerts, J., van Goethem, N.P., Chesworth, R., Shapiro, G., Boess, F.G., Methfessel, C., Reneerkens, O. a H., Flood, D.G., Hilt, D., Gawryl, M., Bertrand, S., Bertrand, D., König, G., 2012. EVP-6124, a novel and selective  $\alpha 7$  nicotinic acetylcholine receptor partial agonist, improves memory performance by potentiating the acetylcholine response of  $\alpha 7$  nicotinic acetylcholine receptors. *Neuropharmacology* 62, 1099–1110.
- Rahman, S., Zhang, J., Engleman, E.A., Corrigan, W.A., 2004. Neuroadaptive changes in the mesoaccumbens dopamine system after chronic nicotine self-administration: a microdialysis study. *Neuroscience* 129, 415–424.
- Ramirez-Latorre, J., Yu, C.R., Qu, X., Perin, F., Karlin, A., Role, L., 1996. Functional contributions of  $\alpha 5$  subunit to neuronal acetylcholine receptor channels. *Nature* 380, 347–351.
- Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-oro, M.C., 2011. Europe PMC Funders Group Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *11*, 760–772.
- Redish, A.D., 2004. Addiction as a computational process gone awry. *Science* 306, 1944–1947.
- Redish, A.D., Jensen, S., Johnson, A., Kurth-Nelson, Z., 2007. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychol. Rev.* 114 (3), 784–805.
- Redish, A.D., Jensen, S., Johnson, A., 2008. A unified framework for addiction: vulnerabilities in the decision process. *Behav. Brain Sci.* 31, 415–437. Discussion 437–87.
- Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. *Class. Cond. II Curr. Res. Theory* 2, 64–99.
- Reynolds, J.N., Hyland, B.I., Wickens, J.R., 2001. A cellular mechanism of reward-related learning. *Nature* 413 (6851), 67–70.
- Reynolds, B., Richards, J.B., Horn, K., Karraker, K., 2004. Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behav. Process.* 65, 35–42.
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron* 70 (6), 1054–1069.
- Saal, D., Dong, Y., Bonci, A., Malenka, R.C., 2003. Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neurons* 37 (4), 577–582.
- Salamone, J.D., Correa, M., Farrar, A.M., Nunes, E.J., Pardo, M., 2009. Dopamine, behavioral economics, and effort. *Front. Behav. Neurosci.* 3, 13.
- Samejima, K., Ueda, Y., Doya, K., Kimura, M., 2005. Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.
- Sarter, M., Parikh, V., Howe, W.M., 2009. Phasic acetylcholine release and the volume transmission hypothesis: time to move on. *Nat. Rev. Neurosci.* 10, 383–390.
- Schoemaker, P.J.H., 1982. The expected utility model: its variants, purposes, evidence and limitations. *J. Econ. Lit.* XX, 529–563.
- Schultz, W., 2002. Getting formal with dopamine and reward. *Neuron* 36, 241–263.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Serreau, P., Chabout, J., Suarez, S.V., Naudé, J., Granon, S., 2011.  $\beta 2$ -containing neuronal nicotinic receptors as major actors in the flexible choice between conflicting motivations. *Behav. Brain Res.* 225, 151–159.
- Simon, D.A., Daw, N.D., 2011. Neural correlates of forward planning in a spatial decision task in humans. *J. Neurosci.* 31, 5526–5539.
- Skjei, K.L., Markou, A., 2003. Effects of repeated withdrawal episodes, nicotine dose, and duration of nicotine exposure on the severity and duration of nicotine withdrawal in rats. *Psychopharmacology* 168, 280–292.
- Stocco, A., 2012. Acetylcholine-based entropy in response selection: a model of how striatal interneurons modulate exploration, exploitation, and response variability in decision-making. *Front. Neurosci.* 6, 18.
- Stolerman, I.P., Jarvis, M.J., 1995. The scientific case that nicotine is addictive. *Psychopharmacology (Berl)* 117, 2–10. Discussion 14–20.
- Sutton, R.S., Barto, A.G., 1998. Introduction to Reinforcement Learning. MIT Press.
- Tang, J., Dani, J.A., 2009. Dopamine enables in vivo synaptic plasticity associated with the addictive drug nicotine. *Neuron* 63, 673–682.
- Tolu, S., Eddine, R., Martí, F., David, V., Graupner, M., Pons, S., Baudonnet, M., Husson, M., Besson, M., Reperant, C., Zemdegs, J., Pagès, C., Hay, Y.A.H., Lambomez, B., Caboche, J., Gutkin, B., Gardier, A.M., Changeux, J.-P., Faure, P., Maskos, U., 2013. Co-activation of VTA DA and GABA neurons mediates nicotine reinforcement. *Mol. Psychiatry* 18, 382–393.
- Thompson, J.A., Felsen, G., 2013. Activity in mouse pedunculopontine tegmental nucleus reflects action and outcome in a decision-making task. *J. Neurophysiol.* 110, 2817–2829.
- Valjent, E., Pagès, C., Hervé, D., Girault, J.-A., Caboche, J., 2004. Addictive and non-addictive drugs induce distinct and specific patterns of ERK activation in mouse brain. *Eur. J. Neurosci.* 19, 1826–1836.
- Verdejo-García, A., Lawrence, A.J., Clark, L., 2008. Impulsivity as a vulnerability marker for substance-use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. *Neurosci. Biobehav. Rev.* 32, 777–810.
- Vinalls, X., Molas, S., Gallego, X., Fernández-Montes, R.D., Robledo, P., Dierssen, M., Maldonado, R., 2012. Overexpression of  $\alpha 3/\alpha 5/\beta 4$  nicotinic receptor subunits modifies impulsive-like behavior. *Drug Alcohol Depend.* 122 (3), 247–252.
- Vlaev, I., Chater, N., Stewart, N., Brown, G.D.A., 2011. Does the brain calculate value? *Trends Cognit. Sci.* 15, 546–554.
- Wallace, T.L., Porter, R.H.P., 2011. Targeting the nicotinic  $\alpha 7$  acetylcholine receptor to enhance cognition in disease. *Biochem. Pharmacol.* 82, 891–903.
- West, R., Brown, J., 2013. Theory of Addiction. John Wiley & Sons.
- Wilson, D.I.G., MacLaren, D.A.A., Winn, P., 2009. Bar pressing for food: differential consequences of lesions to the anterior versus posterior pedunculopontine. *Eur. J. Neurosci.* 30, 504–513.

- Winn, P., 2006. How best to consider the structure and function of the pedunculo-pontine tegmental nucleus: evidence from animal studies. *J. Neurol. Sci.* 248 (1), 234–250.
- Wonnacott, S., 1990. The paradox of nicotinic acetylcholine receptor upregulation by nicotine. *Trends Pharmacol. Sci.* 11.
- Xiao, C., Yang, K., Zhou, C., Jin, G., Wu, J., Ye, J., 2009. Nicotine modulates GABAergic transmission to dopaminergic neurons in substantia nigra pars compacta. *Acta Pharmacol. Sin.* 30, 851–858.
- Xiao, L., Bechara, A., Cen, S., Grenard, J.L., Stacy, A.W., Gallaher, P., Wei, Y., Jia, Y., Johnson, C.A., 2008. Affective decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in 10th-grade Chinese adolescent smokers. *Nicot. Tob. Res.* 10, 1085–1097.
- Yin, H.H., Ostlund, S.B., Balleine, B.W., 2008. Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* 28, 1437–1448.
- Yu, A.J., Dayan, P., 2002. Acetylcholine in cortical inference. *Neural Netw.* 15, 719–730.
- Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.
- Zhang, H., Sulzer, D., 2004. Frequency-dependent modulation of dopamine release by nicotine. *Nat. Neurosci.* 7, 581–582.
- Zhang, L., Dong, Y., Doyon, W.M., Dani, J.A., 2013. Withdrawal from chronic nicotine exposure alters dopamine signaling dynamics in the nucleus accumbens. *Biol. Psychiatry* 71, 184–191.



## 5 | Objectives

As discussed in the previous review the role of nicotinic acetylcholinergic receptors in relation to decision-making is a not much explored research field and could reveal insight into the consequences of nicotine use on daily life behaviours that involve such choice processes. Furthermore, different mental disorders have been brought into relation with excessive smoking. Disorders that affect different higher cognitive processes. These difficulties may partly be relieved or counteracted through smoking, while longterm smoking in non-affected individuals has repetitively solely been correlated with negative cognitive consequences. These findings may indicate that nicotinic acetylcholinergic receptors and its endogenous neurotransmitter acetylcholine can be bidirectionally linked to different cognitive consequences.

During this thesis, I developed approaches to investigate the role of nicotine receptors and the effects of chronic exposure on decision-making that are independent of nicotine seeking behavior. A few attempts have been made to identify the effect of long term nicotine use, and therefore the nicotinic system, on decision-making in humans. However, these studies express a few complications. Firstly, smokers are automatically addicted to nicotine and addiction; the habitual use of drugs or other compulsive behaviours, evokes brain plasticity in itself. Furthermore, a predisposition for addiction has been correlated with different personality and genetic traits that may also explain the findings. Secondly, the tasks used to assess choice behaviour are complex and complicate identification of modified parameters. For example, Addicott's "restless" bandit task [238] can not distinguish between rigidity or reward sensitivity. Various paradigms have been developed for rodents based on the Iowa Gambling task [239] and probabilistic reversal learning tasks [240] and these tasks are probably the most suitable to assess behaviour in a probabilistic environment. No data yet has been published on the effects of long-term nicotine in these tasks. However, these tasks have a few limitations: they are relatively difficult for rodents to learn due to their direct translation of human tasks; and usually require food restriction which may affect the outcome.

For this reason, we designed a behavioural paradigm for mice (*See Methods and Results*

that allows to assess different decision-making parameters in a fixed probabilistic setting, with as reward an electrical intracranial stimulation. This paradigm allows to develop computational models of the task that serve as a powerful tool to identify scales of behavioural modifications. With this multi-armed bandit task we can focus on two specific parameters: uncertainty and reward sensitivity, that can be compared between different experimental groups. The definitions of these parameters will be further discussed in the result section. For the first paper, we aimed at describing the paradigm and studying the role of nicotinic cholinergic system in decision-making through the inactivation and re-expression of the high-affinity  $\beta 2$  subunit in the VTA. For the second paper we analyse the effect of chronic nicotine exposure within this decision-making framework. To learn about the exclusive consequences of nicotine on these parameters without taking the effects of addiction into account we treated mice passively with nicotine and studied the behavioural modifications in the same paradigm again with a focus on the interaction between the dopaminergic reward system and the nicotinic cholinergic system. Since, also the presented multi-armed bandit task may have limitations I will discuss preliminary data obtained from the experiments in a closed economy for mice. These data will be presented in the discussion, with a brief introduction in the methods section. The closed economy for mice provides a more natural setting to study behaviour in a probabilistic environment.

# Methods



# 6 | Methods and materials

## 6.1 Behavioural paradigms

### 6.1.1 Multi-armed ICSS Bandit task

We have developed a task based on intra-cranial self-stimulation (ICSS) rewards to assess Bandit-style problems in mice. The ICSS electrode is implanted in the medial forebrain bundle (*see Figure 6.1a and Chapter 7 & 8*). Mice were trained to perform a sequence of choices between explicit locations in a open field (0.8 m diameter) to obtain ICSSs. Mice cannot receive two consecutive ICSSs at the same location. Consequently, they alternate between rewarding locations by performing a sequence of choices. The task consisted of daily 5-minute sessions in an open-field where three places were explicitly associated with ICSS. ICSS intensity is adjusted so that mice self-stimulate between 50 and 150 times per session at the end of the training (ninth and tenth session). Current intensity is subsequently maintained the same throughout all the experiments. Mice are initially trained in a deterministic/certain setting (i.e in which all locations are associated with a certain ICSS) and then in a probabilistic/uncertain setting (*see Figure 6.1b*), in which each location is associated with a different probability of ICSS delivery (in the first set, reward probabilities are 100, 50 and 25% for the three locations).

To quantify behaviours, sequences of choices (which rewarded site); trajectories and kinetics of the displacement in-between rewarded sites have been analysed. The repartition of the rewarded locations arises from a sequence of binary choices between two respective payoffs (e.g. 100% vs. 50% when the animal is on the 25% site). This type of sequential decision-making task constitutes a Markovian decision process consisting of 3 states (A, B, C) and a transition function expressing the proportion of choices made after leaving this point. The states correspond to the rewarded positions, and the transition function corresponds to the binary choices the mouse has to perform from these locations. These transitions can be expressed in three gambles and this data (i.e the transition function) has been fitted with reinforcement-learning models (RL). Three main models have been tested. These gambles can reflect the exploration-exploitation trade-off: considering



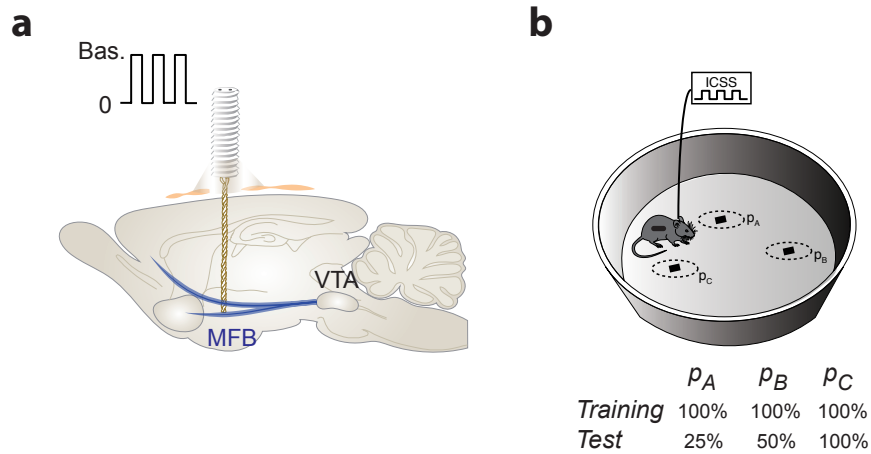


Figure 6.1 – **Multi-armed ICSS bandit task** **a** Unilateral implantation of a stimulation electrode into the medial forebrain bundle **b** A representation of the multi-armed ICSS bandit task

that you have to choose between two rewards  $A$  and  $B$  that are associated with a value  $V_A$  and  $V_B$ . If  $V_A > V_B$ , and you do not explore, but go for the optimal choice, you choose  $A$ . The three models describe the probability to choose  $A$  versus  $B$  and thus how exploration is driven by their values ( $V_A$  and  $V_B$ ).

- The  $\epsilon$ -greedy models. In this case exploration is completely independent of the value  $V_A > V_B$ . You choose  $B$  (and thus explore) with a fixed probability  $\epsilon$ .
- The classic softmax model. The probability to choose  $A$  depends on the difference between  $V_A$  and  $V_B$ . If  $V_A = V_B$ , the probability  $p_A$  to choose  $A$  equal 50% (same for  $p_B$ ). This model formalises that the higher the reward difference ( $V_A - V_B$ ), the higher the probability of choosing the best option  $V_A$ .
- The uncertainty model based on a classic softmax model with the addition of a bonus for exploration. The value associated with a choice is increasing by the uncertainty associated with this choice. This is a bonus. This model formalises that the decision is also biased in favour of choices with the most uncertain consequences by assigning a "bonus" value for the expected uncertainty.

The uncertainty model (third one) is the one that accounted best for the experimental data presented and reproduced the preferences of wild-type mice as explained in the first paper that will be discussed in the results section. This model has two parameters:  $\beta$  (inverse temperature), which represents the sensitivity to the difference of values (see above) and the parameter  $\phi$  (expected uncertainty) which represents the value attributed

to the expected uncertainty. If the WT sequence of choice are well fitted by the third model,  $\beta 2$  KO mice are best fitted by the second one.

### 6.1.2 Closed economy for mice

We use classical operant boxes (Med. Associates) for mice with two levers and a food well in between them (see Figure 6.2). The floor of the cage is adapted from a grid to a solid aluminum plate and covered in a layer of bedding. Water is provided ad libitum. Mice have access to the two non-retractable levers continuously. This set up allows us to study neuroeconomical phenomena driven by natural motivation (no food restriction; no time restrains). We focus on value discrimination by reversing the location of the reinforcing lever. At any point during the experiment one lever is reinforced with one food pellet at a certain fixed ratio (FR) and the other lever is inactive. The reinforcing lever changes position every 24 hours or 96 hours. Experiments take 5-7 days and during this period the mouse is solitary housed in the operant box. Nicotine was administered acutely (I.P. injections, 500 $\mu$ g/kg), or in drinking water (200  $\mu$ g/mL, 2% Saccharine). Under the control conditions, we recorded VTA DA cell activity of WT mice over periods of 24 hours and more with in vivo electrophysiology.

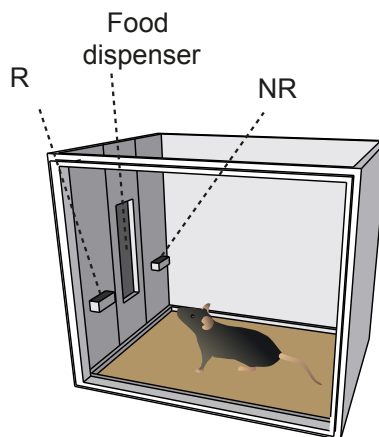


Figure 6.2 – **Closed economy paradigm** Operant box in which the mouse lives continuously for the complete time of the experiment. Food can be obtained through lever pressing. R = rewarding lever, NR = non-rewarding lever

## 6.2 Nicotine and minipump implantation

Implantation of osmotic mini pumps Animals were anaesthetised with a gas mixture of oxygen (1L/min) and 1-3% of isoflurane (IsoVet, Piramal Healthcare,UK). After the administration of a local anaesthetic, an incision was performed at the level of the interscapular zone, to subcutaneously implant an osmotic minipump (Model 2004, ALZET, CA, USA) containing 200  $\mu$ L of either a solution of nicotine hydrogen tartrate salt (Sigma-Aldrich, MO, USA) at a dose of 10mg/kg/d (4.16 mg/kg, free base) or saline water (H<sub>2</sub>O with 0.9% NaCl) for the control group. Both solutions were prepared in the laboratory. The minipumps delivered their content with a flow of 0.25  $\mu$ L/hour over 28 days. The surgical wound was closed with surgical stitches. Animals had two days of rest to recover from the minipump surgery before going further with their behavioural training.

## 6.3 Electrophysiological recordings of dopaminergic neurons in the VTA

### 6.3.1 Juxtacellular recordings

Mice were deeply anaesthetised with chloral hydrate (8%), 400 mg/kg I.P., supplemented as required to maintain optimal anaesthesia throughout the experiment. The scalp was opened and a hole was drilled in the skull above the location of the VTA. The saphenous vein was catheterised for intravenous administration of drugs. Extracellular recording electrodes were constructed from 1.5 mm O.D. / 1.17 mm I.D. borosilicate glass tubing (Harvard Apparatus) using a vertical electrode puller (Narishige). Under microscopic control, the tip was broken to obtain a diameter of approximately 1  $\mu$ m. The electrodes were filled with a 0.5% NaCl solution containing 1.5% of neurobiotin tracer (AbCys) yielding impedances of 6-9 M $\Omega$ . Electrical signals were amplified by a high-impedance amplifier (Axon Instruments) and monitored audibly through an audio monitor (A.M. Systems Inc.). The signal was digitised, sampled at 25 kHz and recorded on a computer using Spike2 software (Cambridge Electronic Design) for later analysis. The electrophysiological activity was sampled in the central region of the VTA (coordinates: between 3.1 to 4 mm posterior to Bregma, 0.3 to 0.7 mm lateral to midline, and 4 to 4.8 mm below brain surface). Individual electrode tracks were separated from one another by at least 0.1 mm

in the horizontal plane. Spontaneously active DAergic neurons were identified based on previously established electrophysiological criteria (see main text, methods section).

### 6.3.2 Development of the micro-system

Steve Didiene, a recent PhD student in the lab, has developed a micro-system that allows us to record multi-units at different depths in the ventral tegmental area. Two octrodes are connected to a "descender" (see *Figure 6.3*). The descender is a guide cannula connected to a screw that leads the electrodes deeper in the brain area of interest when turned. Every turn of the screw will descend the two octrodes with  $300\ \mu\text{m}$ . This descender system is connected to an electronic interface board (EIB, *Figure 6.3a*) by means of a 3D-printed piece that also protects the electrodes and simplifies the implantation. An octrode exists out of a twisted loop of 4 strings of nichrome wire ( $0.1\text{mm}$ ) that undergoes approximately 20-30 turns in the neurospinner. Subsequently the twisted electrodes are

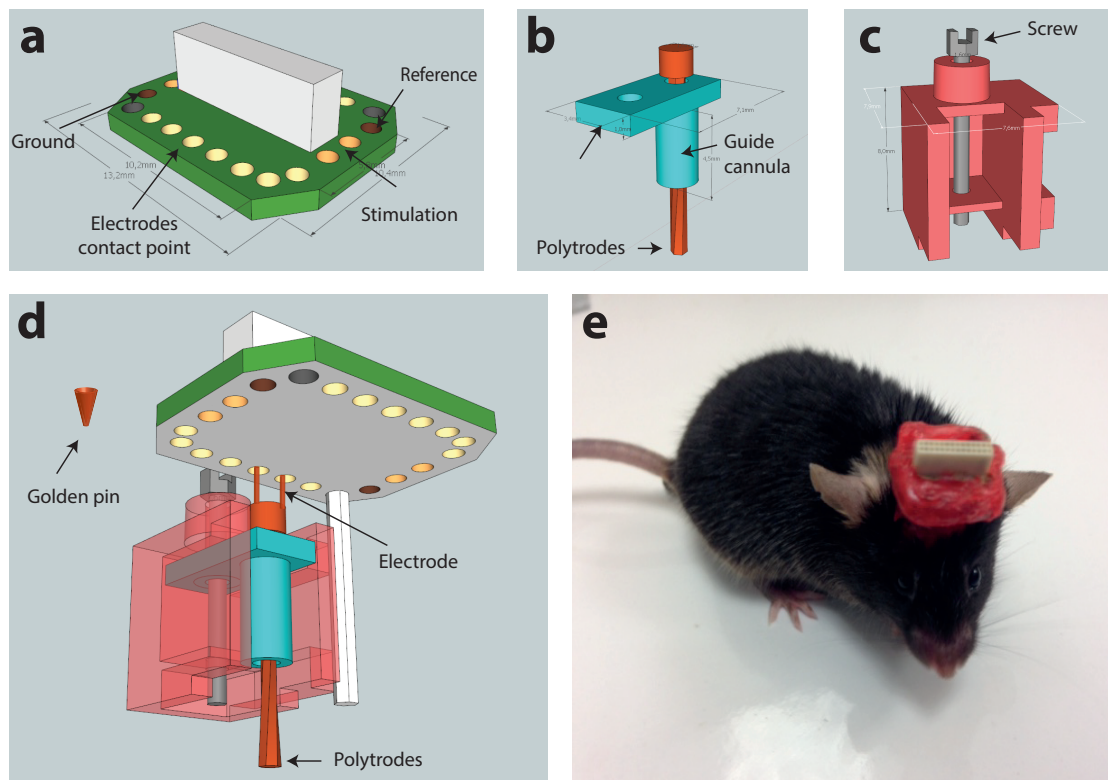


Figure 6.3 – Overview of the implantable electrophysiology micro-drive system a Electronic interface board (EIB) b Electrode cannula and bridge to the screw of the descender c Outside of descender with build-in screw d Complete microdrive e Implanted mouse

heated to melt the insulation, creating a stiff bundle of 8 microwires. Two octrodes are placed in the cannula guide. The ends of all 16 electrode wires are carefully put in the separate pin connectors of the EIB and fixed with golden pins. After completion (*Figure 6.3d*), the octrodes are glued together and cut to length of  $\pm 6$ mm measured from the lower end of the 3D-printed piece. Then the impedance of all separate electrodes is tested and subsequently unified and lowered (150-400 k $\Omega$ ) through an electroplating process with gold. The ends of the two octrodes are placed in a gold solution (gold in polyethylene glycol (PEG)), while NanoZ (Neuralynx software) sends constant-current pulses through the circuit. This leaves a gold residue attached to the ends of the octrodes, which increases the recording surface and the conductance of the octrodes. This process is also used to identify problems in the circuits. After electroplating, but before implantation, the descender and the EIB are wrapped into parafilm (except for the connector and octrodes) and covered with a layer of dental acrylic cement.

### 6.3.3 Surgery

After anaesthesia with a gas mixture of oxygen (1l/min) and 3% isoflurane, a mouse is placed in a stereotaxic frame (David Kopf). The cranial bone is exposed by a midline incision of the skin. The skull is then drilled and the recordings octrodes are placed just above the VTA (coordinates: 3.2  $\pm$  0.1 mm behind the Bregma, 0.5  $\pm$  0.1 mm lateral and 4.0  $\pm$  0.1 mm deep from the surface of the brain [241]). Monopolar ground electrodes are laid over the cortical layer of the cerebellum and the olfactory bulb. Dental acrylic cement is used to fix the main body of the micro-drive to the skull during the surgery. (*Figure 6.3e*)

### 6.3.4 Data acquisition and treatment

After a recovery of at least 2-3 weeks the extracellular potentials are tested in the homecage of the mouse. The EIB is connected to the digital acquisition system (Digital Lynx SX; Neuralynx). A cable with a preamplifier originating from the rotator is connected to the EIB. The rotator responds to the movements of the mice and reduces tension in the cable. The rotator has a data output towards the Neuralynx system, interpreted by the Cheetah Software. Broadband signals from each wire are filtered between 0.1 and 9000 Hz and recorded continuously at 32 kHz. To extract spike timing, signals are band-

pass-filtered between 600 and 6000 Hz. Once extracellular potentials are identified the mouse is placed in the closed economy for a period of 2-4 days depending on the protocol to get habituated to the environment and learn the instrumental behaviour. In general, the lever press-food reward association is established over night. After habituation, the extracellular potentials are recording for 24-72 hours. The data is saved in 3 hour files that are automatically initiated by trial control (Neuralynx software). Times of behavioural actions are aligned with the neurophysiological data through binary transistor-transistor logic (TTL) pulses. At the moment of a lever press or nose poke, the Med Associates system sends a TTL signal to the Neuralynx system. After a completed experiment, the octrodes can be lowered to test a different population of extracellular potentials.

### **6.3.5 Spike extraction and sorting**

Spike times were manually sorted offline with Spike Extractor (Neuralynx software). Sorted spikes were manually clustered with SpikeSort3D (Neuralynx software). SpikeSort3D offers different waveform parameters, like peak, valley, area and energy that can be included in the principal component analyses. Data clusters are represented in a 3D space (x, y, z axes) to simplify the cluster selection. Spike times are exported per cluster. Due to the size of the 3 hour files maximum 1-2 channels can be extracted and sorted simultaneously.

### **6.3.6 Dopamine neuron identification**

Extracellular identification of putative DA neurons was based on their location, as well as on the set of unique electrophysiological properties that characterise these cells in vivo: 1) a typical triphasic action potential with a marked negative deflection; 2) a characteristic long duration ( $>2.0$  ms) action potential; 3) an action potential width from start to negative trough  $>1.1$  ms; 4) a slow firing rate ( $<10$  Hz) with an irregular single spiking pattern and occasional short, slow bursting activity. After completion of the experiment, brain slices were prepared and the location of the recording electrodes were verified under the microscope.

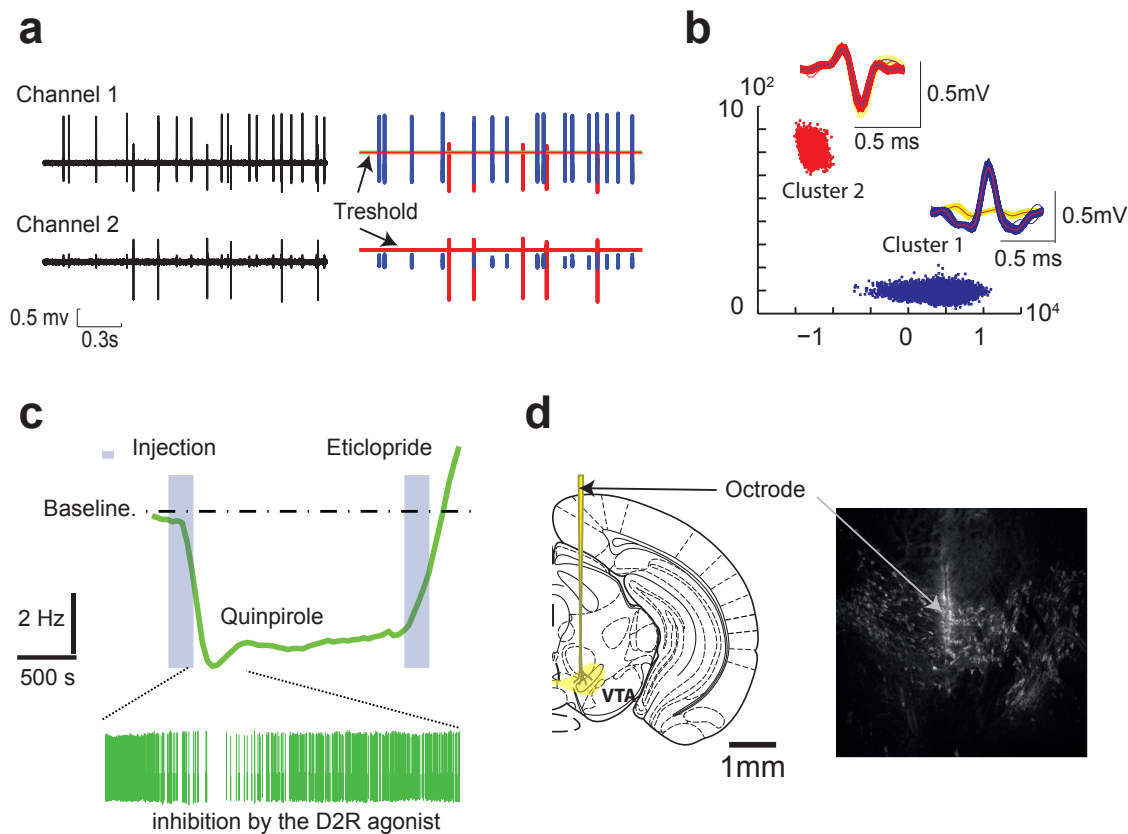


Figure 6.4 – **Spike time analyses** **a** Two recordings obtained with the Neuralynx recording system **b** Representation of the spikesorting process **c** An inactivation and re-activation of a dopaminergic cell by respectively Quinpirole and Eticlopride **d** A microscopic image of an electrode trace placed in the VTA

## Pharmacological identification

D2 receptors (D2R) pharmacology was also used for confirmation of DA neurons identification: after a baseline (10 min) and a saline (dose, 10 min) injection, quinpirole (D2R antagonist) was injected (10 min), followed by an eticlopride (D2R agonist) injection (10 min). As most DA neurons, but not GABA neurons, express these inhibitory D2 auto-receptors, neurons were considered as confirmed as DA neurons if quinpirole induced at least 30% decrease in their firing rates, while eticlopride at least restored the firing above baseline. Nevertheless, as continuous D2 pharmacology could have affected both baseline DA neurons firing and decision-making, we allowed the mice to recover two days after this experiment.

## 6.4 $\beta$ 2 knock out mice and local re-expression

### 6.4.1 $\beta$ 2 knock out mice

The encoding of the of neuronal nAChR  $\beta$ 2-subunit is disrupted as described in Picciotto et al., 1995 [155]. Initiator methionine and the rest of the first exon were replaced with the coding region of NLS-lacZ and the MC1 neo<sup>R</sup> expression cassette [242]. The knockout SOPF HO ACNB2 ( $\beta$ 2KO) and wildtypes were not littermates in our experiment and this could be a potential caveat of the study. However, mutant mice were generated almost 20 years ago, the line has been backcrossed more than 20 generations with the wild-type C57BL6/J line, and the  $\beta$ 2KO line was confirmed to be at more than 99.99% C57BL/6J [243].

### 6.4.2 Stereotaxic injection of lentivirus

The lentiviral expression vectors  $\beta$ 2 subunit IRES-eGFP cDNAs and the eGFP cDNA (control) are under the control of the ubiquitous mouse phosphoglycerate kinase (PGK) promoter as described in Maskos et al., 2005 [25].

$\beta$ 2KO mice aged 8 weeks were anaesthetised using isoflurane. The mouse was introduced into a stereotaxic frame adapted for mice. Lentivirus (2  $\mu$ l at 75 ng of p24 protein per  $\mu$ l) was injected bilaterally at: anteroposterior = +3.4 mm, mediolateral = +/-0.5 mm from bregma and dorsoventral = 4.4 mm from the surface for VTA injection. Mice were implanted with stimulation electrodes 4-5 weeks after viral injection. At the end of the behavioural experiments, lentiviral re-expression in the VTA was verified using fluorescence immunohistochemistry. As a control for  $\beta$ 2VEC mice, another group of  $\beta$ 2KO mice were injected with lentivirus expressing eGFP only [243].

## 6.5 Optogenetic methods

### 6.5.1 Virus injection and optogenetic experiments

DAT-iCre mice were anaesthetised (Isoflurane 1%) and implanted with an ICSS electrode as described above. The DAT-iCre mice were then intracranially injected (1  $\mu$ L injected into the VTA [coordinates from bregma AP: +3.2 mm; ML: +0.5 mm; DV: +4.6



mm] with an adeno-associated virus (AAV5.EF1.DIO.CatCh.eYFP). A double-floxed inverse open reading frame (DIO) allowed to restrain the expression of CatCh (Ca<sup>2+</sup> translocating channelrhodopsin) to VTA dopamine neurons (DANs). An optical fiber (200  $\mu$ m core, NA=0.5, Thor Labs) coupled to a ferule (1.25 mm) was implanted just above the VTA [coordinates from bregma AP: +3.2 mm; ML: +0.5 mm; DV: +4.4 mm], and fixed to the skull with dental cement (SuperBond, Sun medical). An ultra-high power LED (470 nm from Prizmatix) coupled to a patchcord (500  $\mu$ m core, NA=0.5, Prizmatix) was used for optical stimulation (output intensity of 10 mW). Optical stimulation was delivered at a frequency of 8hz with 5ms/pulse, starting 5 min before and during the behavioural experiment, following a schedule of 4 paired ON and OFF days after the training phase of behavioural task was finished, at least 4 weeks after virus injection, to allow the construct to be integrated in the target cells genome. The optical stimulation cable was plugged onto the ferule during all experimental sessions to prepare the animals and control for latent experimental effects.

For a functional check of CatCh expression, 10-12 week old male DAT-iCre mice were injected with the same virus described above. After 4 weeks, mice were deeply anesthetized with an I.P. injection of a mix of ketamine/xylazine. Coronal midbrain sections (250  $\mu$ m) were sliced using a Compressstome (VF-200; Precisionary Instruments) after intracardial perfusion of cold (4°C) sucrose-based artificial cerebrospinal fluid (SB-aCSF) containing (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 5.9 MgCl<sub>2</sub>, 26 NaHCO<sub>3</sub>, 25 Sucrose, 2.5 Glucose, 1 Kynurenate (pH 7.2, 325 mOsm). After 10 min to 1h at 35°C for recovery, slices were transferred into oxygenated aCSF containing (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 2 CaCl<sub>2</sub>, 1 MgCl<sub>2</sub>, 26 NaHCO<sub>3</sub>, 15 Sucrose, 10 Glucose (pH 7.2, 325 mOsm) at room temperature for the rest of the day and individually transferred to a recording chamber continuously perfused at 2 ml/min with oxygenated aCSF. Patch pipettes (4-8 M $\Omega$ ) were pulled from thin wall borosilicate glass (G150TF-3, Warner Instruments) using a micropipette puller (P-87, Sutter Instruments, Novato, CA) and filled with a KGlu based intra-pipette solution containing either (in mM): 144 KGlu, 3 MgCl<sub>2</sub>, 10 HEPES, 0.5 EGTA and 3 mg/ml biocytin (pH 7.2, 295mOsm) or: 116 K-gluconate, 10 20 HEPES, 0.5 EGTA, 6 KCl, 2 NaCl, 4 ATP, 0.3 GTP and 2 mg/mL biocytin (pH adjusted to 7.2). Transfected VTA DANs were visualised using an upright microscope coupled with a Dodt contrast lens and illuminated with a white light source

(Scientifica). A 460 nm LED (Cooled) was used for both visualizing eYFP positive cells (using a bandpass filter cube) and for optical stimulation through the microscope (1s continuous for light-evoked current in voltage-clamp mode and 10Hz with 5ms/pulses to drive neuronal firing in current-clamp mode). Whole-cell recordings were performed using a patch-clamp amplifier (Axoclamp 200B, Molecular Devices) connected to a Digidata (1550 LowNoise acquisition system, Molecular Devices). Signals were lowpass filtered (Bessel, 2 kHz) and collected at 10 kHz using the data acquisition software pClamp 10.5 (Molecular Devices). All the electrophysiological recordings were extracted using Clampfit (Molecular Devices) and analysed with R.

For an *in vivo* functional check of CatCh expression, transfected DANs were recorded in anesthetized animals as previously described. In that case, an optical fiber (500  $\mu\text{m}$  core, Prizmatix) was inserted in the glass pipette electrode and coupled to a 470 nm LED (Prizmatix). Light-pulses trains (1, 2, 5, 10, 20Hz with 5ms/pulse) were applied to test the functional expression of CatCh in VTA DANs.

## 6.6 Fluorescence immunohistochemistry

After euthanasia, brains were rapidly removed and fixed in 4% paraformaldehyde. Following a period of at least 3 d of fixation at 4°C, serial 60- $\mu\text{m}$  sections were cut from the midbrain with vibratome. Immunohistochemistry was performed as follows: free-floating VTA brain sections were incubated 1 h at 4°C in a blocking solution of phosphate-buffered saline (PBS) containing 3% Bovine Serum Albumin (BSA, Sigma; A4503) (vol/vol) and 0.2% Triton X-100 (vol/vol) and then incubated overnight at 4°C with a mouse anti-tyrosine hydroxylase antibody (TH, Sigma, T1299) at 1:200 dilution and a rabbit anti-GFP antibody (Molecular Probes, A-6455) at 1:5,000 dilution in PBS containing 1.5% BSA and 0.2% Triton X-100. The following day, sections were rinsed with PBS and then incubated 3 h at 22-25°C with Cy3-conjugated anti-mouse and Cy2-conjugated anti-rabbit secondary antibodies (Jackson ImmunoResearch, 715-165-150 and 711-225-152) at 1:200 dilution in a solution of 1.5% BSA in PBS. After three rinses in PBS, slices were wet-mounted using Prolong Gold Antifade Reagent (Invitrogen, P36930). Microscopy was carried out with a fluorescent microscope, and images captured using a camera and ImageJ imaging software.

In the case of electrophysiological recordings, an immunohistochemical identification of the recorded neurons was performed as described above, with the addition of 1:200 AMCA-conjugated Streptavidin (Jackson ImmunoResearch) in the solution. Neurons labelled for both TH and neurobiotin in the VTA allowed to confirm their neurochemical phenotype.

In the case of optogenetics experiments on DAT-iCre mice, an immunohistochemical identification of the transfected neurons was performed as described above, with the addition of 1:500 Chicken-anti-YFP primary IgG (ab13970, Abcam) in the solution. A Goat-anti-chicken AlexaFluor 488 (1:500, Life Technologies) was then used as secondary IgG. Neurons labelled for TH, eYFP and neurobiotin/biocityn in the VTA allowed to confirm their neurochemical phenotype and the transfection success.

## 6.7 Overview of performed experiments

### 6.7.1 Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking

The first paper presented in the results section describes the model fitting process that describes the behaviour of WT mice in the multi-armed ICSS bandit task. Three experimental groups are compared: wild type mice;  $\beta 2$  knock out mice and  $\beta 2$  knock out mice with re-expression of the  $\beta 2$  subunit specifically in the VTA. All mice were implanted with the ICSS a week before the training in the multi-armed bandit task took place. Virus injections took place 4-5 weeks before ICSS implantation for the third experimental group. Juxtacellular recordings were performed in parallel in non-implanted animals (*see Figure 6.5*).

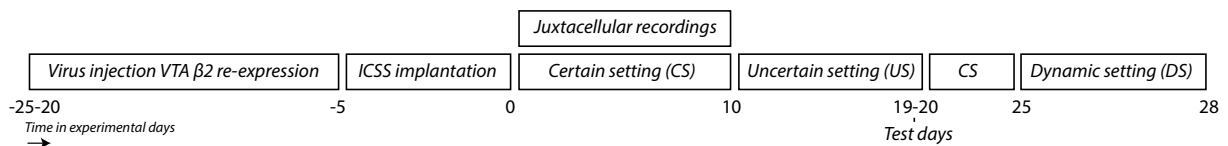


Figure 6.5 – **Timeline of performed experiments in Chapter 7**

After the two test days in the uncertain setting (Day 19 & 20), mice were retrained in the certain setting in order to restore and test baseline behaviour. After 5 training

days mice were assessed in the dynamic setting (DS) in which flexibility behaviour was compared between the different groups. A session was split up in three parts in which the probabilities (100%, 100%, 0%) of the three points altered every sub-session (*see Chapter 7 for the full description of the methods*).

### 6.7.2 Chronic nicotine enhances value sensitivity

The second paper in the results section compares two experimental groups in the multi-armed bandit task: mice implanted with saline minipumps and nicotine minipumps. These two groups were similarly implanted with an ICSS electrode a week before the first training session. Minipumps were implanted at day 5 of the training to allow recovery for 2 days (*see Figure 6.6*). After training and test days (Day 19 & 20) in the probabilistic setting, mice were either assigned to the juxtacellular recording experiment; intensity increase experiment for SAL mp mice or the intensity decrease experiment for the NIC mp mice. To control for the effect of ICSS stimulation, 2 control groups (SAL and NIC minipumps, without ICSS) were also tested in the juxtacellular recording experiment. In parallel, a group of DAT-iCRE mice was injected with the AAV5.EF1a.DIO.CatCh.eYFP virus in the VTA and implanted with a fiber in the VTA and ICSS electrode in the MFB 3 weeks before training to allow the virus to be expressed. These mice underwent a series of manipulations during day 21-28 in which the VTA was alternately photostimulated and not photostimulated (*see Chapter 8 for the full description of the methods*).

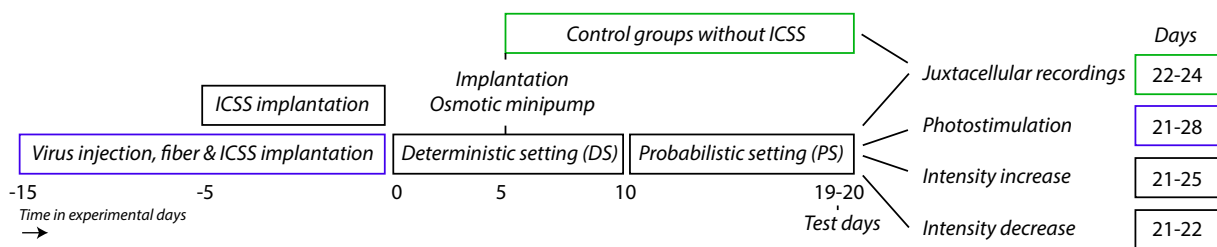


Figure 6.6 – **Timeline of the performed experiments in Chapter 8**



# Results



## 7 | Nicotine receptors in the ventral tegmental area promote uncertainty-seeking

As described in the introduction, acetylcholine (ACh) is a known modulator of the activity of DAergic neurons and consequently of DA release. This modulation functions in particular through the stimulation of nicotinic ACh receptors. Understanding normal habit learning is relevant for the comprehension of modification underlying nicotine addiction. Theories of decision-making suggest that individuals analyse potential benefits and costs to guide their actions. Thus, making appropriate choices requires learning the value of available options from experience. This process is being hypothesised to depend on DA cells activity. One crucial question is whether these basic computations underlying choices can be persistently modified by drug exposure and/or the manipulation of nAChR distribution. This chapter introduces a paper where we describe a task aiming at addressing the role of nAChR in uncertainty seeking.

A key element in decision making is exploration and response to uncertainty. Uncertainty governs our daily choices, as it motivates to explore apparently less advantageous alternatives. Cigarette smokers seem to exhibit alterations of exploratory behaviours and of sensitivity to uncertainty, which may favour the tendency to focus on the familiar drug reward, despite possible negative consequences. However, the mechanisms through which nicotine affects exploration of uncertain outcomes remain unknown. In this work, we adapted the classic *multi-armed bandit task* to mice. We demonstrated that mice lacking the  $\beta 2^*$ -subunit of the nicotinic acetylcholine receptors ( $\beta 2^*$ -nAChRs) displayed a decrease in uncertainty-based exploratory choices and locomotion. Viral re-expression of functional  $\beta 2^*$ -nAChRs in the ventral tegmental area (VTA) restored the value given to uncertainty. Hence, the nicotinic regulation of the VTA translates expected uncertainty into exploration. This paradigm and the findings will also serve as a basis to identify the role of nicotine in a value-based decision-making task (*see Chapter 4*). Some key concepts are introduced below.



## 7.1 Expected uncertainty

Uncertainty is a state of having limited knowledge about the current situation, a future outcome or multiple possible outcomes. Uncertainty can be measured when probabilities are assigned to the different possible outcomes. Risk, is a state of uncertainty where some possible outcomes have an undesired effect or significant loss. In daily life when making sense of the world and making predictions about the future one is struck by a lot of uncertainties. Therefore, uncertainty signals are critical when measuring information and assessing the accuracy of predictions.

As mentioned before, the dopaminergic signal is indispensable for the encoding of values in value-based decision-making and modification of this DA signal can therefore have detrimental effects on choice. The value ('utility') of a reward can hold an assemble of different properties, like magnitude, probability and subjective preference. The DA system is capable of ordering values (including these properties) on a common scale and update them in an adaptive manner when changes occur [41, 244–246].

Fiorillo and colleagues (2003) have demonstrated a positive correlation between the magnitude of the phasic activation and the reward probability a cue predicted (0; 0.25; 0.5; 0.75 or 1.00). Similarly, an inversed correlation was found between the DA phasic activity and the actual reward delivery. For example, when a reward with a probability of  $P = 0.25$  was delivered, a phasic DA activity was observed at the moment of reward, a signal similar to an unexpected reward (as seen in the non-probabilistic version of this study [41]). However, when a fully predicted reward ( $P = 1.00$ ) was delivered no difference in DA activity was observed (RPE=0, [246]). Moreover, DA neurons also encode expected uncertainty. Uncertainty can be presented as a U-curve reflecting the -variances- of the different available reward probabilities. Expected uncertainty thus is highest at a probability of  $P=0.5$  and 0 at  $P=0$  and  $P=1$ . The level of uncertainty is associated with slowly increasing phasic activity (ramping activity) between the cue presentation and the moment of reward delivery. The slope of these ramps represents the expected uncertainty and is steepest when  $P=0.5$ . Different studies have demonstrated a link between expected uncertainty and acetylcholine [229, 247]. In a Posner task<sup>1</sup>, relevant and irrelevant cues are being presented before a target stimulus. The validity effect (VE) is the difference between the time to respond to the target stimulus after a valid cue (stimulus is presented

---

1. A neuropsychological task that assesses the ability to perform an attentional shift

in location predicted by the cue) and an incorrect cue (inversed). VE varies inversely with the levels of ACh: nicotine reduces reaction times, while ACh depletion lengthens the reaction times, suggesting an increase in attention and uncertainty processing through nAChR activation.

## 7.2 Exploration-exploitation trade-off

Value-based decision making is largely dependent on learning and having experienced the different alternatives before being able to assign values to these alternatives. However, when an environment is changing and organisms are presented with unlearned alternatives or uncertainties, these unknown alternatives need to be explored in order to assign a value (in comparison to the other available options). The exploration-exploitation trade-off is therefore very dependent on the stability of the environment. Stability favours exploiting knowledge to maximise gains, while changeability favours exploring new options and discovering new outcomes [234].

The action-selection probability of selecting a high probability reward (optimal choice) versus a low probability reward can be illustrated with a classic softmax function, a variant of the matching law<sup>2</sup>. In the softmax function, the inverse temperature parameter ( $\beta$ ) represents the sensitivity to the difference of values. The larger the inverse temperature, the more likely one chooses the expected optimal action. A low inverse temperature parameter can be considered as risk-taking (too much exploration), while a high inverse temperature reflects risk aversion (too much exploitation) [249,250].

### 7.2.1 Bandit-style problems

Bandit problems illustrate the fundamental difficulty of decision-making when facing uncertainty. A decision maker must choose between exploiting what seems to be the best choice or exploring an alternative, hoping to discover a better option. There is large range of fields in which bandit-style problems can be applied [251]. The classical example to

---

2. The matching law describes a quantitative relationship between the relative rates of response and the relative rates of reinforcement. I.e. the bigger the chance of obtaining a reward, the more likely the selection of the option [248]

test this problem neuroscientifically is where the subject has access to a virtual row of slot machines. The subject has to decide which machine to play and when to switch to another machine. Each machine provides a random reward from a probability distribution specific to that machine. In the 'restless' version of the bandit task these probability distributions change over time, making an optimal choice even more difficult. The objective is to maximise the sum of reward by pulling the levers of the slot machines for a certain number of trials.

The RL  $\epsilon$ -greedy strategies are developed to solve the bandit problem. They all have in common a greedy algorithm, which follows the problem-solving heuristic of making the locally optimal choice at each stage, i.e. choosing the best lever in every trial of the multi-armed bandit task. Epsilon ( $\epsilon$ ) is a parameter reflecting the proportion of randomness during the task, with or without a learning or contextual effect [235,252]. The deficiency of the  $\epsilon$ -greedy models is that it does not take state transitions into account. Optimistic algorithms that make use of Markovian decision processes are considered more aggressive and optimal to solve a bandit problem. However, these methods can only be applied to stationary bandit problems (non-restless) as will be presented in this paper.

### **7.3 Nicotine receptors in the VTA promote uncertainty-seeking**

Overall, in this paper we demonstrate a role for  $\beta 2^{-/-}$  in uncertainty seeking and exploration. Importantly, our results show that WT mice behave in this task as if they give a supplementary value to actions associated with the highest variance in the outcome. Otherwise stated, they visit more often than expected, compared to the value based on the probability, the point associated with 50%. WT mice seem attracted by uncertainty. This property disappeared in  $\beta 2$  KO mice.

# Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking

Jérémy Naudé<sup>1-3</sup>, Stefania Tolu<sup>1-3</sup>, Malou Dongelmans<sup>1-3</sup>, Nicolas Torquet<sup>1-3</sup>, Sébastien Valverde<sup>1-3</sup>, Guillaume Rodriguez<sup>1-3</sup>, Stéphanie Pons<sup>4</sup>, Uwe Maskos<sup>4</sup>, Alexandre Mourot<sup>1-3</sup>, Fabio Marti<sup>1-3</sup> & Philippe Faure<sup>1-3</sup>

**Cholinergic neurotransmission affects decision-making, notably through the modulation of perceptual processing in the cortex. In addition, acetylcholine acts on value-based decisions through as yet unknown mechanisms. We found that nicotinic acetylcholine receptors (nAChRs) expressed in the ventral tegmental area (VTA) are involved in the translation of expected uncertainty into motivational value. We developed a multi-armed bandit task for mice with three locations, each associated with a different reward probability. We found that mice lacking the nAChR  $\beta 2$  subunit showed less uncertainty-seeking than their wild-type counterparts. Using model-based analysis, we found that reward uncertainty motivated wild-type mice, but not mice lacking the nAChR  $\beta 2$  subunit. Selective re-expression of the  $\beta 2$  subunit in the VTA was sufficient to restore spontaneous bursting activity in dopamine neurons and uncertainty-seeking. Our results reveal an unanticipated role for subcortical nAChRs in motivation induced by expected uncertainty and provide a parsimonious account for a wealth of behaviors related to nAChRs in the VTA expressing the  $\beta 2$  subunit.**

Acetylcholine (ACh) has a well-studied role in arousal, learning and attention<sup>1,2</sup> and modulates perceptual decision-making, notably through its influence over prefrontal cortices<sup>3</sup>. Decisions are not only driven by sensory information, but also by the animal's expectation of the values associated with alternative choices<sup>4,5</sup>. ACh also affects cost-benefit decision-making<sup>6,7</sup>, albeit through unknown mechanisms. Notably, effects on value-based decisions induced by pharmacological manipulations of ACh or dopamine (DA) often mirror each other<sup>5</sup>. Systemic pharmacological manipulation of either DA or ACh receptors affects the choices between alternatives associated with different delays, costs or risk<sup>5-7</sup>. Disentangling the respective implications of ACh and DA in decision-making is of utmost interest, as psychological diseases such as tobacco addiction or schizophrenia involve alterations of both decision-making and ACh-DA interactions<sup>2,8</sup>.

By opposition to ACh, DA exerts a well-defined role in motivation and reinforcement<sup>9</sup>. DA neurons encode reward prediction errors as bursts of action potentials<sup>9,10</sup>. These bursts may be used as a teaching signal to learn the value of actions<sup>11</sup> or as an incentive signal biasing the ongoing decisions<sup>12</sup>. The bursting activity of DA neurons from the VTA is influenced by ACh, notably through nicotinic acetylcholine receptors containing the  $\beta 2$  subunit ( $\beta 2^*$ -nAChRs)<sup>2,13-15</sup>. Thus, the similarity between the effects of DA and ACh on decision-making may arise from a nicotinic regulation of the VTA. We hypothesized that endogenous ACh, released from mesopontine nuclei to the VTA<sup>2,5,15</sup>, may be involved in value-based decisions.

In the context of decision-making, the concept of exploration is opposed to that of exploitation with regard to a known reward

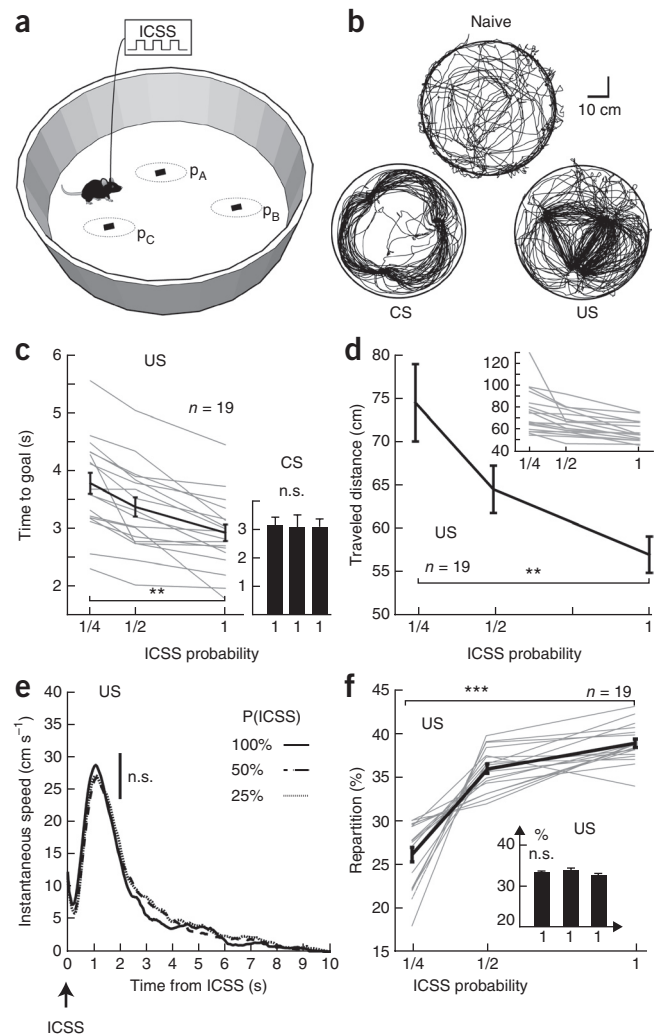
source<sup>16,17</sup>. Exploration occurs when an animal actively gathers information about alternative choices with the aim of reducing the uncertainty level on the consequences of possible actions<sup>18-21</sup>. This typically happens in a learning setting when the statistics of an outcome given a specific action, or its 'uncertainty', are in the process of being estimated. Once the consequences of possible actions have been estimated, the animal can use this knowledge of the environment to exploit reward sources efficiently. However, when the outcome of an action is probabilistic, uncertainty remains as to what will be the outcome of an action every time it is performed. This known variability of the outcome of an action, as in a repeated lottery, is referred to as expected uncertainty or reward risk<sup>22,23</sup>.

The motivation to perform an action can be modulated by expected uncertainty and lead to uncertainty-seeking or risk-taking. In animals, it is challenging to distinguish between a motivation to explore or exploit a probabilistic reward source, as it cannot be easily inferred whether animals might still try to reduce expected uncertainty by exploring<sup>19,24,25</sup> or whether they are attracted by this 'known-unknown' and thus exploit. Nevertheless, the influence of expected uncertainty on motivational value is experimentally tractable. It has been proposed that expected uncertainty may be signaled by ACh<sup>22</sup> in the context of perceptual decision-making, but this theory has never been connected to an involvement of DA in value-based decision-making. Moreover, the neural basis underlying the motivation given to choices associated with expected uncertainty is not known. We computationally characterized the influence of VTA  $\beta 2^*$ -nAChRs on seeking probabilistic rewards in a multi-armed bandit task for mice and found that these receptors are involved in translating expected uncertainty into motivational value.

<sup>1</sup>Sorbonne Universités, UPMC University Paris 06, Institut de Biologie Paris Seine, UM 119, Paris, France. <sup>2</sup>CNRS, UMR 8246, Neuroscience Paris Seine, Paris, France. <sup>3</sup>INSERM, U1130, Neuroscience Paris Seine, Paris, France. <sup>4</sup>Institut Pasteur, CNRS UMR 3571, Unité NISC, Paris, France. Correspondence should be addressed to P.F. (phfaure@gmail.com).

Received 9 October 2015; accepted 9 December 2015; published online 18 January 2016; doi:10.1038/nn.4223

**Figure 1** Decisions under uncertainty in a mouse bandit task using intracranial self-stimulations. **(a)** Illustration of the spatial multi-armed bandit task design. Three explicit square locations were placed in the open field (0.8-m diameter), forming an equilateral triangle (50-cm side). Mice received an intracranial self-stimulation each time they were detected in the area of one of the rewarding locations. Animals, which could not receive two consecutive stimulations at the same location, alternated between rewarding locations. **(b)** Trajectories of one mouse (5 min) before (left) and after (middle) learning in the CS and US (right). **(c)** Time to goal (average duration from the last location to the goal) in the US as a function of the reward probability of the goal. Inset, times to goal were identical for the three locations in the CS ( $F_{(2,18)} = 0.53$ ,  $P = 0.59$ , one-way ANOVA). Inset, individual curves.  $N = 19$  mice. **(d)** Traveled distance between two consecutive locations. In the US, WT mice traveled more distance when going toward less probable ICSS reward. Light gray, individual curves. **(e)** Instantaneous speed: in the US, the maximal speed of WT mice did not depend on the expected probability of the reward, contrary to what was observed in the DS with increasing intensity. Data are presented as mean  $\pm$  s.e.m. Time 0 corresponds to the last time of ICSS delivery or omission. **(f)** Proportion of choices of the three rewarding locations as a function of reward probability in the US. Light gray, individual curves. Inset, proportion of choices were identical for the three locations in the CS ( $F_{(2,18)} = 0.16$ ,  $P = 0.86$ , one-way ANOVA). Error bars represent mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . n.s., not significant at  $P > 0.05$ .



## RESULTS

### Mice-adapted multi-armed bandit task based on ICSS

In uncertain environments, living beings have to decide when to exploit known resources and when to explore alternatives. This exploitation-exploration dilemma is often studied in the multi-armed bandit task<sup>16,18</sup>, in which humans choose between different slot machines to discover the richest option. To assess the implication of nAChRs in decision-making under uncertainty, we designed a spatial version of the bandit task adapted to mice. Studies of animal choices often rely on food restriction, even though the satiation level is known to affect decisions under uncertainty<sup>26</sup>. To circumvent this issue, we trained mice to perform a sequence of choices in an open-field in which three locations were explicitly associated with intra-cranial self-stimulation (ICSS) rewards<sup>27,28</sup> (Fig. 1a and Online Methods). Mice could not receive two consecutive ICSS at the same location. Consequently, they alternated between rewarding locations by performing a sequence of choices. Mice mostly went directly to the next rewarding location, but sometimes wandered around in the open field before reaching the goal (Fig. 1b). At each location, mice had to choose which next rewarding location to go to (amongst the two alternatives) and how directly they should get there.

We compared the behavior of wild-type (WT) mice under two settings of ICSS delivery: a certain setting (CS) in which all locations were associated with a given ICSS, and an uncertain setting (US), in which each location was associated with a different probability of ICSS delivery (Fig. 1a). Although trajectories in the CS were stereotyped, reward uncertainty induced a markedly different behavioral pattern in the US (Fig. 1b). The time to goal was identical for the three locations in the CS (Fig. 1c), but was greater for locations associated with lower reward probabilities in the US ( $F_{(2,18)} = 6.8$ ,  $P = 0.002$ , one-way ANOVA; Fig. 1c). More precisely, the reward probability of the goal affected the traveled distance ( $F_{(2,18)} = 7.3$ ,  $P = 0.002$ ; Fig. 1d), but not the traveling speed ( $F_{(2,18)} = 0.48$ ,  $P = 0.62$ ; Fig. 1e) or the dwell times (Supplementary Fig. 1a). This contrasts with the effect of reward intensity, which affected the speed profiles between two rewarding locations (Supplementary Fig. 1b). Thus, in this setup, reward intensity affected the invigoration of goal-directed movements, whereas

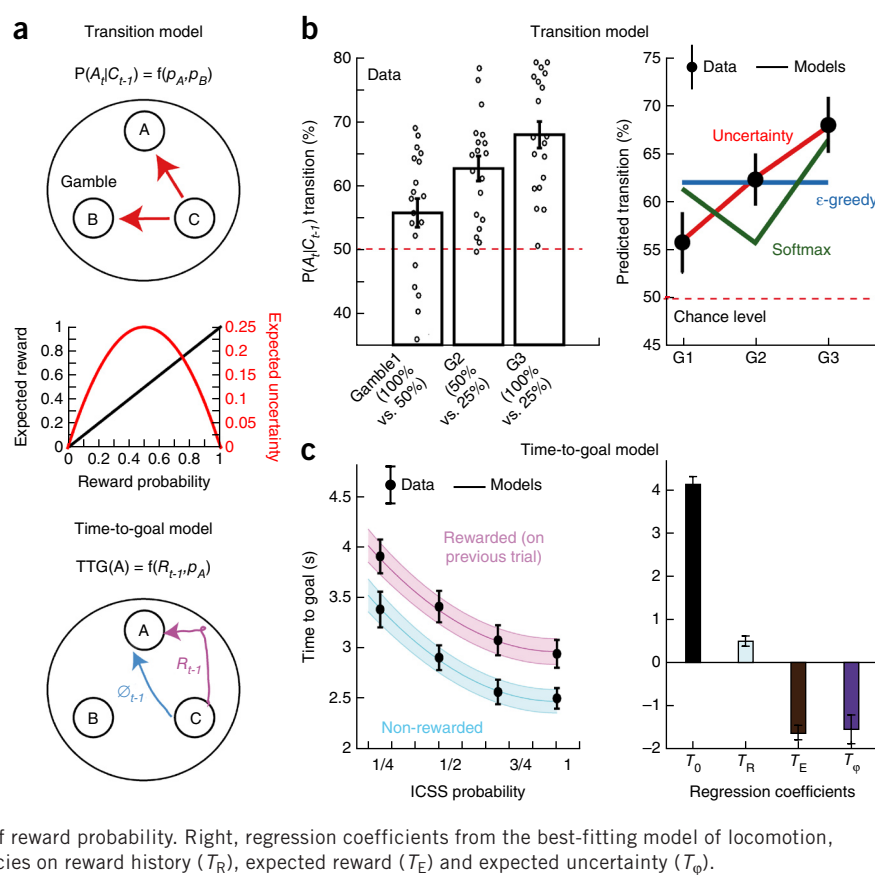
reward probability affected the extent of locomotion, reflecting the tendency to explore the open field between two visits.

In addition, mice distributed their choices of ICSS according to the reward probability associated to each location. As expected, in the CS, mice treated each rewarding location the same way (Fig. 1f). In the US, however, mice visited the locations associated with higher ICSS probability more often ( $F_{(2,18)} = 113$ ,  $P < 0.001$ , one-way ANOVA; Fig. 1f). Because mice could not receive two consecutive ICSSs, the repartition on the rewarding locations (Fig. 1f) arose from a sequence of binary choices in three gambles (G1, G2, G3) between two respective payoffs (here, G1 = {100 versus 50%}, G2 = {50 versus 25%}, G3 = {100 versus 25%}; Fig. 2a,b). For each gamble, mice chose the optimal location (associated with the highest probability of reward; Fig. 2b) more than 50% of the time, but less than 100% of the time. When they had to choose between a certain (100%) and an uncertain (50%) ICSS, mice displayed a low preference (56%) for the optimal location, suggesting a positive inclination toward reward uncertainty (Fig. 2a)<sup>29,30</sup>.

### A positive motivational value to expected uncertainty

In standard rodent decision tasks in which there is only a single choice, the relative influence of expected value and uncertainty on choices is difficult to dissect, as both parameters vary with reward probability. For binary outcomes (the choice is rewarded or not), the expected mean reward corresponds to the reward probability  $p$ ,

**Figure 2** Model-based analysis of decisions shows motivation for expected uncertainty. (a) Illustration of the modeling of the task. Top, transition model of animal choices. Each rewarding location is modeled as a state, labeled {A,B,C}. The probability of transition from one state to another depends on the reward probabilities of the two available options. Middle, expected reward and uncertainty as a function of reward probabilities (curves). In the three gambles, the differences in expected values (0.5 in G1, 0.25 in G2, 0.75 in G3) and expected uncertainties (−0.25 in G1, 0.0625 in G2, −0.1875 in G3) are distinct. Bottom, model of locomotion. The time to goal depends on both reward history (whether the mouse received a reward in the previous location or not) and reward expectation at the goal. (b) Left, proportions of exploitative choices (choice of the most valuable alternative, that is, with the highest probability of reward in a given gamble) of the mice, for the three gambles. Dots, individual data points. Right, predicted transition of the three decision models (lines) corresponding to the experimental data (dots, same value as in the right panel). Error bars represent mean  $\pm$  s.e.m. (c) Left, time to goal (experimental data and model fit, mean  $\pm$  s.e.m.) as a function of reward probability of the goal and reward history. Data merged from experiments with different sets of reward probability. Right, regression coefficients from the best-fitting model of locomotion, corresponding to a constant ( $T_0$ ) and the dependencies on reward history ( $T_R$ ), expected reward ( $T_E$ ) and expected uncertainty ( $T_\phi$ ).



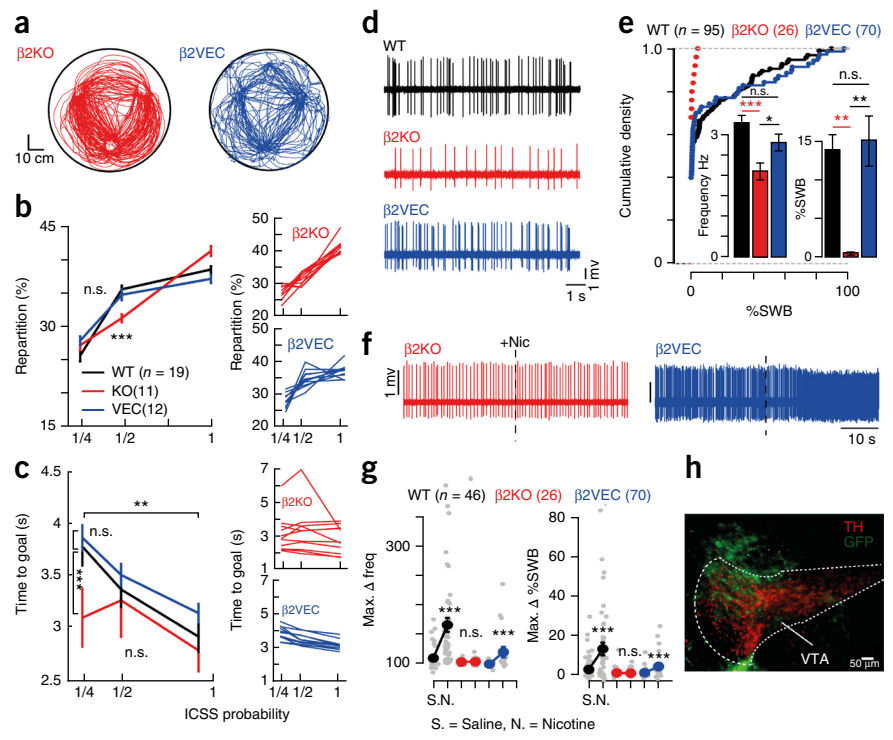
whereas expected uncertainty is related to reward variance,  $p(1-p)$  (Fig. 2a). Expected uncertainty is zero for predictable outcomes (100% or 0% probability) and maximal at 50% probability (the most unpredictable outcome). In our setup, the difference in expected uncertainty and value between the outcomes was distinct for each of the three gambles (Fig. 2a), which provides enough constraints to differentiate between the influence of two co-varying parameters (reward mean and variance). We compared computational models of decision-making<sup>16,31</sup> (Online Methods), each representing a different influence of expected reward and uncertainty on choices, to assess which model best explained the experimental data (Supplementary Fig. 2). In the epsilon-greedy model, animals always choose the best option, minus a fixed probability. In this model, the choices for the optimal reward are identical whatever the gamble is (Fig. 2b), which did not correspond to the experimental data. In the softmax model, choices depend on the difference between the expected rewards of the two alternatives. The softmax model formalizes that the larger the difference in rewards is, the higher the probability to select the best option will be. This model predicts that the proportions of optimal choices would be sorted in the following order {G2 < G1 < G3}, differing from what was found experimentally {G1 < G2 < G3}. Finally, in the uncertainty model, decision is biased toward actions with the most uncertain consequences by assigning a bonus value<sup>32</sup> to their expected uncertainties<sup>19,21,24,29</sup>. This last model accurately reproduced the pattern of mice preferences (Fig. 2b) and best accounted for our experimental data (Supplementary Fig. 2), as shown by model comparison (likelihood penalized for the number of parameters, Bayesian Information Criterion (BIC); Online Methods). Furthermore, the two parameters of the ‘uncertainty bonus’ model disentangle two determinants of decision-making: the inverse temperature parameter  $\beta$

depicts the randomness in choices, whereas the uncertainty-seeking parameter  $\phi$  represents the value given to expected uncertainty. The positive uncertainty bonus ( $\phi = 1.01 \pm 0.24$ , mean  $\pm$  s.e.m.) explains the great attractiveness of the 50% choice in G1 by a powerful motivation induced by its expected uncertainty. We assessed the robustness of the data and of the model by fitting four sets of probabilities, with multiple different differences of expected reward and uncertainties (Supplementary Fig. 3), and compared alternative models (matching law<sup>33</sup> and uncertainty-normalized temperature<sup>34</sup>; Supplementary Fig. 2). Overall, we found that expected uncertainty positively biased the choices in WT mice.

As stated above, two types of decisions are nested in the task: the sequence of choices (“which goal?”) and the locomotion (“how to reach the goal?”). To investigate the influence of uncertainty on the latter, we performed multiple linear regressions of time to goal. Comparison of linear models (BIC; Online Methods and Supplementary Fig. 2) revealed that the time to goal depended on the reward probability of the goal, but not on the alternative (the location not chosen in the gamble). These observations suggest a dual-stage process in which animals first choose which location to go to and then how to reach it. Furthermore, the dependence on reward history ( $T_R = 0.49 \pm 0.21$ , mean  $\pm$  s.e.m.) suggests that when mice had just gotten rewarded, they traveled further in the open field (Fig. 2c). We also found that the time to goal was decreased by the expected reward ( $T_E = -1.63 \pm 0.16$ ; Fig. 2c) and by the expected uncertainty ( $T_\phi = -1.56 \pm 0.33$ ). This suggests that expected uncertainty increased motivation to go straight toward the rewarding goal. Thus, model-based analyses suggest that, in the two decision problems (“which location” and “how to get there”), mice assign a positive motivational value ( $\phi$  and  $T_\phi$ ) to the expected uncertainty of the goal.

**Figure 3**  $\beta 2^*$ -nAChRs in the VTA affect choices and locomotion. (a) Behavioral trajectories after learning in the US for  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. (b) Proportion of choices of the three rewarding locations plotted as a function of reward probability in the US for the WT (black),  $\beta 2$ KO (red,  $n = 11$ ) and  $\beta 2$ VEC (blue,  $n = 12$ ) mice. Insets, individual curves for the  $\beta 2$ KO (top, red) and  $\beta 2$ VEC (bottom, blue) mice. (c) Time to goal (in seconds) as a function of reward probability of the goal for the WT (black),  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. Insets, individual curves for the  $\beta 2$ KO (top, red) and  $\beta 2$ VEC (bottom, blue) mice. (d) Examples of *in vivo* juxtacellular recordings of the firing pattern of DA neurons from anesthetized WT (black),  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. (e) Cumulative distribution of percent of spikes in a burst (%SWB). Insets, mean frequency (left) and %SWB (right) of VTA DA neurons from the three genotypes (obtained from 22 WT mice, 13  $\beta 2$ KO mice and 13  $\beta 2$ VEC mice). (f) Typical electrophysiological recording illustrating the effect of intravenous injection of nicotine on the firing pattern of DA neurons in  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. Dots, individual data points. (g) Relative variation in firing frequency (left) and absolute variation in %SWB of DA neurons from the three genotypes (obtained from 14

WT mice, 13  $\beta 2$ KO mice and 13  $\beta 2$ VEC mice) in response to nicotine. Error bars represent mean  $\pm$  s.e.m. (h) Coronal sections of the VTA showing the site of lentivirus injection revealed that  $\beta 2$ -eGFP colocalized with TH, a dopaminergic marker. Transduction of  $\beta 2$ -eGFP virus was efficient in both dopaminergic and non-dopaminergic cells. Dots, individual data points. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . n.s., not significant at  $P > 0.05$ .



### VTA $\beta 2^*$ -nAChRs are involved in motivation by uncertainty

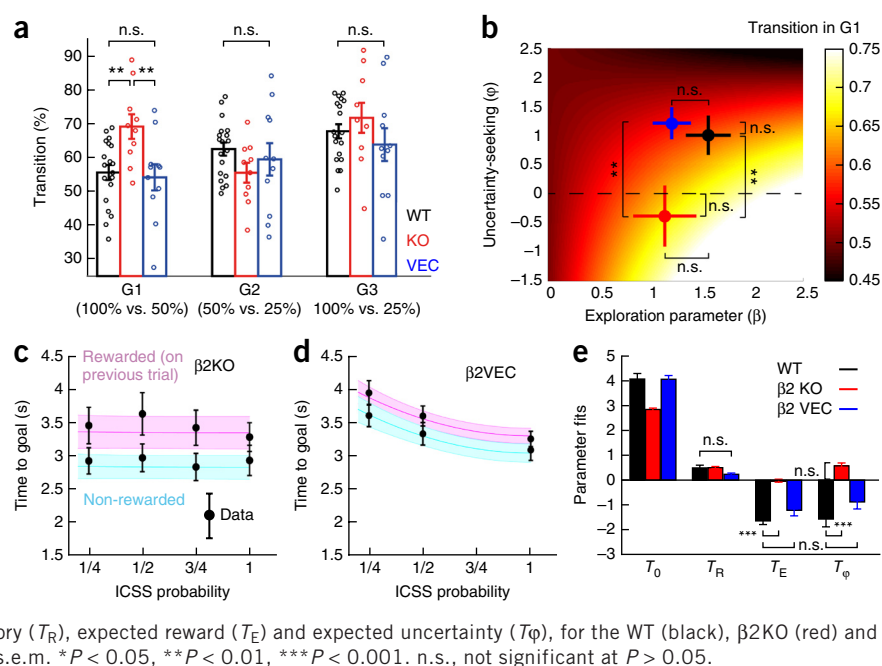
In the ICSS bandit task, WT mice displayed a robust preference for uncertain outcomes. Thus, mice estimate expected uncertainty to direct their decisions and locomotion<sup>29,30</sup>. The suggested role of ACh in signaling expected uncertainty<sup>22</sup> prompted us to investigate whether nAChRs are involved in uncertainty-driven motivation. We used mice in which the  $\beta 2$  subunit, the most abundant nicotinic subunit in the brain<sup>1,2</sup>, was deleted ( $\beta 2$ KO mice), in our ICSS bandit task. In the CS,  $\beta 2$ KO mice ( $\beta 2$ KO and  $\beta 2$ GFP ( $\beta 2$ KO mice injected with a lentivirus expressing just eGFP); Online Methods) learned the task and responded to different ICSS current intensities similarly to WT mice (Supplementary Fig. 4), confirming the modest implication of nAChRs in decision-making with certain rewards<sup>35,36</sup>. In contrast, in the US (Fig. 3a),  $\beta 2$ KO mice systematically chose the location associated with the highest uncertainty level (that is, 50% probability) to a lower extent than WT mice ( $T_{(28)} = -5.4$ ,  $P < 0.001$ , unpaired  $t$  test; Fig. 3b). Furthermore, the relationship between time to goal and reward probability of the goal ( $F_{(2,10)} = 0.33$ ,  $P = 0.72$ , one-way ANOVA; Fig. 3c and Supplementary Fig. 4) was abolished in  $\beta 2$ KO animals. These results suggest a role for  $\beta 2^*$ -nAChRs in decision-making under uncertainty.

We next tested whether  $\beta 2^*$ -nAChRs could affect motivation by expected uncertainty by acting on VTA DA neurons, which are important for value-based decision-making<sup>9,12</sup>. Extracellular *in vivo* single-unit recordings in anesthetized animals (Fig. 3d) confirmed that, when compared with those of WT mice, DA neurons from  $\beta 2$ KO mice displayed a decreased firing frequency (2.1 versus 3.2 Hz,  $T_{(74)} = 2.4$ ,  $P < 0.001$ , Welch  $t$  test), lacked bursting activity ( $U = 1,637$ ,  $P = 0.002$ , Mann-Whitney test; Fig. 3e) and did not respond to a systemic injection of nicotine (104.6  $\pm$  1.34% from baseline frequency,  $V = 103$ ,  $P = 0.95$ , Wilcoxon test; Fig. 3f,g and Supplementary Fig. 5e,f)<sup>13,14</sup>. If  $\beta 2^*$ -nAChRs underlie uncertainty-driven motivation in the VTA, then restoring expression of these receptors in the VTA of  $\beta 2$ KO mice should restore both the sensitivity to expected uncertainty and DA activity. We achieved selective re-expression of the  $\beta 2$  subunit in the VTA of  $\beta 2$ KO mice ( $\beta 2$ VEC) using a lentiviral vector<sup>13</sup> strategy (Online Methods). Coronal sections revealed that viral re-expression was restricted to the VTA (Fig. 3h and Supplementary Fig. 5a-d). DA cells from  $\beta 2$ VEC mice displayed a spontaneous firing frequency ( $T_{(156)} = 1.6$ ,  $P = 0.1$ , unpaired  $t$  test) and bursting activity ( $U = 3,288$ ,  $P = 0.9$ , Mann-Whitney test) similar to those observed in WT mice, and responded to nicotine (120.2  $\pm$  4.78% from baseline frequency,

**Table 1** Behavioral measures and model parameters in the uncertain setting

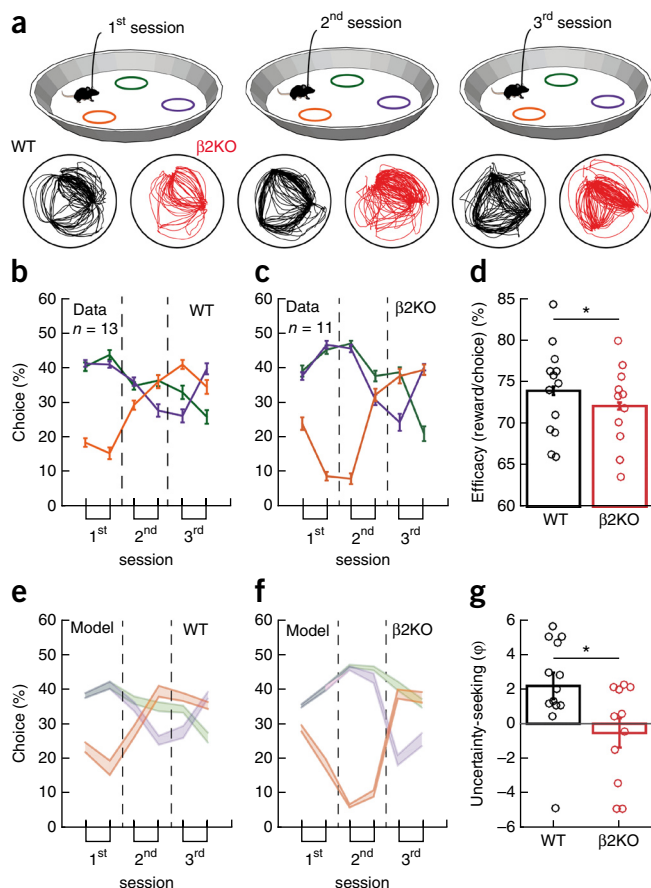
	WT (mean $\pm$ s.e.m.)	$\beta 2$ KO (mean $\pm$ s.e.m.)	$\beta 2$ VEC (mean $\pm$ s.e.m.)	ANOVA	$\beta 2$ VEC versus $\beta 2$ KO	$\beta 2$ VEC versus WT
Repartition at $P = 1/2$ (Fig. 3b)	35.6 $\pm$ 0.5%	31.3 $\pm$ 0.5%	34.8 $\pm$ 0.7%	$F_{(2,39)} = 13.45$ , $P < 0.001$	$T_{(21)} = -3.86$ , $P < 0.001$	$T_{(29)} = 0.96$ , $P = 0.35$
Gamble 1 (Fig. 4a)	55.6 $\pm$ 2.2%	69.1 $\pm$ 3.5%	54.1 $\pm$ 3.7%	$F_{(2,39)} = 6.95$ , $P = 0.002$	$T_{(21)} = 3.04$ , $P = 0.006$	$T_{(29)} = 0.30$ , $P = 0.77$
Uncertainty-seeking parameter (Fig. 4b)	1.01 $\pm$ 0.24	-0.38 $\pm$ 0.47	1.21 $\pm$ 0.23	$F_{(2,39)} = 6.89$ , $P = 0.003$	$T_{(21)} = 3.1$ , $P = 0.005$	$T_{(29)} = -0.6$ , $P = 0.56$
Inverse temperature parameter (Fig. 4b)	1.57 $\pm$ 0.16	1.14 $\pm$ 0.28	1.21 $\pm$ 0.16	$F_{(2,39)} = 1.6$ , $P = 0.22$		
Reward history coefficient (Fig. 4e)	0.49 $\pm$ 0.12	0.52 $\pm$ 0.11	0.26 $\pm$ 0.16	$F_{(2,39)} = 1.6$ , $P = 0.22$		
Reward expectation coefficient (Fig. 4e)	-1.63 $\pm$ 0.16	-0.02 $\pm$ 0.17	-1.21 $\pm$ 0.23	$F_{(2,39)} = 18.4$ , $P < 0.001$	$T_{(21)} = 4.0$ , $P < 0.001$	$T_{(29)} = 1.55$ , $P = 0.13$
Uncertainty expectation coefficient (Fig. 4e)	-1.56 $\pm$ 0.33	0.58 $\pm$ 0.37	-0.88 $\pm$ 0.28	$F_{(2,39)} = 9.8$ , $P = 0.001$	$T_{(21)} = 3.2$ , $P = 0.005$	$T_{(29)} = 1.42$ , $P = 0.17$

**Figure 4** Model-based analysis reveals a role for VTA  $\beta 2$ -nAChR in uncertainty-driven motivation. (a) Transition (proportions of exploitative choices) in the three gambles, for the WT (black),  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. Dots, individual data points. (b) Value of the parameters ( $\beta$  and  $\varphi$ ) derived from the model-based analysis (uncertainty model) of the transition functions for the WT (black),  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. The color code indicates the predicted transition in gamble 1 (100 versus 50% reward probability) as a function of the parameters of the model. (c,d) Time to goal as a function of reward probability of the goal and reward history for  $\beta 2$ KO (c) and  $\beta 2$ VEC (d) mice. Experimental data (black dots with error bars) and model fit (stripes) are displayed as mean  $\pm$  s.e.m. Data are merged from experiments with four sets of reward probabilities. (e) Regression coefficients from the best-fitting model of locomotion, corresponding to a constant ( $T_0$ ) and the dependencies on reward history ( $T_R$ ), expected reward ( $T_E$ ) and expected uncertainty ( $T_\varphi$ ), for the WT (black),  $\beta 2$ KO (red) and  $\beta 2$ VEC (blue) mice. Data are presented as mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . n.s., not significant at  $P > 0.05$ .



$V = 960$ ,  $P < 0.001$ , Wilcoxon test), suggesting that, as previously established<sup>13,14</sup>, physiological functions were also restored. Notably,  $\beta 2$ VEC mice differed from  $\beta 2$ KO animals, but not from WT mice (Table 1), when analyzing the uncertainty-related choices (Fig. 3b) and the times to goal (Fig. 3c and Supplementary Fig. 4), indicating a restoration of the WT phenotype following re-expression of  $\beta 2$  in the VTA.

We next used the model-based analysis to characterize the role of VTA  $\beta 2^*$ -nAChRs in decision-making. Transition functions of  $\beta 2$ KO and WT mice differed in particular in G1 (100 versus 50%,  $T_{(28)} = 3.54$ ,  $P = 0.001$ , unpaired  $t$  test; Fig. 4a), suggesting an alteration of decisions under uncertainty. Indeed, the behavior of  $\beta 2$ KO mice was best explained (Supplementary Fig. 6) either by the softmax model or the uncertainty model in which the sensitivity to uncertainty was null on average ( $T_{(11)} = -0.8$ ,  $P = 0.44$ ,  $t$  test; Fig. 4b). Both models point toward the same interpretation:  $\beta 2^*$ -nAChRs are necessary for translating uncertainty signals into motivational value. Accordingly, uncertainty-seeking was significantly different in  $\beta 2$ KO and WT mice ( $T_{(29)} = 2.9$ ,  $P = 0.007$ , unpaired  $t$  test). Notably, the model-based analysis supports the conclusion that  $\beta 2^*$ -nAChRs selectively re-expressed in the VTA restored the positive value of expected uncertainty (Table 1 and Supplementary Fig. 7). Moreover, the analysis of the trajectories in-between goals indicates that neither expected reward nor expected uncertainty of the next goal influenced the time to goal in  $\beta 2$ KO mice, whereas both parameters affected time to goal in  $\beta 2$ VEC mice (Table 1 and Fig. 4c–e). Mice from each genotype all traveled more distance when the previous trial was rewarded, compared to when it was not ( $F_{(2,39)} = 0.02$ ,  $P = 0.98$ ). Together with the transition model, where the temperature parameter  $\beta$  was not significantly different between genotypes ( $F_{(2,39)} = 1.6$ ,  $P = 0.2$ ,



**Figure 5**  $\beta 2^*$ -nAChRs affect decision-making under uncertainty in a dynamical foraging task. (a) Top, illustration of the task design. During each session, animals receive stimulations in two (of three) potential locations, with the two rewarding locations (indicated by an 'R' in the colored circle) changing between sessions. Bottom, behavioral trajectories in the three 2-min sessions for the WT (black) and  $\beta 2$ KO (red) mice. (b,c) Repartition (in %) on the three locations (color-coded as in a) for the three sessions. Calculation is divided by half-session durations for the WT (b) and  $\beta 2$ KO (c) mice. (d) Proportion of rewarded choices averaged on three sessions for WT (black) and  $\beta 2$ KO (red) mice. Dots, individual data points. (e,f) Model fits of the experimental data shown in b and c. (g) Uncertainty-seeking parameter (that is, value given to uncertainty) of the models for the WT (black) and  $\beta 2$ KO (red). Dots, individual data points. Data are presented as mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .



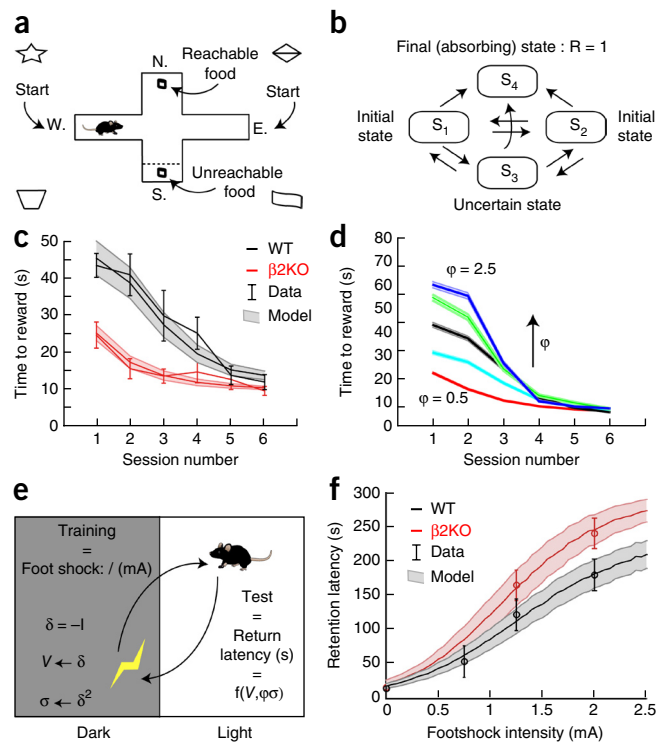
**Figure 6** New interpretation of behaviors related to VTA nAChRs using the uncertainty model. **(a)** Spatial learning task<sup>40</sup>. In a cross maze, the north arm contained a reachable food reward and the south arm contained an unreachable food. The initial position of the animal was variable (east or west). **(b)** Discretized representation of the task.  $S_1$ – $S_4$  are the four possible states in the task;  $R = \{0,1\}$  indicates whether the food reward was attained or not. Arrows represent the possible transitions between the states. Data adapted with permission from ref. 40. **(c)** Simulation (stripes) of the time to reach the food (data: lines with error bars) along the learning sessions for the WT (black) and  $\beta 2\text{KO}$  (red) mice. Parameters were  $\alpha = 0.54$ ,  $\beta = 7.75$ ,  $\varphi = 1.51$  for WT mice,  $\alpha = 0.59$ ,  $\beta = 7.93$ ,  $\varphi = 0.04$  for  $\beta 2\text{KO}$  mice. **(d)** Effect of the uncertainty-seeking parameter in the simulation of the time to reach the food. **(e)** Passive avoidance task<sup>41</sup> consisting in a single training trial in which the mouse was delivered a foot shock upon entrance in the dark chamber. Data adapted with permission from ref. 41. **(f)** Simulation (stripes) of retention latencies (data: lines with error bars) in response to various intensities of foot shock. Parameters were  $\beta = 1.81$ ,  $\varphi = 0.53$ ,  $\theta = 1.62$  for WT,  $\beta = 1.62$ ,  $\varphi = 0.15$ ,  $\theta = 1.65$  for  $\beta 2\text{KO}$ . Error bars represent mean  $\pm$  s.e.m.

one-way ANOVA; **Fig. 4b**), the time-to-goal model strongly suggests that  $\beta 2^*$ -nAChRs do not affect the global motivation to explore, but rather specifically affect expected uncertainty on choices (“which goal”) and locomotion (“how to reach it”).

### $\beta 2^*$ -nAChRs and uncertainty-seeking in a dynamic environment

Having characterized the role of  $\beta 2^*$ -nAChRs in motivation by expected uncertainty at steady state, we next asked whether our results could be extended to unstable environments. We analyzed the behavior of WT and  $\beta 2\text{KO}$  mice during the learning sessions of the CS (**Supplementary Fig. 8a,b**), when reward probabilities were not known yet, and modeled it with reinforcement-learning (RL) models<sup>16,29,31,37,38</sup> (Online Methods and **Supplementary Fig. 8c–f**). In the standard RL model, animals learn the expected value of the three rewarding locations using reward prediction errors (the difference between actual reward and predicted value)<sup>31</sup>. In the model, animals use these values to select the next action using a softmax decision rule. We extended the standard RL model to uncertainty learning. Animals can use reward prediction errors to estimate reward uncertainty<sup>21,23,24,39</sup>: the larger the errors (positive or negative) are, the more uncertain the outcomes will be. This uncertainty RL model best explained the behavior of WT mice (**Supplementary Fig. 8c,e**). By contrast, the behavior of  $\beta 2\text{KO}$  mice was best accounted for by a standard RL model, that is, without uncertainty bonus (**Supplementary Fig. 8d,f**).

To further test the importance of  $\beta 2^*$ -nAChRs for translating expected uncertainty into motivational value, we compared the behavior of WT and  $\beta 2\text{KO}$  mice in a dynamic setting (DS) in which the locations delivering the ICSS reward changed over time (Online Methods). In the DS, mice underwent three consecutive sessions in which only two of the three locations delivered the ICSS. Overall, WT and  $\beta 2\text{KO}$  mice adapted their strategies to these changes in reward contingencies (**Fig. 5a**). Starting from a random strategy, both WT and  $\beta 2\text{KO}$  mice learned the position of the two rewarding locations in the first session (**Fig. 5b,c**). However,  $\beta 2\text{KO}$  mice persevered in their earlier choices throughout the changes in outcomes (**Fig. 5b,c**), resulting in slightly fewer rewarded choices than for WT mice ( $T_{(22)} = 2.7$ ,  $P = 0.01$ , unpaired  $t$  test; **Fig. 5d**). Model comparison (RL models; **Supplementary Fig. 9** and Online Methods) suggested that an uncertainty bonus model best explained the behavior of WT mice (**Supplementary Fig. 9**). This uncertainty model reproduced the choices of WT animals during the changes in rewarding outcomes (**Fig. 5e**), with a positive bonus given to uncertainty ( $\varphi = 2.18 \pm 0.77$ ). This is consistent with the results in the



CS task (**Fig. 4** and **Supplementary Fig. 8**). Moreover, model comparison (**Supplementary Fig. 9**) suggested that experimental data was not better explained by indirect effects arising from learning, that is, asymmetric (different for reward and omission<sup>38</sup>) or adaptive (uncertainty-dependent) learning rates<sup>37</sup>. In summary, our results and models support the idea that, in WT mice, expected uncertainty exerts a direct motivational effect. By contrast, the behavior of  $\beta 2\text{KO}$  mice could be explained by either the standard RL model or the expected uncertainty model (**Fig. 5f** and **Supplementary Fig. 8b,d,f**). In this latter model, the uncertainty-seeking parameter in the  $\beta 2\text{KO}$  mice was significantly lower than that of WT mice ( $T_{(22)} = 2.4$ ,  $P = 0.027$ , unpaired  $t$  test; **Fig. 5g**) and not significantly different from zero ( $T_{(22)} = -0.6$ ,  $P = 0.54$ ). These results provide further evidence that  $\beta 2^*$ -nAChRs are involved in uncertainty-seeking.

### Uncertainty-seeking in other nAChRs-related behaviors

Finally, using computational approaches, we assessed whether the role of VTA  $\beta 2^*$ -nAChRs in uncertainty-seeking might pervade other decisions about natural rewards, punishments and salient aspects of the environment. Paradoxically, it has been found that mice lacking the  $\beta 2$  subunit perform seemingly better than WT mice, displaying ‘improved’ spatial learning<sup>40</sup> and passive avoidance<sup>41</sup>. The spatial learning test consists of a maze with a reachable food reward at one of the four arms and an unreachable food at the opposite arm (**Fig. 6a**). We simulated this task with a RL model embedding uncertainty-seeking (**Fig. 6b** and Online Methods). The model fitted the behavior of both strains, as the time to reach the food was greater for WT (with an uncertainty bonus) than for  $\beta 2\text{KO}$  mice (without bonus) in the early trials<sup>40</sup> (**Fig. 6c**). This slowly decreasing time to reward progressively emerges in RL models embedding uncertainty-seeking (**Fig. 6d**), but cannot be easily explained in terms of differences in initial value (novelty-seeking<sup>31,32</sup>), learning rates or a combination of both (**Supplementary Fig. 10**). Hence, interest for the unreachable reward may arise in WT mice from uncertainty, integrated at the level of the VTA. We also assessed whether the same explanation

holds in the case of punishment. We simulated the passive avoidance task, where animals were in a box divided in two compartments (light and dark).  $\beta 2\text{KO}$  mice avoided the dark compartment, which was associated with a foot shock, for a longer time than WT mice<sup>41</sup> (Fig. 6e). Uncertainty-seeking can also explain this difference, as the foot shock induces a single negative prediction error, which results in uncertainty (Fig. 6f). Expected uncertainty may in that case motivate WT mice, but not  $\beta 2\text{KO}$  mice, to explore the dark part of the box in spite of potential negative consequences. Finally, these models can be extended to neutral, but potentially uncertain, outcomes. The deficits of  $\beta 2\text{KO}$  mice in locomotion in an open-field without rewards<sup>13,42</sup> can be understood as a lack of uncertainty-seeking (Supplementary Fig. 11a–e). Exploration in the open-field is composed of action patterns related to information-seeking (scanning, rearing and sniffing<sup>42,43</sup>). The apparent lack of object recognition observed in  $\beta 2\text{KO}$  mice<sup>40</sup> can also be interpreted as a lack of curiosity for the objects, that is, an absence of uncertainty-seeking (Supplementary Fig. 11f,g), rather than a memory deficit. The uncertainty-seeking model not only generalizes our results to positive, aversive and neutral natural outcomes, but also provides a parsimonious interpretation for a wealth of behaviors associated with  $\beta 2^*$ -nAChRs<sup>13,40–42</sup>.

## DISCUSSION

Our findings reveal a role for VTA  $\beta 2^*$ -nAChRs in translating expected uncertainty into motivational value and suggest that this receptor is involved in exploratory decisions. Three broad theoretical types of exploration have been proposed. At one extreme, exploration is seen as randomness or noise in the choices (as in the softmax or epsilon greedy models), which is problematic, as rodents, similar to primates, display curiosity and refined forms of exploration<sup>29,42,43</sup>. At the other extreme, a dichotomy has been postulated between subcortical systems (such as the VTA) and frontal cortices. Frontal cortices would mediate flexible exploration by overriding<sup>16,17</sup> the influence of exploitive value, underpinned by DA neurons<sup>9,11</sup>. Our results lie in between these two extremes and are consistent with theoretical work on optimal exploration<sup>18,20,21</sup> and intrinsic motivation<sup>19,24</sup>. In this view, exploration and exploitation are entangled: uncertainty is given a value that can be compared to and added to the value of primary rewards<sup>18,20,21,32</sup>. Our findings further suggest that motivation driven by expected uncertainty may be sufficient to explain exploration in unstable environments. This contrasts with neuroeconomics studies, where expected uncertainty is defined as risk and corresponds to the exploitation of the irreducible variability of the outcomes, whereas exploration is driven only by reducible uncertainty<sup>23,44</sup>. However, assigning a given choice to exploration or exploitation is tricky in non-human animals. It relies on phenomenological models of behavior in the absence of direct reports of decision strategies. Thus, our data clearly show that VTA  $\beta 2^*$ -nAChRs affect motivation from expected uncertainty in both stable and unstable environments, but whether this corresponds to motivation to explore, exploit or both remains unclear. Nevertheless, our results are consistent with a causal role for the VTA in decisions under uncertainty via a common currency (a motivational metric) that integrates at least the values of both expected reward and expected uncertainty<sup>20,21,25,32</sup>.

DA neurons not only encode reward prediction errors<sup>9,10</sup>, but also surprise<sup>45</sup>, risk<sup>25</sup> (that is, expected uncertainty) and resolution of uncertainty<sup>46</sup>, which are all linked to information. DA neuron bursting related to reward is thought to constitute a teaching signal for actions<sup>11</sup> or an incentive signal<sup>12</sup> biasing the ongoing behavior. DA activity not related to expected rewards *per se* could also act as

an ‘intrinsic’ reinforcement signal<sup>24</sup> (or an intrinsic incentive) for which gathering information would be self-satisfactory, helping the animal to better predict its environment<sup>19,45</sup>. ACh is closely related to information processing<sup>1,2</sup>. We found that the cholinergic control of DA could underpin the motivational properties of information. This finding could explain the observed similarities when ACh or DA are pharmacologically manipulated during value-based decisions<sup>5</sup>. Several mechanisms may underlie functional ACh-DA interactions in the brain. Mesopontine ACh might directly signal expected uncertainty ( $\sigma^2$  in our model), as proposed for forebrain ACh<sup>22</sup>. Alternatively, our data suggest a contribution of  $\beta 2^*$ -nAChRs to the spontaneous excitability of DA neurons, with anesthetized  $\beta 2\text{KO}$  animals lacking bursting of DA neurons<sup>14</sup>. In this case, cholinergic signaling onto the VTA via  $\beta 2^*$ -nAChRs could serve as a permissive gate<sup>15</sup>, rendering DA neurons more responsive (that is, affecting  $\phi$ ) to uncertainty signals generated elsewhere in the mesocorticolimbic loop<sup>23</sup>. A strong prediction of these interpretations would be that ACh is implicated in the encoding of expected uncertainty by DA neurons<sup>25</sup>. Nevertheless, we cannot totally exclude, with our lentiviral strategy, downstream adaptations in  $\beta 2\text{KO}$  mice or an effect at the level of axon terminals, where  $\beta 2^*$ -nAChRs also influence the transfer function between DA firing and release in the striatum<sup>47</sup>.

Nicotine hijacks endogenous cholinergic signaling by exerting its reinforcing effects through  $\beta 2^*$ -nAChRs in the VTA<sup>13</sup>. But nicotine also affects decisions that are not related to nicotine intake itself. Smokers are actually known to display alterations of the exploration-exploitation tradeoff<sup>48</sup> and of risk-sensitivity<sup>49</sup>. Notably, tobacco addiction and pathological gambling, which can be seen as excessive uncertainty-seeking<sup>25,30</sup>, display a high comorbidity<sup>8</sup>. Thus, we suggest that  $\beta 2^*$ -nAChRs in the VTA, in addition to mediating reinforcement to nicotine, constitute a key neural component in the alterations of decision-making under uncertainty observed in nicotine addicts<sup>48,49</sup>.

## METHODS

Methods and any associated references are available in the [online version of the paper](#).

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## ACKNOWLEDGMENTS

We thank E. Guigon for discussions, C. Prévost-Solié for technical support, and J.-P. Changeux, E. Ey, G. Dugué and A. Boo for comments on the manuscript. This work was supported by the Centre National de la Recherche Scientifique CNRS UMR 8246, the University Pierre et Marie Curie (Programme Emergence Blanc 2012 for J.N. and P.F.), the Agence Nationale pour la Recherche (ANR Programme Blanc 2012 for P.F., ANR JCJC to A.M.), the Neuropole de Recherche Francilien (NeRF) of Ile de France, the Foundation for Medical Research (FRM, Equipe FRM DEQ2013326488 to P.F.), the Bettencourt Schueller Foundation (Coup d’Élan 2012 to P.F.), the Ecole des Neurosciences de Paris (ENP) to P.F., the Fondation pour la Recherche sur le Cerveau (FRC et les rotariens de France, “espoir en tête” 2012) to P.F. and the Brain & Behavior Research Foundation for a NARSAD Young Investigator Grant to A.M. The laboratories of P.F. and U.M. are part of the École des Neurosciences de Paris Ile-de-France RTRA network. P.F. and U.M. are members of the Laboratory of Excellence, LabEx Bio-Psy, and P.F. is member of the DHU Pepsy.

## AUTHOR CONTRIBUTIONS

J.N. and P.F. designed the study. S.T. and J.N. performed the virus injections. M.D., N.T., G.R. and J.N. performed the behavioral experiments. S.V. and F.M. performed the electrophysiological recordings. S.P. and U.M. provided the genetic tools. J.N., F.M. and P.F. analyzed the data. J.N., A.M. and P.F. wrote the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Everitt, B.J. & Robbins, T.W. Central cholinergic systems and cognition. *Annu. Rev. Psychol.* **48**, 649–684 (1997).
2. Dani, J.A. & Bertrand, D. Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annu. Rev. Pharmacol. Toxicol.* **47**, 699–729 (2007).
3. Guillem, K. *et al.* Nicotinic acetylcholine receptor  $\beta 2$  subunits in the medial prefrontal cortex control attention. *Science* **333**, 888–891 (2011).
4. Rangel, A., Camerer, C. & Montague, P.R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
5. Fobbs, W.C. & Mizumori, S.J. Cost-benefit decision circuitry: proposed modulatory role for acetylcholine. *Prog. Mol. Biol. Transl. Sci.* **122**, 233–261 (2014).
6. Kolokotroni, K.Z., Rodgers, R.J. & Harrison, A.A. Acute nicotine increases both impulsive choice and behavioral disinhibition in rats. *Psychopharmacology (Berl.)* **217**, 455–473 (2011).
7. Mendez, I.A., Gilbert, R.J., Bizon, J.L. & Setlow, B. Effects of acute administration of nicotinic and muscarinic cholinergic agonists and antagonists on performance in different cost-benefit decision making tasks in rats. *Psychopharmacology (Berl.)* **224**, 489–499 (2012).
8. McGrath, D.S. & Barrett, S.P. The comorbidity of tobacco smoking and gambling: a review of the literature. *Drug Alcohol Rev.* **28**, 676–681 (2009).
9. Schultz, W. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).
10. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
11. Montague, P.R., Dayan, P. & Sejnowski, T.J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
12. Berridge, K.C. From prediction error to incentive salience: mesolimbic computation of reward motivation. *Eur. J. Neurosci.* **35**, 1124–1143 (2012).
13. Maskos, U. *et al.* Nicotine reinforcement and cognition restored by targeted expression of nicotinic receptors. *Nature* **436**, 103–107 (2005).
14. Mameli-Engvall, M. *et al.* Hierarchical control of dopamine neuron-firing patterns by nicotinic receptors. *Neuron* **50**, 911–921 (2006).
15. Grace, A.A., Floresco, S.B., Goto, Y. & Lodge, D.J. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* **30**, 220–227 (2007).
16. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B. & Dolan, R.J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
17. Frank, M.J., Doll, B.B., Oas-Terpstra, J. & Moreno, F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat. Neurosci.* **12**, 1062–1068 (2009).
18. Gittins, J.C. & Jones, D.M. A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika* **66**, 561–565 (1979).
19. Scott, P.D. & Markovitch, S. Learning novel domains through curiosity and conjecture. *IJCAI (US)* **1**, 669–674 (1989).
20. Kaelbling, L.P. *Learning in Embedded Systems* (MIT Press, 1993).
21. Meuleau, N. & Bourgin, P. Exploration of multi-state environments: Local measures and back-propagation of uncertainty. *Mach. Learn.* **35**, 117–154 (1999).
22. Yu, A.J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
23. Bach, D.R. & Dolan, R.J. Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat. Rev. Neurosci.* **13**, 572–586 (2012).
24. Oudeyer, P.-Y. & Kaplan, F. What is intrinsic motivation? A typology of computational approaches. *Front. Neurobot.* **1**, 6 (2007).
25. Fiorillo, C.D., Tobler, P.N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
26. Schuck-Paim, C., Pompilio, L. & Kacelnik, A. State-dependent decisions cause apparent violations of rationality in animal choice. *PLoS Biol.* **2**, e402 (2004).
27. Carlezon, W.A. Jr. & Chartoff, E.H. Intracranial self-stimulation (ICSS) in rodents to study the neurobiology of motivation. *Nat. Protoc.* **2**, 2987–2995 (2007).
28. Kobayashi, T., Nishijo, H., Fukuda, M., Bureš, J. & Ono, T. Task-dependent representations in rat hippocampal place neurons. *J. Neurophysiol.* **78**, 597–613 (1997).
29. Funamizu, A., Ito, M., Doya, K., Kanzaki, R. & Takahashi, H. Uncertainty in action-value estimation affects both action choice and learning rate of the choice behaviors of rats. *Eur. J. Neurosci.* **35**, 1180–1189 (2012).
30. Anselme, P., Robinson, M.J.F. & Berridge, K.C. Reward uncertainty enhances incentive attribution as sign-tracking. *Behav. Brain Res.* **238**, 53–61 (2013).
31. Sutton, R.S. & Barto, A.G. *Reinforcement Learning: an introduction* (MIT Press, 1998).
32. Kakade, S. & Dayan, P. Dopamine: generalization and bonuses. *Neural Netw.* **15**, 549–559 (2002).
33. Herrnstein, R.J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4**, 267–272 (1961).
34. Ishii, S., Yoshida, W. & Yoshimoto, J. Control of exploitation-exploration meta-parameter in reinforcement learning. *Neural Netw.* **15**, 665–687 (2002).
35. Yeomans, J. & Baptista, M. Both nicotinic and muscarinic receptors in ventral tegmental area contribute to brain-stimulation reward. *Pharmacol. Biochem. Behav.* **57**, 915–921 (1997).
36. Serreau, P., Chabouat, J., Suarez, S.V., Naudé, J. & Granon, S. Beta2-containing neuronal nicotinic receptors as major actors in the flexible choice between conflicting motivations. *Behav. Brain Res.* **225**, 151–159 (2011).
37. Krugel, L.K., Biele, G., Mohr, P.N., Li, S.-C. & Heekeren, H.R. Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci. USA* **106**, 17951–17956 (2009).
38. Niv, Y., Edlund, J.A., Dayan, P. & O'Doherty, J.P. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J. Neurosci.* **32**, 551–562 (2012).
39. Balasubramani, P.P., Chakravarthy, V.S., Ravindran, B. & Moustafa, A.A. An extended reinforcement learning model of basal ganglia to understand the contributions of serotonin and dopamine in risk-based decision making, reward prediction, and punishment learning. *Front. Comput. Neurosci.* **8**, 47 (2014).
40. Granon, S., Faure, P. & Changeux, J.-P. Executive and social behaviors under nicotinic receptor regulation. *Proc. Natl. Acad. Sci. USA* **100**, 9596–9601 (2003).
41. Picciotto, M.R. *et al.* Abnormal avoidance learning in mice lacking functional high-affinity nicotine receptor in the brain. *Nature* **374**, 65–67 (1995).
42. Maubourguet, N., Lesne, A., Changeux, J.-P., Maskos, U. & Faure, P. Behavioral sequence analysis reveals a novel role for  $\beta 2^*$  nicotinic receptors in exploration. *PLoS Comput. Biol.* **4**, e1000229 (2008).
43. Gordon, G., Fonio, E. & Ahissar, E. Emergent exploration via novelty management. *J. Neurosci.* **34**, 12646–12661 (2014).
44. Payzan-LeNestour, E. & Bossaerts, P. Risk, unexpected uncertainty and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Comput. Biol.* **7**, e1001048 (2011).
45. Redgrave, P. & Gurney, K. The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* **7**, 967–975 (2006).
46. Bromberg-Martin, E.S. & Hikosaka, O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119–126 (2009).
47. Rice, M.E. & Cragg, S.J. Nicotine amplifies reward-related dopamine signals in striatum. *Nat. Neurosci.* **7**, 583–584 (2004).
48. Addicott, M.A., Pearson, J.M., Wilson, J., Platt, M.L. & McClernon, F.J. Smoking and the bandit: a preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Exp. Clin. Psychopharmacol.* **21**, 66–73 (2013).
49. Galván, A. *et al.* Greater risk sensitivity of dorsolateral prefrontal cortex in young smokers than in nonsmokers. *Psychopharmacology (Berl.)* **229**, 345–355 (2013).

## ONLINE METHODS

**Animals.** 40 male C57BL/6J (WT) mice and 47 male knockout SOPF HO ACNB2 ( $\beta$ 2KO) mice obtained from Charles Rivers Laboratories France were used.  $\beta$ 2KO mice were generated as described previously<sup>41</sup>. WT and  $\beta$ 2KO mice are not littermates and this could be a potential caveat of the study. However, mutant mice were generated almost 20 years ago, the line has been backcrossed more than 20 generations with the wild-type C57BL/6J line, and the  $\beta$ 2KO line was confirmed to be at more than 99.99% C57BL/6J. Mice arrived to the animal facility at 8 weeks of age, and were housed individually for at least 2 weeks before the electrode implantation. Behavioral tasks started 1 week after implantation to insure full recovery. Intracranial self-stimulation (ICSS) does not require food deprivation; as a consequence all mice had ad libitum access to food and water except during behavioral sessions. The temperature (20–22 °C) and humidity was automatically controlled and a circadian light cycle of 12/12-h light-dark cycle (lights on at 8:30 a.m.) was maintained in the animal facility. All experiments were performed during the light cycle, between 09:00 a.m. and 5:00 p.m. Experiments were conducted at Université Pierre et Marie Curie. All procedures were performed in accordance with the recommendations for animal experiments issued by the European Commission directives 219/1990 and 220/1990, and approved by Université Pierre et Marie Curie.

**Stereotaxic injection of lentivirus.** The lentiviral expression vectors  $\beta$ 2 subunit-IRES-eGFP cDNAs and the eGFP cDNA (control) are under the control of the ubiquitous mouse phosphoglycerate kinase (PGK) promoter. Further details can be found in ref. 13.  $\beta$ 2KO mice aged 8 weeks were anesthetized using isoflurane. The mouse was introduced into a stereotaxic frame adapted for mice. Lentivirus (2  $\mu$ l at 75 ng of p24 protein per  $\mu$ l) was injected bilaterally at: anteroposterior = -3.4 mm, mediolateral =  $\pm$ 0.5 mm from bregma and dorsoventral = 4.4 mm from the surface for VTA injection. Mice were implanted with electrodes 4–5 weeks after viral injection. At the end of the behavioral experiments, lentiviral re-expression in the VTA was verified using fluorescence immunohistochemistry. As a control for  $\beta$ 2VEC mice, another group of  $\beta$ 2KO mice were injected with lentivirus expressing eGFP only. We did not observe any difference between  $\beta$ 2KO (without lentiviral injections,  $n = 6$ ) and  $\beta$ 2-eGFP mice ( $n = 6$ ) in either choices ( $P = 0.76$ , unpaired  $t$  test) and time-to-goal ( $P = 0.34$ , unpaired  $t$  test). We thus pooled the data from both groups to serve as control for  $\beta$ 2VEC data.

**In vivo electrophysiological recordings.** Extracellular recording electrodes were constructed from borosilicate glass tubing (1.5 mm O.D. / 1.17 mm I.D.) using a vertical electrode puller (Narishige). Tip was broken and electrodes were filled with a 0.5% sodium acetate solution (wt/vol) and 1.5% neurobiotin (wt/vol), yielding impedances of 6.9 M $\Omega$ .

Animals were anesthetized with chloral hydrate (400 mg per kg of body weight, intraperitoneal, supplemented as required to maintain optimal anesthesia throughout the experiment), and placed in a stereotaxic apparatus (Kopf Instruments). The left saphenous vein was catheterized for intravenous administration of nicotine and the right saphenous vein was catheterized for intravenous administration of saline solution (NaCl 0.9%, wt/vol). The electrophysiological activity was sampled in the central region of the VTA (coordinates: 3.1–3.5 mm posterior to Bregma,  $\pm$ 0.3–0.6 mm lateral to midline and 4–4.7 mm below the brain surface)<sup>50</sup>. Spontaneously active DAergic neurons were identified on the basis of previously established electrophysiological criteria: (1) a typical triphasic action potential with a marked negative deflection; (2) a characteristic long duration ( $>2.0$  ms); (3) an action potential width from start to negative trough  $>1.1$  ms; (4) a slow firing rate (between 1 and 10 Hz) with an irregular single spiking pattern and occasional short, slow bursting activity<sup>51</sup>. At least 5 min of spontaneous baseline electrophysiological activity was recorded before intravenous injection of nicotine (30  $\mu$ g per kg). At the end of the recording period, the neurons were stimulated by application of positive currents steps to electroporate neurobiotin into the neurons to allow DA neurons identification.

**Analysis of electrophysiological data.** DA cell firing was analyzed with respect to the average firing rate and the percentage of spikes within bursts (%SWB, number of spikes within bursts, divided by total number of spikes). Bursts were identified as discrete events consisting of a sequence of spikes such that: their onset is defined by two consecutive spikes within an interval  $<80$  ms and they terminated with an

interval  $>160$  ms (ref. 51). Firing rate and %SWB were evaluated on successive windows of 60 s, with a 45-s overlapping period<sup>14</sup>. For each cell, firing frequency was rescaled as a percentage of its baseline value averaged during the 2 min before nicotine injection. The effect of nicotine was assessed as a comparison between the maximum of variation of firing rate and %SWB observed during the first 3 min after saline and nicotine injection. The results are presented as mean  $\pm$  s.e.m. of the difference of maximal variation before and after nicotine.

**Fluorescence immunohistochemistry.** Following the death of all the lentivirus-injected mice (GFP and VEC animals), brains were rapidly removed and fixed in 4% paraformaldehyde. Following a period of at least 3 d of fixation at 4 °C, serial 60- $\mu$ m sections were cut from the midbrain with vibratome. Immunohistochemistry was performed as follows: free-floating VTA brain sections were incubated 1 h at 4 °C in a blocking solution of phosphate-buffered saline (PBS) containing 3% Bovine Serum Albumin (BSA, Sigma; A4503) (vol/vol) and 0.2% Triton X-100 (vol/vol) and then incubated overnight at 4 °C with a mouse anti-tyrosine hydroxylase antibody (TH, Sigma, T1299) at 1:200 dilution and a rabbit anti-GFP antibody (Molecular Probes, A-6455) at 1:5,000 dilution in PBS containing 1.5% BSA and 0.2% Triton X-100. The following day, sections were rinsed with PBS and then incubated 3 h at 22–25 °C with Cy3-conjugated anti-mouse and Cy2-conjugated anti-rabbit secondary antibodies (Jackson ImmunoResearch, 715-165-150 and 711-225-152) at 1:200 dilution in a solution of 1.5% BSA in PBS. After three rinses in PBS, slices were wet-mounted using Prolong Gold Antifade Reagent (Invitrogen, P36930). Microscopy was carried out with a fluorescent microscope, and images captured using a camera and ImageJ imaging software.

In the case of electrophysiological recordings, an immunohistochemical identification of the recorded neurons was performed as described above, with the addition of 1:200 AMCA-conjugated Streptavidin (Jackson ImmunoResearch) in the solution. Neurons labeled for both TH and neurobiotin in the VTA<sup>50</sup> allowed to confirm their neurochemical phenotype.

**Electrode implantation and ICSS training.** Mice were introduced into a stereotaxic frame and implanted unilaterally with bipolar stimulating electrodes for ICSS in the medial forebrain bundle<sup>27,28</sup> (MFB, anteroposterior = 1.4 mm, mediolateral =  $\pm$ 1.2 mm, from the bregma, and dorsoventral = 4.8 mm from the dura). After recovery from surgery (1 week), the efficacy of electrical stimulation was verified in an open field with an explicit square location (side = 1 cm) at its center. Each time a mouse was detected in the area ( $D = 3$  cm) of the location, a 200-ms train of 20 0.5-ms biphasic square waves pulsed at 100 Hz was generated by a stimulator<sup>28</sup>. Mice self-stimulating at least 50 times in a 5-min session were kept for the behavioral sessions (3 mice were excluded at this stage, due to improper electrode implantation). In the certain setting (see below), ICSS intensity was adjusted so that mice self-stimulated between 50 and 150 times per session at the end of the training (ninth and tenth session). Current intensity was subsequently maintained the same throughout the uncertainty setting.

**Behavioral data acquisition.** Decision-making and locomotor activity were recorded in a 1-m diameter circular open-field. Experiments were performed using a video camera, connected to a video-track system, out of sight of the experimenter. A home-made software (Labview National instrument) tracked the animal, recorded its trajectory (20 frames per s) for 5 min and sent TTL pulses to the ICSS stimulator when appropriate (see below).

**Markovian decision problem by ICSS conditioning.** We considered two complementary aspects of motivation: direction and locomotion of the mice. We thus developed a protocol allowing to record simultaneously the sequential choices between differently rewarding locations (that is, associated with intracranial self-stimulation) and the locomotor activity of the mice in between these locations. After validation of ICSS behavior<sup>27</sup>, conditioning tasks took place in the 0.8-m diameter circular open-field. Three explicit square locations were placed in the open field, forming an equilateral triangle (side = 50 cm). Each time a mouse was detected in the area of one of the rewarding locations, a stimulation train was delivered. Animals received stimulations only when they alternate between rewarding locations. In separate experiments, the intensity or the probability of stimulation delivery differed for the three rewarding locations. Precise parameters (for example, reward probabilities) were pseudo-randomly assigned to each

rewarding location for each mouse. For each set of (consecutive) experiments conditioning consisted in one daily session of 5 min, during 10 d. Decision-making was analyzed by expressing data as a series of choices between rewarding locations (labeled A, B, C). We only considered choices made in an interval of 10s after visiting the previous rewarding location. This restriction is based on the observation that choices made after 10 s were random (that is, uniformly distributed) for every condition, and thus probably reflect a disengagement from the task. This led to the exclusion of fewer than 3% of the total choices made by the animals (all groups), which suggests that incorporating these late choices would not significantly change the results. This game implements a Markovian Decision Process (MDP<sup>31</sup>) consisting of three states (A, B, C), corresponding to each rewarding locations, and a transition function, corresponding to the proportions of choices in the three gambles. The repartition is defined as the proportion of states visited by the animal during a session. The transition matrix describes the proportion of transitions from one state to another. Because animals receive stimulations only when they alternate between rewarding locations, there is no repetition of states in the sequence and the  $3 \times 3$  transition matrix has null diagonal elements. The training consisted of a block (10 daily sessions of 5 min) in a CS where all locations were associated with an ICSS delivery. The test phase consisted of a block (10 daily sessions of 5 min) assessing choice organization under an US, by associating each location with a different probability of ICSS, a validated protocol for studying risky choices<sup>52</sup>. The foraging phase was performed after the uncertain setting, and five supplementary sessions of deterministic setting. The foraging phase assessed the exploratory strategy in a dynamic setting (DS), which consisted in three consecutive 5-min sessions. In each session, two out of three locations delivered the ICSS reward, and the identity of the two rewarding locations changed every session.

**Analysis of locomotion.** Locomotor activity toward the rewarding locations was measured in terms of time-to-goal, speed profile, dwell time and traveled distance. Time-to-goal measures the duration between one choice and the next one. The speed profile corresponds to the instantaneous speed as a function of time (expressing it as a function of the distance between two locations did not give any additional information). We averaged the speed profiles on a 10-s interval (the same used for restricting the choices considered in the analysis), which was zero-padded if the reward location was attained before 10 s. The dwell time is defined as the duration between the moment of the detection in the last rewarding location and the moment when the animal's speed is greater than  $10 \text{ cm s}^{-1}$ . The traveled distance corresponds to the summation of the local distances between two points of the mouse's trajectory (20 frames per s) between the last and the next choice. A multiple linear regression was performed on the time-to-goal, in the different sets of probabilities of the US setting. We compared models with increasing number of explanatory variables. As potential explanatory variables, we included reward history (whether the animal just got rewarded or not, as a binary variable), the expected reward of the goal, the expected uncertainty of the goal, the expected reward of the alternative (that is, the location not chosen in the gamble), and the expected uncertainty of the alternative. We compared these linear models based on their summed squared errors, penalized for complexity (Bayesian information criterion):  $BIC(TTG) = n \ln(SSE/n) + k \ln(n)$ , where  $n$  is the number of observations (time-to-goal,  $n$  is the same for all regressions),  $k$  the number of explanatory variables, and  $SSE$  the summed squared errors from the multiple linear regressions. Constant terms were omitted from the formula for simplicity, as the BICs of the linear regressions were only used for comparisons.

**Computational models of decision-making.** In the US, we investigated how well the transition function (that is, choices) from both genotypes can fit to variants of decision-making models. At the end of the US, since mice are trained and choice behavior is at steady-state, we only modeled decision-making, and used the settings of the task (that is, reward probabilities) as fixed parameters for the values of the options (see below). In the DS and in the learning phase of the US, we modeled both learning (see below) and decision-making, and we evaluated how well the models fits the animals' choices, which were not at steady-state. These models are thus based on the estimation of the expected payoffs ("value") and uncertainties of the options, rather than on objective parameters of the task.

Decision-making models determined the probability  $P_i$  of choosing the next state  $i$ , as a function (the choice rule) of a decision variable. Because mice could not return to the same rewarding location, they had to choose between

the two remaining locations. Accordingly, we modeled decisions between two alternatives. We considered five choice rules<sup>31</sup>: local matching law<sup>33</sup> Herrnstein, softmax, epsilon-greedy, uncertainty bonus<sup>19,21,39</sup> and uncertainty-controlled randomness<sup>34</sup>.

- In the local matching law, the probability to choose an action  $i$  (amongst two rewarding location) is given by

$$P_i = \frac{V_i}{\sum_j V_j} \quad (1)$$

where  $V_i$  is defined as the value of an option, that is, the expected reward (see below).

- The epsilon-greedy choice rule is

$$\begin{cases} P_i = 1 - \epsilon & i = \text{argmax}(V_i) \\ \epsilon & \text{otherwise} \end{cases} \quad (2)$$

where  $\epsilon$  is the probability of choosing less valuable options, reflecting undirected exploration.

- The softmax choice rule is

$$P_i = \frac{\exp(\beta V_i)}{\sum_j \exp(\beta V_j)} \quad (3)$$

where  $\beta$  is an inverse temperature parameter reflecting the sensitivity of choice to the difference between decision variables.

In standard reinforcement learning<sup>31</sup>, the value of an option is the expected (average) reward. In the US, where the choices are at steady-state, the expected reward is taken as the reward probability

$$V_i = E_i(\text{ICSS}) = p_i(\text{ICSS}) \quad (4)$$

In models embedding an exploration bonus, the value depends on both expected reward and uncertainty<sup>16,17,29</sup>. Uncertainty may refer to estimation uncertainty (due to incomplete knowledge or sampling of the outcome), to the expected uncertainty (or reward risk), related to the estimated variability of the outcome, or to unexpressed uncertainty, that is, uncertainty greater than expected<sup>22,23,44</sup>. The expected uncertainty scheme is similar to the mean-variance approach used in neuroeconomic studies<sup>53</sup> and it has also been proposed to drive exploration<sup>19,21,24,25,30</sup>. In the US, as mice are trained and choice behavior is at steady-state, we used this version of the model, where the decision variable is a compound of the true (that is, not estimated by a learning algorithm) mean and variance of the payoff

$$V_i = E_i(\text{ICSS}) + \phi \sigma_i^2(\text{ICSS}) = p_i(\text{ICSS}) + \phi p_i(\text{ICSS})(1 - p_i(\text{ICSS})) \quad (5)$$

This compound value is then nested in the softmax choice rule. Note that expected uncertainty ( $\sigma_i^2$ ) can also be estimated through learning (see equation (10)).

Finally, in the uncertainty-based temperature model (or local control of randomness<sup>34</sup>), uncertainty associated with all the possible actions at a state controls the randomness of choices (that is, the temperature parameter). In this strategy, the randomness of action selection does not depend on the variability of the possible outcomes. In the softmax model (equation (3)), in case where different choices may yield comparable outcomes, the decision process is random even with large  $\beta$ ; while a large difference in values results in greedy action selection even for small  $\beta$ . To circumvent this issue, it is possible to normalize the temperature parameter  $\beta_i$  for each state  $i$ .

$$\beta_i = \frac{\beta_0}{\sqrt{E(V_j^2) - E(V_j)^2}} \quad (6)$$

where  $\beta_0$  is a constant (free) parameter, whereas  $\sqrt{E(V_j^2) - E(V_j)^2}$  represents the uncertainty (or variability) of the state  $i$  (over all the possible actions  $j$ ) rather than reward uncertainty associated with a particular action.

Reinforcement learning models determined the evolution of the decision variables, which are in this case estimations of the task parameters. The values of

the rewarding locations were estimated using standard reinforcement models<sup>31</sup>, which are based on trial-and-error learning. First, the model computes the discrepancy between the predicted value of the chosen location ( $V_i$ ) and the actual reward  $R$  at the trial  $t$

$$\delta_i(t) = R_{i,t} - V_{i,t-1} \quad (7)$$

where  $R_i(t) = 1$  or 0 depending on whether the animal was rewarded or not. This reward prediction error is then used to adapt the estimation of the value  $V_i$  of the chosen location only, that is, the values of the other locations are not changed

$$V_{i,t} = V_{i,t-1} + \alpha \delta_{i,t} \quad (8)$$

where  $\alpha$  is the learning rate. To test whether nicotinic receptors differentially affected the sensitivity to reward and reward omission, we used an asymmetric version of reinforcement learning<sup>38</sup>

$$\begin{cases} V_{i,t} = V_{i,t-1} + \alpha^+ \delta_{i,t} & \delta_{i,t} > 0 \\ V_{i,t} = V_{i,t-1} + \alpha^- \delta_{i,t} & \delta_{i,t} < 0 \end{cases} \quad (9)$$

where  $\alpha^+$  and  $\alpha^-$  are the learning rates for better- or worse-than-expected outcomes.

We also used an extended version of reinforcement-learning model<sup>23,39</sup> to evaluate the expected uncertainty of the rewarding locations. The rationale behind this model is that uncertain and unpredictable outcomes produce large prediction errors (positive and negative), by definition. Hence squared prediction errors (equation (7)) can be used to estimate unpredictability or uncertainty  $\sigma_{i,t}^2$

$$\sigma_{i,t}^2 = \sigma_{i,t-1}^2 + \alpha^\varphi \xi_{i,t} \quad (10)$$

where  $\alpha^\varphi$  is the learning rate for uncertainty, and  $\xi_{i,t}$  is the uncertainty (or risk) prediction error of the option  $i$  at trial  $t$ , that is,

$$\xi_{i,t} = \delta_{i,t}^2 - \sigma_{i,t-1}^2 \quad (11)$$

The uncertainty prediction error corresponds to unexpected uncertainty (uncertainty larger than expected) and we tested whether exploration might be directed by unexpected form of uncertainty, by assigning a bonus to this error term

$$V_{i,t}^* = V_{i,t} + \xi_{i,t} \quad (12)$$

Finally, uncertainty may exert an indirect effect through learning. It has been shown in humans that learning rate itself can increase with sudden changes in uncertainty<sup>37,54</sup>. We tested the following adaptive learning rate model<sup>37</sup>, where learning rate increases when there is a recent increase  $m$  in absolute prediction errors

$$\begin{cases} \alpha_t = \alpha_{t-1} + f(m_t)(1 - \alpha_{t-1}) & m_t > 0 \\ \alpha_t = \alpha_{t-1} + f(m_t)\alpha_{t-1} & m_t < 0 \end{cases} \quad (13)$$

where  $f(m)$  is a double sigmoid function  $f(m_t) = \text{sign}(m_t)(1 - e^{-(m/\lambda)^2})$  where  $m$  is the slope of the (recent) smoothed absolute reward prediction errors,

$m_t = 2 \frac{\delta_t^{abs} - \delta_{t-1}^{abs}}{\delta_t^{abs} + \delta_{t-1}^{abs}}$ . Smoothing of absolute prediction errors is achieved by

$\delta_t^{abs} = \delta_{t-1}^{abs}(1 - \alpha_1) + |\delta_t| \alpha_1$ . The free parameter  $\lambda$  determines the degree to which uncertainty (absolute prediction errors) affects the learning rate, and the other free parameter,  $\alpha_1$ , determines the initial learning rate and the speed of  $\delta_t^{abs}$  updating.

In the US, at steady-state, we fitted the free parameters of the four decision-making models (none for the matching law,  $\epsilon$  for  $\epsilon$ -greedy,  $\beta$  for softmax,  $\beta$  and  $\varphi$  for uncertainty model). In the learning phase of the US, we fitted the free parameters of these 4 models: standard RL ( $\alpha, \beta$ ), RL with uncertainty learning and expected uncertainty bonus ( $\alpha, \beta, \alpha^\varphi, \varphi$ ), RL with adaptive (uncertainty-dependent) learning rate ( $\alpha, \beta, \lambda$ ), and RL with uncertainty learning and unexpected uncertainty bonus ( $\alpha, \beta, \alpha^\varphi, \varphi$ ). We fixed the initial conditions ( $V(0) = 1$ , and  $\sigma(0) = 0$ ), because the mice underwent the certain setting just beforehand.

In the DS, we fitted the free parameters and initial conditions of these 7 models: standard RL ( $\alpha, \beta, V(0)$ ), asymmetric learning rates RL ( $\alpha^-, \alpha^+, \beta, V(0)$ ), RL with uncertainty bonus ( $\alpha, \beta, \varphi, V(0), \sigma(0)$ ), RL with separate learning for value and uncertainty ( $\alpha, \alpha^\varphi, \beta, \varphi, V(0), \sigma(0)$ ), RL with asymmetric learning rates learning for value and separate uncertainty learning ( $\alpha^-, \alpha^+, \alpha^\varphi, \beta, \varphi, V(0), \sigma(0)$ ), RL with uncertainty learning and unexpected uncertainty bonus ( $\alpha, \alpha^\varphi, \beta, \varphi, V(0), \sigma(0)$ ), RL with adaptive (uncertainty-dependent) learning rate ( $\alpha, \beta, \lambda, V(0)$ ).

In each case, we searched for the free parameters maximizing the respective likelihood of the observed choices  $c$  at all trials  $t$  ( $\prod_c P_{c,t}$ ). We performed the

fits of all the parameters individually for each animal  $a$ , using the population fit (that is, fit of the average probabilities of choices) as initial conditions. We checked that the mean of individual fits stayed close to the population fit, and that the optima was non-local (by examining the Hessian matrix<sup>55</sup>). We used the *fmincon* function in Matlab to perform the fits, with the constraints that learning rates and temperature could not be negative and that learning rates could not exceed 1. To assess goodness-of-fit, we report negative log likelihoods penalized for model complexity (Bayesian information criterion; BIC). Smaller BIC values indicate a better fit. Each of these models has been found to fit experimental data in at least one given experimental condition (for example, behavioral task or species<sup>16,17,29,38,39</sup>). Here, we aimed at accounting for the difference observed between genotypes, to propose a computational role for the nicotinic modulation of the VTA. Hence, once the best model is determined, possible differences in the free parameters (for example,  $\epsilon, \beta, \varphi$ ) between genotypes or conditions point at the computational role of the  $\beta 2$  subunit-containing nAChRs expressed in the VTA in decision-making processes.

#### Extension of the uncertainty model to previous experiments on $\beta 2$ KO mice.

We also aimed at extending our framework by modeling the results from previous studies focusing on the behavioral differences between WT and  $\beta 2$ KO mice with reinforcement learning models embedding an uncertainty-based exploratory bonus (equations (5, 7, 8, 10 and 11)). In these experiments, uncertainty was not explicitly controlled but was yet present, as in most decision tasks. We thus used the main difference found in the model-based analysis of our decision task, that is, a positive value given to uncertainty in WT, but not to  $\beta 2$ KO, mice, and explored the values for uncertainty estimation to qualitatively match the data. All experiments were modeled as MDPs with a discretization of the relevant states for the animals.

In the open-field experiment<sup>13,42</sup>, we used the symbolic decomposition of the behavior proposed in ref. 42, by splitting the locomotion of the mice into "active" versus "inactive" states, and their positions into "center" and "periphery" states. The active state corresponds to high-speed navigation, while the inactive state corresponds to low-speed exploration, mainly composed of rearing, scanning and sniffing behaviors<sup>42,43</sup>. This double dichotomy gives rise to four states, that we modeled as an MDP with all transitions possible, except for the "stay" transitions (that is, of one state on itself) and the transitions between periphery-inactive (PI) and center-inactive (CI) states, which were not found in the data<sup>13,42</sup>. The duration of one state was 1 s. We modeled the difference between WT and  $\beta 2$ KO mice by adding an exploratory bonus to the inactive states in WT mice only, which we deduced from the experimental (average) transition probability and the softmax decision rule with bonus as follows. In the center-active state, the probability of going the center-inactive state is given by  $P(CI|CA) = e^{\beta V_{CI}} / (e^{\beta V_{CI}} + e^{\beta V_{PA}})$ , where  $V_{CI}$  and  $V_{PA}$  represent the values associated with the center-inactive and the periphery-active states, so we computed the relations between  $V_{PA}$  and  $V_{CI}$ , and between  $V_{PI}$  and  $V_{CA}$ , and fitted  $\beta, V_{PI}$  and  $V_{CI}$  to reproduce the data.

In the object recognition task<sup>40</sup>, two objects are placed in an open-field, and the time spent in the objects area is measured as a function of the behavioral sessions. We modeled this task as an MDP using a discretization of space, consisting in 25 states corresponding to the open-field without objects, and two states corresponding to the objects. The duration of one state was 1 s. We used the uncertainty model (no reward being present in the task, we modeled the uncertainties but not the values) and we fitted the values of  $\alpha, \beta, \varphi$ , and the initial uncertainties of the objects and of the open-field to reproduce the data.

In the spatial maze<sup>40</sup>, we modeled an idealized version of this conditioning task, consisting of four states, corresponding to the arms of the maze. One of them delivered a reachable food reward ( $R = 1$  if reached), and was absorbing:

the simulation stopped if the agent (the modeled mouse) reached it. The duration of one state (the mean duration of visiting one arm) was 10 s. We used the uncertainty model with a single learning rate ( $\alpha$ ,  $\beta$ ,  $\phi$ ,  $\zeta(0)$ ,  $\sigma(0)$ ) for simplicity. We simulated the model until the food was reached, and measured the time to reach the food, as done in the experiment.

In the passive avoidance task<sup>41</sup>, animals are in a box divided in two (light and dark) compartments. The learning phase (which was not modeled) consists in a single foot shock given in the dark compartment, which arguably induces a negative prediction error for this state. We simulated this experiment by considering a sequential evaluation model representing incentive motivation<sup>56</sup>, in which the agent sequentially evaluates the probability to go to the dark compartment until it decides to accept it. The probability to go to the dark part of the box at any time is given by

$$P(D) = \frac{1}{e^{-\beta(V_D - \theta)}} \quad (14)$$

where  $\beta$  is the inverse temperature (sensitivity to value) and  $\theta$  a threshold representing the basal locomotor activity of the animal. In this model, the agent evaluates the probability of going to the dark part, based on its single experience of a foot shock, which induced a single, negative, reward prediction error (equation (7)), resulting both in a decrease in value (equation (8)) and an increased uncertainty (equation (10)). The time-step for each evaluation was 1 s. We measured the time before the agents go to the dark part, as done in the experiment<sup>41</sup>. For each model experiment, standard errors were obtained following a bootstrap procedure, using the sample size of the original data.

**Statistical analysis.** No statistical methods were used to predetermine sample sizes. Our sample sizes are comparable to many studies using similar techniques and animal models. We used a pseudo-randomization procedure, in the sense that in the behavioral experiments, precise parameters (for example, reward probabilities) were pseudo-randomly assigned to each rewarding location for

each mouse. The experiments were blind, in the sense that the experimenters (both in behavioral and electrophysiological experiments) were not aware of which genotype each mouse belonged to.

Behavioral and model data were analyzed and fitted using Matlab (The MathWorks) Electrophysiological data was analyzed using R (The R Project). Code is available on request. Data are plotted as mean  $\pm$  s.e.m. Total number ( $n$ ) of observations in each group and statistics used are indicated in figure captions. Classically comparisons between means were performed using parametric tests (Student for two groups, or ANOVA for comparing more than two groups) when parameters followed a normal distribution (Shapiro test  $P > 0.05$ ), and non-parametric tests (here, Wilcoxon or Mann-Whitney) when this was not the case. Homogeneity of variances was tested preliminarily and the  $t$  tests were Welch-corrected if needed. Multiple comparisons were Bonferroni corrected. All statistical tests were two-sided.  $P > 0.05$  was considered to be not statistically significant.

A **Supplementary Methods Checklist** is available.

50. Paxinos, G. & Franklin, K.B. *The Mouse Brain in Stereotaxic Coordinates* (Gulf Professional Publishing, 2004).
51. Grace, A.A. & Bunney, B.S. Intracellular and extracellular electrophysiology of nigral dopaminergic neurons--1. Identification and characterization. *Neuroscience* **10**, 301–315 (1983).
52. Rokosik, S.L. & Napier, T.C. Intracranial self-stimulation as a positive reinforcer to study impulsivity in a probability discounting paradigm. *J. Neurosci. Methods* **198**, 260–269 (2011).
53. D'Acromont, M. & Bossaerts, P. Neurobiological studies of risk assessment: a comparison of expected utility and mean-variance approaches. *Cogn. Affect. Behav. Neurosci.* **8**, 363–374 (2008).
54. Behrens, T.E.J., Woolrich, M.W., Walton, M.E. & Rushworth, M.F.S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
55. Daw, N.D. Trial-by-trial data analysis using computational models. in *Decision Making, Affect, and Learning: Attention and Performance XXIII* (eds. Delgado, M.R., Phelps, E.A. & Robbins, T.W.) 3–38 (2011).
56. McClure, S.M., Daw, N.D. & Montague, P.R. A computational substrate for incentive salience. *Trends Neurosci.* **26**, 423–428 (2003).

# 8 Chronic nicotine enhances value sensitivity in mice through altered VTA DA activity

## 8.1 Reward sensitivity and nicotine

Normative theories and recent data report that the phasic dopamine signal provides a teaching signal necessary for learning, while the background or tonic dopamine tone and its fluctuations may underlie the expression of previously acquired behaviours [253]. Long-term nicotine can alter the spontaneous activity of VTA DA cells and therefore hypothetically change phasic and tonic signaling.

Neuroeconomical behavioural tasks can provide a measure for shifts in choice behaviour. In a probabilistic delay discounting task<sup>1</sup> both a high doses of acute and sub-chronic nicotine increased the preference for the immediate less probable reward [50], which is in conformity with simple delay discounting tasks in humans [254–256] and rats [257]. This suggests that nicotine enhances impulsivity and a slight preference for risk. In a probabilistic reversal learning task<sup>2</sup> saline and chronic nicotine rats performed similarly on reversals and response latency. Differences in cognitive flexibility were only observed during withdrawal. The authors did not report any differences in probability (value) preference. The limitation of this paradigm is that flexibility, effort, value evaluation and impulsivity are difficult to distinguish. Furthermore, the task was already pre-trained before mini-pump implantation [258] and (dopamine dependent) learned behaviour is more difficult to manipulate.

Recently, Lesage and colleagues (2017) did demonstrate a double dissociation between reward sensitivity, the ability to interpret valence information, and cognitive flexibility during abstinence and acute nicotine administration in smokers and non smokers. After

---

1. In an operant box, one lever was reinforced immediately, but with a lower probability, the other lever delivered a delayed reward with variable interval

2. In an operant box, two nose-poke holes are respectively reinforced 20% or 80% with a food pellet. The probabilities shifted every so many trials



nicotine administration smokers showed a lower lose-shift percentage, a higher win-stay percentage and perseverative errors, suggesting less cognitive flexibility. At the same time, acute nicotine in smokers increased the reward sensitivity, increasing the contrast between the available option values. Interestingly, smokers displayed less reward sensitivity activity in the striatum and dorsal anterior cingulate cortex than non-smokers, but enhanced cognitive flexibility contrasts in the striatum, anterior insula, dorsal anterior cingulate cortex and ventromedial prefrontal cortex compared to non-smokers [259]. This study shows a direct link between an altered mesocorticolimbic dopaminergic system and modified choice behaviour due to chronic nicotine, a finding that supports the work that will be presented in this paper.

# **Chronic nicotine enhances value sensitivity in mice through altered VTA DA activity**

M.L. Dongelmans *MSc*<sup>1</sup>, D. Lemoine *PhD*<sup>1</sup>, R. Durand-de Cuttoli *MSc*<sup>1</sup>, C. Ngyuen, *MSc*<sup>1</sup>, E.K.L. Duranté *MSc*<sup>1</sup>, R. Britto *BSc*<sup>1</sup>, F. Marti *PhD*<sup>1</sup>, N. Torquet *MSc*<sup>1</sup>, D. Dalkara *PhD*<sup>2</sup>, J. Naudé *PhD*<sup>1</sup>, A. Mourot *PhD*<sup>1</sup>, Ph. Faure *PhD*<sup>1\*</sup>

<sup>1</sup>Sorbonne Universités, UPMC Univ Paris 06, INSERM, CNRS, Neurosciences Paris Seine - Institut de Biologie Paris Seine (NPS - IBPS), 75005, Paris, France.

<sup>2</sup> Sorbonne Universités, UPMC Univ Paris 06, INSERM, U968, 75012, Paris, France.

\*Correspondence and requests for materials should be addressed to P.F. (email: phfaure@gmail.com)

## **ABSTRACT**

Dopamine (DA) neuron activity in the ventral tegmental area (VTA) reflects reward and value encoding. We demonstrate with an intracranial self-stimulation (ICSS) driven multi-armed bandit task for mice that long-term nicotine exposure increases value sensitivity. By changing the intensity of the ICSS such an increase is reversible and inducible in respectively nicotine-treated and control mice. In order to understand the basis of this change in value sensitivity we analyse VTA DA cell activity. *In vivo* electrophysiological recordings of VTA DA activity demonstrate a significant burst reduction after the ICSS multi-armed bandit task. However, due to the nicotine administration the tonic frequency is augmented. To confirm that the altered choice behaviour is linked to an increase in spontaneous tonic VTA DA neuron activity, we acutely mimicked the high tonic DA tone with specific optical stimulation of the VTA DA neurons and reversibly altered choices. Our results suggest that high tonic VTA DA neuron activity changes the incentive value of a reward.

**Running title:** Chronic nicotine enhances value sensitivity in mice through increased VTA DA activity

**Keywords:** nicotine, dopamine, ventral tegmental area, decision-making, bandit task, optical stimulation, *in vivo* electrophysiology, patch-clamp, C57BL/6JRJ mice

## INTRODUCTION

Nicotine, the main active component in tobacco, is highly addictive<sup>1</sup>. Like many addictive substances, nicotine is hypothesized to perpetuate addiction through plasticity in the mesocorticolimbic dopaminergic system<sup>2-7</sup>. The dopamine (DA) signal is essential in reinforcement and reward processing<sup>8</sup>. Nicotine exposure is associated with a shift in incentive salience<sup>9,10</sup>, as well as an increase in reward sensitivity in animals<sup>11,12</sup> and smokers<sup>13</sup>. However, reward sensitivity in these experiments refers to a motivation for visual reinforcement or a modification in intra-cranial self-stimulation threshold. These studies did not assign a utility to the possible outcomes, nor did they take the relative value of each outcome into account. Here we use a neuroeconomical behavioural paradigm to assess and quantify nicotine-induced shifts in value sensitivity<sup>14,15</sup> and electrophysiological measurement to correlate ventral tegmental area (VTA) DA cell activity to modifications in choice behaviour. DA neurons in the VTA exhibit different patterns of activity: a tonic activity that is associated with directional motivation<sup>16-18</sup> and phasic patterns that play a pivotal role in learning<sup>19-21</sup>. These activity patterns result in respectively constant or increased transient DA release in the nucleus accumbens (NAc) and the frontal cortices, projection sites of the VTA<sup>22,23</sup>. Nicotine alters the DA signal depending on its state<sup>24-28</sup>. Phasic DA is essential in novel and dynamic situations in which the values of environmental stimuli need to be updated regularly to adapt responses<sup>16,29,30</sup>. As illustrated by the reward prediction error theory<sup>20</sup>, little or no phasic activity implies a stable environment. In a stable environment, exploitation is the optimal strategy, while dynamic environments require exploration<sup>31</sup>.

The incentive-salience perspective theorizes that tonic dopamine can scale the incentive value associated with environmental stimuli, influencing behavioural choice in stable and dynamic environments<sup>32-34</sup>. In the classic reinforcement learning softmax action selection, the inverse temperature ( $\beta$ ) parameters measures the bias towards exploitation in a behavioural choice and reflects the trade-off between exploration-exploitation in behavioural paradigms<sup>15,35,36</sup>. Elevated tonic levels of DA were associated with an increased inverse temperature, a bias towards optimal

choices, in a stable environment<sup>16,37–39</sup>. However, a nicotine-induced increase in exploitation in a dynamic environment reduced behavioural flexibility<sup>40</sup>.

We demonstrate with a novel multi-armed bandit task for mice<sup>41</sup> using intracranial self-stimulation (ICSS) rewards that a nicotine-induced increase in tonic VTA DA activity enhances the incentive value of the available rewards by increasing the inverse temperature, a result that was mimicked by tonically photostimulating DA cells in the VTA. We quantify the augmented incentive value in nicotine-treated mice by altering the intensity of the ICSS bi-directionally. Our results demonstrate that passively administered nicotine augments value perception (incentive value) through alteration of the DA system and essentially makes the DA system more vulnerable for addiction.

## **METHODS AND MATERIALS**

### **Animals**

C57Bl/6j male mice (Janvier Labs, France) aged 8-16 weeks weighing 25-35 grams were used for all experiments, except for the optogenetic manipulation, which will be specified later. Mice were kept in an animal facility where temperature (21 $\pm$  1°C) and humidity were automatically monitored and where a circadian light cycle of 12/12-h light-dark cycle was maintained.

### **Intracranial self-stimulation electrode implantation**

Mice were anaesthetised with a gas mixture of oxygen (1L/min) and 1-3 % of isoflurane (Piramal Healthcare, UK), then introduced into a stereotaxic frame (Kopf Instruments, CA, USA). After the administration of a local anaesthetic, a median incision of the forebrain skin revealed the skull, which was after being scraped with a rongeur, was drilled at the level of the Median Forebrain Bundle (MFB). A bipolar stimulating electrode for ICSS was then implanted either on the left or on the right side of the brain (Stereotaxic coordinates after Paxinos: AP=1.4mm; ML= $\pm$  1.2mm, DV=4.8mm). Dental cement was used to fix the implant to the skull. After stitching and administration of a dermal antiseptic, mice were then placed back in their home-cage and had at least 5 days to recover from the surgery. The efficacy of electrical stimulation was verified through the rate of acquisition of the deterministic setting.

### **Drugs**

#### *Implantation of osmotic mini pumps*

Animals were anesthetized with a gas mixture of oxygen (1L/min) and 1-3 % of isoflurane (IsoVet, Piramal Healthcare, UK). After the administration of a local anaesthetic, an incision was performed at the level of the interscapular zone, to subcutaneously implant an osmotic minipump (Model 2004, ALZET, CA, USA) containing 200  $\mu$ L of either a solution of nicotine hydrogen tartrate salt (Sigma-Aldrich, MO, USA) at a dose of 10mg/kg/d (4.16 mg/kg, free base) or saline water (H<sub>2</sub>O with 0.9 % NaCl) for the control group. Both solutions were prepared in the laboratory. The

minipumps delivered their content with a flow of 0.25  $\mu\text{L}/\text{hour}$  over 28 days. The surgical wound was closed with surgical stitches. Animals had two days of rest to recover from the minipump surgery before going further with their behavioural training.

### **Virus injection and optogenetics experiments**

DAT-iCre mice (N=4) were anaesthetized (Isoflurane 1%) and implanted with an ICSS electrode as described above. The DAT-iCre mice were then intracranially injected (1  $\mu\text{L}$  injected into the VTA [coordinates from bregma AP: +3.2 mm; ML: +0.5 mm; DV: +4.6 mm] with an adeno-associated virus (AAV5.EF1a.DIO.CatCh.eYFP). A double-floxed inverse open reading frame (DIO) allowed restraining the expression of CatCh ( $\text{Ca}^{2+}$  translocating channelrhodopsin) to VTA dopamine neurons (DANs). An optical fiber (200  $\mu\text{m}$  core, NA=0.5, Thor Labs) coupled to a ferule (1.25 mm) was implanted just above the VTA [coordinates from bregma AP: +3.2 mm; ML: +0.5 mm; DV: +4.4 mm], and fixed to the skull with dental cement (SuperBond, Sun medical). An ultra-high power LED (470 nm from Prizmatix) coupled to a patchcord (500  $\mu\text{m}$  core, NA=0.5, Prizmatix) was used for optical stimulation (output intensity of 10 mW). Optical stimulation was delivered at a frequency of 8hz with 5ms/pulse, starting 5 min before and during the behavioural experiment, following a schedule of 4 paired ON and OFF days after the training phase of behavioural task was finished, at least 4 weeks after virus injection, to allow the construct to be integrated in the target cells genome. The optical stimulation cable was plugged onto the ferule during all experimental sessions to prepare the animals and control for latent experimental effects.

For a functional check of CatCh expression, 10-12 week old male DAT-iCre mice were injected with the same virus described above. After 4 weeks, mice were deeply anesthetized with an i.p. injection of a mix of ketamine/xylazine. Coronal midbrain sections (250  $\mu\text{m}$ ) were sliced using a Compressstome (VF-200; Precisionary Instruments) after intracardial perfusion of cold ( $4^{\circ}\text{C}$ ) sucrose-based artificial cerebrospinal fluid (SB-aCSF) containing (in mM): 125 NaCl, 2.5 KCl, 1.25  $\text{NaH}_2\text{PO}_4$ , 5.9  $\text{MgCl}_2$ , 26  $\text{NaHCO}_3$ , 25 Sucrose, 2.5 Glucose, 1 Kynurenate (pH 7.2, 325 mOsm). After 10 min to 1h at  $35^{\circ}\text{C}$  for recovery, slices were transferred into oxygenated aCSF containing

(in mM): 125 NaCl, 2.5 KCl, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 2 CaCl<sub>2</sub>, 1 MgCl<sub>2</sub>, 26 NaHCO<sub>3</sub>, 15 Sucrose, 10 Glucose (pH 7.2, 325 mOsm) at room temperature for the rest of the day and individually transferred to a recording chamber continuously perfused at 2 ml/min with oxygenated aCSF. Patch pipettes (4–8 M $\Omega$ ) were pulled from thin wall borosilicate glass (G150TF-3, Warner Instruments) using a micropipette puller (P-87, Sutter Instruments, Novato, CA) and filled with a K<sub>2</sub>Glu based intra-pipette solution containing either (in mM): 144 K<sub>2</sub>Glu, 3 MgCl<sub>2</sub>, 10 HEPES, 0.5 EGTA and 3 mg/ml biocytin (pH 7.2, 295mOsm) or: 116 K-gluconate, 10 HEPES, 0.5 EGTA, 6 KCl, 2 NaCl, 4 ATP, 0.3 GTP and 2 mg/mL biocytin (pH adjusted to 7.2). Transfected VTA DANs were visualised using an upright microscope coupled with a Doty contrast lens and illuminated with a white light source (Scientifica). A 460 nm LED (Cooled) was used for both visualizing eYFP positive cells (using a bandpass filter cube) and for optical stimulation through the microscope (1s continuous for light-evoked current in voltage-clamp mode and 10Hz with 5ms/pulses to drive neuronal firing in current-clamp mode). Whole-cell recordings were performed using a patch-clamp amplifier (Axoclamp 200B, Molecular Devices) connected to a Digidata (1550 LowNoise acquisition system, Molecular Devices). Signals were lowpass filtered (Bessel, 2 kHz) and collected at 10 kHz using the data acquisition software pClamp 10.5 (Molecular Devices). All the electrophysiological recordings were extracted using Clampfit (Molecular Devices) and analysed with R.

For an *in vivo* functional check of CatCh expression, transfected DANs were recorded in anesthetized animals as previously described. In that case, an optical fiber (500  $\mu$ m core, Prizmatix) was inserted in the glass pipette electrode and coupled to a 470 nm LED (Prizmatix). Light-pulses trains (1, 2, 5, 10, 20Hz with 5ms/pulse) were applied to test the functional expression of CatCh in VTA DANs.

### **In vivo electrophysiology**

Mice were deeply anaesthetized with chloral hydrate (8%), 400 mg/kg I.P., supplemented as required to maintain optimal anaesthesia throughout the experiment. The scalp was opened and a



whole was drilled in the skull above the location of the VTA. The saphenous vein was catheterized for intravenous administration of drugs. Extracellular recording electrodes were constructed from 1.5 mm O.D. / 1.17 mm I.D. borosilicate glass tubing (Harvard Apparatus) using a vertical electrode puller (Narishige). Under microscopic control, the tip was broken to obtain a diameter of approximately 1  $\mu\text{m}$ . The electrodes were filled with a 0.5% NaCl solution containing 1.5% of neurobiotin tracer (AbCys) yielding impedances of 6-9 M $\Omega$ . Electrical signals were amplified by a high-impedance amplifier (Axon Instruments) and monitored audibly through an audio monitor (A.M. Systems Inc.). The signal was digitized, sampled at 25 kHz and recorded on a computer using Spike2 software (Cambridge Electronic Design) for later analysis. The electrophysiological activity was sampled in the central region of the VTA (coordinates: between 3.1 to 4 mm posterior to Bregma, 0.3 to 0.7 mm lateral to midline, and 4 to 4.8 mm below brain surface). Individual electrode tracks were separated from one another by at least 0.1 mm in the horizontal plane. Spontaneously active DAergic neurons were identified based on previously established electrophysiological criteria (see main text, methods section).

### **Fluorescence immunohistochemistry.**

After euthanasia, brains were rapidly removed and fixed in 4% paraformaldehyde. Following a period of at least 3 d of fixation at 4 °C, serial 60- $\mu\text{m}$  sections were cut from the midbrain with vibratome. Immunohistochemistry was performed as follows: free-floating VTA brain sections were incubated 1 h at 4 °C in a blocking solution of phosphate-buffered saline (PBS) containing 3% Bovine Serum Albumin (BSA, Sigma; A4503) (vol/vol) and 0.2% Triton X-100 (vol/vol) and then incubated overnight at 4 °C with a mouse anti-tyrosine hydroxylase antibody (TH, Sigma, T1299) at 1:200 dilution and a rabbit anti-GFP antibody (Molecular Probes, A-6455) at 1:5,000 dilution in PBS containing 1.5% BSA and 0.2% Triton X-100. The following day, sections were rinsed with PBS and then incubated 3 h at 22–25 °C with Cy3-conjugated anti-mouse and Cy2-conjugated anti-rabbit secondary antibodies (Jackson ImmunoResearch, 715-165-150 and 711-225-152) at 1:200 dilution in a solution of 1.5% BSA in PBS. After three rinses in PBS, slices were wet-mounted using Prolong Gold Antifade Reagent (Invitrogen, P36930). Microscopy was carried out

with a fluorescent microscope, and images captured using a camera and ImageJ imaging software.

In the case of electrophysiological recordings, an immunohistochemical identification of the recorded neurons was performed as described above, with the addition of 1:200 AMCA-conjugated Streptavidin (Jackson ImmunoResearch) in the solution. Neurons labelled for both TH and neurobiotin in the VTA allowed to confirm their neurochemical phenotype.

In the case of optogenetics experiments on DAT-iCre mice, an immunohistochemical identification of the transfected neurons was performed as described above, with the addition of 1:500 Chicken-anti-YFP primary IgG (ab13970, Abcam) in the solution. A Goat-anti-chicken AlexaFluor 488 (1:500, Life Technologies) was then used as secondary IgG. Neurons labelled for TH, eYFP and neurobiotin/biocytin in the VTA allowed to confirm their neurochemical phenotype and the transfection success.

## **Behavioural data acquisition.**

### *Behavioural set up*

The ICSS bandit task took place in a circular open field with a diameter of 70 cm. Three explicit marks, in the shape of squares, were placed in the open field, forming an equilateral triangle (side = 50 cm). Entry of the circular zones (diameter = 9cm) around each mark was associated with the delivery of a rewarding ICSS stimulation. Experiments were performed using a video camera, connected to a video-track system, out of sight of the experimenter. A Labview application precisely tracked and recorded the animal's position by means of a camera (20 frames/sec). At day one when a mouse was detected in one of the circular rewarding zones an electrical stimulator received a TTL signal from the software application and generated a 200-ms train of 20 0.5-ms biphasic square waves pulsed at 100 Hz. ICSS intensity was adjusted, within a range of 20 to 200  $\mu$ A, during the first week of training (see training settings) and then kept constant, so that mice would achieve between 50 and 150 zone visits per session.

### *Training settings*

The training consisted out of two settings: the deterministic setting (DS) and the probabilistic setting (PS) both consisting out of 10 daily sessions of 5 min. In the CS, all zones were associated with an ICSS delivery ( $P=1$ ). However, no two consecutive rewards could be delivered in the same zone, which motivated the mice to stay in motion. Mice with insufficient scores were excluded during the DS. In the PS, the zones were associated with three different probabilities ( $P=.25$ ,  $P=.5$ ,  $P=1$ ) to obtain an ICSS stimulation. The probabilities were pseudo-randomly assigned per mouse.

### *Data acquisition per experimental group*

Different experimental groups underwent the ICSS bandit task. Firstly, locomotive and choice behaviour of the mice which had been implanted with osmotic minipumps (SAL = 28; NIC = 29) were analysed and compared between the last two days of both training settings (d9&10 (DS) + d19&20 (PS)). Another group of animals ( $N=12$ ) underwent the full training followed by 5 days in which the intensity was increased in bins of  $20\mu\text{A}$  ( $+20$ - $100\mu\text{A}$ ) in a random order. These test days were compared with the baseline (Day 10 of PS). Similarly, a group of animals implanted with minipumps (NIC= 5) completed the training sessions, followed by 2 days of a decreased ICSS intensity in bins of  $20\mu\text{A}$  ( $-20$ - $40\mu\text{A}$ ). Finally, the DATiCre mice ( $N=4$ ) completed the training succeeded by a schedule of 8 days of sessions with photostimulation (ON) altered with days without photostimulation (OFF). The averages of the ON and OFF days were compared in a paired manner.

### *Behavioural measures*

For all groups the following measures were analysed and compared in the US, and for the SAL vs NIC experiment as well as in the DS: i) number of visits; ii) average speed; iii) choice repartition (proportional visits); iv) percentage of directional changes ( $n^{\text{th}}$  visit= $n^{\text{th}}$  visit+2). Furthermore, the ICSS bandit task enforces a Markovian decision process. Every transition between zones can be considered as a binary choice between two probabilities, since the occupied zone cannot be reinforced twice. The sequence of choices per session can be resumed as the proportional

outcome of the sum of three specific binary choices (gambles, f.e. total visits zone 1/total visits zone 1+2). These gambles (G) are G1 = 100 % vs 50 %, G2 = 50 % vs 25 % and G3 = 100 % vs 25 %. The outcome of these gambles illustrates the balance between exploitative (choosing the most valuable option) and exploratory (choosing the least valuable option) choices. With a soft-max based decision-making model fitted in the laboratory<sup>41</sup> we compute two parameters the value sensitivity (the power to discriminate between values in a binary choice) and the uncertainty bonus (the preference for expected uncertainty, considering the reward variance of every option in a binary choice) per animal. In this experiment, the values of the choices are equal to the probabilities, except for the groups in which we changed the intensity of the ICSS during the test days.

#### *Locomotor experiment*

Two naïve groups of animals were implanted with osmotic minipumps (SAL=12 & NIC=14). 22-24 days after the implantation their locomotor activity was assessed in a circular open field (diameter=80cm) during a 30-minute session. Distance travelled was analysed and compared between the groups per 5-minute bins.

#### **Statistical analysis.**

All statistical analyses were computed with R statistical software (The R Project). Results are plotted as a mean  $\pm$  s.e.m. The total number ( $n$ ) of observations in each group and the statistics used are indicated in figure captions. Classically comparisons between means were performed using parametric tests (Student's T-test, or ANOVA for comparing more than two groups) when parameters followed a normal distribution (Shapiro test  $P > 0.05$ ), and non-parametric tests (here, Wilcoxon or Mann-Whitney) when the distribution was skewed. Multiple comparisons were Bonferroni corrected. All statistical tests were two-sided.  $P > 0.05$  was considered not to be statistically significant.

## RESULTS

### The multi-armed ICSS bandit task

We used the multi-armed ICSS bandit task<sup>41</sup> (Fig 1a and Methods) to assess the choice distribution in a probabilistic environment of mice under nicotine. Moving past the marked locations within the daily 5-min session resulted in the delivery of electrical stimulations in the medial forebrain bundle (Figure 1a). During training, every visit to a mark was reinforced ( $P=1.00$ ) and followed by an inactivation for one trial to encourage the mice to explore the open field environment. After a learning period, mice alternate from one point to the other (Figure 1b). Trajectories at the end of the training period were stereotyped, almost circular, in both mice under saline and nicotine exposure. Both groups distributed their visits equally over the three locations (Figure 1c left) and their probability of directional change is equal (Figure 1c right). However, the total number of rewards (i.e ICSS) received by mice under chronic nicotine was higher than those received by mice under saline (Figure 1d left). This difference seems to rely exclusively on the animal's velocity (Figure 1d right). Overall, mice showed a higher peak of instantaneous velocity when they move from one target to the other (Figure 1e). A similar effect on velocity is only observed in the beginning of an open-field exposure (Supp Fig 1), thus suggesting that it could arise from a combination of ICSS and the effect of nicotine.

### Nicotine increases value sensitivity in multi-armed ICSS bandit task

We next compared behaviour of mice under saline (SAL) and chronic nicotine (NIC) in a probabilistic setting (PS) where each location was associated with a different probability of ICSS delivery (Figure 2a). As already demonstrated, expected uncertainty in the PS induced a markedly different behavioural pattern when compared to the training period in which all rewards are certain<sup>41</sup>. In the PS, differences between SAL and NIC mice are more prominent. Compared to mice under saline, mice under chronic nicotine showed an increase in the number of transitions (Figure 2b left), in the number of rewards received (Figure 2b middle) and in the percentage of directional change (Figure 2c right). Overall time between two targets is decreased for NIC mice

while traveled distance is almost the same (Supp figure 2). This suggests a modified speed profile and that the instantaneous speed is higher in mice under chronic nicotine. More specifically, the peak of the instantaneous speed between two targets is specifically decreased in mice under saline after reward delivery. This is not the case for mice under nicotine (Figure 2c). The directional change is an indication that the NIC mice choose differently compared to the SAL group. This effect cannot be accounted for by an increase in the speed. We thus analysed the animal's repartitions and sequences of choice. In the probabilistic setting, mice visited the locations associated with higher ICSS probability more often. Overall both groups distribute their choices depending on the probabilities to receive a reward (Figure 2d). However, mice under chronic nicotine focus more on the most rewarded place (i.e 50 and 100 %). We then analysed the sequence of choice statistics using transition function<sup>41</sup>. Since mice could not receive two consecutive ICSSs on the same target, the repartition on the rewarding locations arose from a sequence of binary choices (Figure 2a) in three gambles (G1, G2, G3) between two respective payoffs (here, G1 = {100 versus 50%}, G2 = {50 versus 25%}, G3 = {100 versus 25%}). For each gamble, mice chose the optimal location (associated with the highest probability of reward; Figure 2e) more than 50% of the time. This repartition is different in WT mice under chronic nicotine, that favour optimal choice in gamble G2 and G3 (i.e when sub-optimal is 25%). The probability to choose the optimal choice in each gamble is defined by a transition function. We fit this function with the uncertainty model<sup>41</sup>. In this model, decision is biased toward actions with the most uncertain consequences by assigning a bonus value to their expected uncertainties. The model accurately reproduced the pattern of mice sequence of choice and suggests an increase in the inverse temperature parameter  $\beta$  (Figure 2f). This modification in the behaviour is obtained without a difference in the mean intensity associated with ICSS (Figure 2g). This difference developed with time (Supp. Figure 3).

### **Value sensitivity is directly modified by altered ICSS values**

In the softmax model<sup>41</sup>, an increase in the value of the ICSS is equivalent to an increase in beta ( $\beta$ ). Otherwise stated, it is impossible with this model to differentiate an increase in  $\beta$  and an

increase in the reward value associated with the ICSS. Even if stimulation intensities are identical, a modification of the “subjective value” cannot be distinguished with this approach. We thus test the effect of a controlled modification of the value associated with the reward. After training, mice were tested for five days in which the intensity of the ICSS was increased, in bins of 20 $\mu$ A (+20-100 $\mu$ A), in a random order. The continuous change in choice repartitions corresponded to the increasing ICSS values (Supp Figure 4a). For an increase of +80 $\mu$ , the total number of transitions and rewards did not vary (Figure 3b left and middle), but the percentage of directional change did increase (Figure 3b right), indicating a modification in the choice sequence. This is confirmed by a significant modification of the repartition. Mice exposed to an increased intensity favour optimal choice (100% and 50%). This is accordingly associated with an increase in  $\beta$  (Figure 3c left). An increase in the uncertainty parameter is also observed. These results indicate that a modification of the intensity can mimic some aspects of the effect of chronic exposure to nicotine. By decreasing the intensities in chronic nicotine treated mice, after a completed training, we aimed at replicating this effect in the inversed sense (Figure 3d,e,f, Supp Figure 4b). These results are preliminary ( $N=5$ ) but a tendency seems to emerge. We observed a reversed effect to what has been observed in mice exposed to saline, with in particular a decrease in the  $\beta$  parameter. These results suggest that at some point chronic nicotine exposure can be associated with an increase in the subjective value of the reward.

### **Value sensitivity can be induced by VTA DA tonic stimulation**

We analysed the spontaneous activity of VTA DA cells in WT exposed to saline or nicotine (Figure 4a). DA cell firing was analysed with respect to the average firing rate and the percentage of spikes within a burst<sup>42</sup>. Chronic exposure to nicotine does not modify firing frequency (Figure 4b left), but produces an increase in VTA DA cells bursting (Figure 4b right). Mice exposed to chronic nicotine and ICSS stimulation in the task exhibited an increase in firing frequency (Figure 4b left) and bursting patterns (Figure 4c right). Compared to mice only exposed to saline, ICSS and chronic nicotine exposure seems to produce a specific alteration of VTA DA cell activity, which results in a high tonic activity. Therefore, we manipulated VTA DA cells activity using optogenetics

approach (Figure 5a) to link the modified behaviour, as observed in the task, with the modifications in VTA DA cell activity. To tonically activate VTA DA neurons, we selectively expressed CatCh (Ca<sup>2+</sup> translocating channel rhodopsine) in these cells in DAT-iCRE mice (Figure 5b). Patch-clamp recordings demonstrate that light can evoke a current and activate firing (Figure 5c). Juxtacellular in-vivo recordings and stimulation of VTA DA cells also demonstrate that neurons can be driven by a 5ms pulse at 1 to 10 Hz. For higher stimulation frequencies, the % of spikes emitted after a pulse decreases and falls to approximately 50% (Figure 5d). After the training period, mice went through a schedule of 8 days of sessions with photostimulation (ON) alternating with days without photostimulation (OFF). Four mice have been analysed. Number of transitions and rewards are increased in the ON condition; when the light is on. Furthermore, repartitions among the three targets also show that WT mice favour the higher probabilities in gamble G2 and G3, i.e. when sub-optimal is 25%, during photostimulation.



## DISCUSSION

In this non-transient multi-armed ICSS bandit task, nicotine-treated (NIC) mice learned progressively to visit the higher probabilities ( $P=0.5$  and  $P=1.0$ ) significantly more frequently, compared to saline-treated (SAL) animals that followed a choice pattern proportional to the probabilities available, as described by the matching law theory<sup>43</sup>, but with an additional bonus for uncertainty<sup>41</sup>. The preference for higher probabilities resulted in significantly more rewards for NIC mice, through different choice patterns and resulted in an increased inverse temperature. We were able to induce a similar increase in inverse temperature by either augmenting the values of the ICSS reward (gradual increases in intensity ( $\mu\text{A}$ )) or by tonically photostimulating the VTA DA neurons. We then observed a tendency to decrease the inverse temperature in mice under nicotine after weakening the ICSS reward. This manipulation of the DA cell activity mimics the effect of nicotine exposure. Indeed, long-term nicotine administration during the ICSS task increased the frequency and bursting activity of VTA DA compared to saline. However, compared to the SAL and NIC control groups the global bursting activity seemed to be lessened due to an interaction effect with the ICSS. This data suggests that nicotine, through modification of the DA signal, shifts the subjective value of the ICSS reward, while the magnitude of the reward remains constant. This shift in inverse temperature is similar to an actual increase in ICSS intensity (i.e a direct manipulation of the value).

The augmented inverse temperature reflects increased exploitative behaviour, an effect that has been previously linked to enhanced tonic dopamine activity. Tonic dopamine is hypothesized to modulate the bias towards optimal choice<sup>16,44,45</sup>. In this paper we demonstrate with electrophysiological and optogenetic techniques a direct link between DA cell activity and exploitation. The multi-armed ICSS bandit task enables, through a clear distinction between speed and choice analyses, to distinguish the modified components of value-based decision-making that have been assigned to increased tonic DA levels<sup>39,46,47</sup>, induced by nicotine. We explicitly demonstrate an increase in value sensitivity due to nicotine-induced DA activity alterations.

Previous ICSS studies have observed that drugs sensitize ICSS stimulation and require lower stimulation frequencies to increase self-stimulation<sup>12</sup>. Our result goes along the same line, but here we quantify the effect of this increased value sensitivity on choice and link this modification to an augmented tonic activity of VTA DA cells.

Long-term nicotine exposure increases VTA DA cell activity<sup>28,48</sup> through desensitization and up-regulation of nicotinic acetylcholinergic receptors (nAChRs). However, we observed a slight increase in VTA DA cell frequency between ICSS mice and controls in the Saline and Nicotine groups. Non-contingency studies with yoked nicotine exposure have previously shown to increase the incentive salience of non-nicotine stimuli<sup>55</sup>, similarly to the sensitization for ICSS reward<sup>56</sup>, suggesting an underlying mechanism to explain the essential role of contextual cues in smoking and a nicotine-induced increased reward sensitivity. Neuroeconomical studies have linked smoking with increased impulsivity (delay discounting task<sup>57</sup>) and decreased behavioural flexibility ("Restless" bandit task<sup>40,58</sup>). We aimed at providing a better understanding, as well as a valuation of the effect nicotine has on value sensitivity in a stable environment to define the altered reward processing in smokers more precisely. These data imply altered choice behaviours in smokers that are not also limited to addiction, a hypothetical addictive brain-state, can have big implications on daily life and has been suggested to increase vulnerability for other addictions, like gambling and other drugs<sup>59</sup>.

**ACKNOWLEDGEMENTS** This work was supported by the Centre National de la Recherche Scientifique CNRS UMR 8246, the Foundation for Medical Research (FRM, Equipe FRM DEQ2013326488 to P.F. ; fourth year thesis to MD and R.D), the region île de France (Dim Cerveau et pensée to P.F.), the French National Cancer Institute Grant TABAC-16-022 (to P.F.) and The LabEx Bio-Psy. P.F.'s laboratory is part of the École des Neurosciences de Paris Ile-de-France RTRA network. P.F. is a member of LabEx Bio-Psy and of DHU Pepsy.

**AUTHOR CONTRIBUTIONS:** M.D. and P.F. designed the study. M.D. and P.F. analysed the behavioural data. F.M., and C.N. performed the electrophysiological juxtacellular recordings. E.D. and R.B. contributed to behavioural experiments. R.D, A.M, D.L and D.D developed optogenetic tools and performed electrophysiological cell recordings. J.N. contributed to the analyses of the behavioural data. M.D and P.F. wrote the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

## REFERENCES

1. Benowitz, N. L. Pharmacology of Nicotine: Addiction, Smoking-Induced Disease, and Therapeutics, *Annu Rev Pharmacol Toxicol.* **49**, 57–71 (2009). doi:10.1146/annurev.pharmtox.48.113006.094742.Pharmacology
2. Pidoplichko, V. I., DeBiasi, M., Williams, J. T. & Dani, J. a. Nicotine activates and desensitizes midbrain dopamine neurons. *Nature* **390**, 401–4 (1997).
3. Dani, J. A., Ji, D. & Zhou, F. Synaptic plasticity and Nicotine Addiction. *Neuron* **31**, 349–352 (2001).
4. Dani, J. A., Jenson, D., Broussard, J. I. & Biasi, M. De. NIH Public Access. **1**, 1–12 (2012).
5. Subramanian, M. & Dani, J. A. Dopaminergic and cholinergic learning mechanisms in nicotine addiction. *Ann. N. Y. Acad. Sci.* **1349**, 46–63 (2015).
6. Koob, G. F. & Le Moal, M. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* **24**, 97–129 (2001).
7. Pistillo, F., Clementi, F., Zoli, M. & Gotti, C. Nicotinic, glutamatergic and dopaminergic synaptic transmission and plasticity in the mesocorticolimbic system: Focus on nicotine effects. *Prog. Neurobiol.* **124**, 1–27 (2015).
8. Averbeck, B. B. & Costa, V. D. Motivational neural circuits underlying reinforcement learning. *Nat. Neurosci.* **20**, 505–512 (2017).
9. Chaudhri, N. *et al.* Complex interactions between nicotine and nonpharmacological stimuli reveal multiple roles for nicotine in reinforcement. *Psychopharmacology (Berl)*. **184**, 353–366 (2006).
10. Donny, E. C. *et al.* Operant responding for a visual reinforcer in rats is enhanced by noncontingent nicotine: Implications for nicotine self-administration and reinforcement. *Psychopharmacology (Berl)*. **169**, 68–76 (2003).
11. Hilario, M. R. F., Turner, J. R. & Blendy, J. a. Reward sensitization: effects of repeated nicotine exposure and withdrawal in mice. *Neuropsychopharmacology* **37**, 2661–70 (2012).
12. Wise, R. A. Addictive drugs and brain stimulation reward. *Annu. Rev. Neurosci.* **19**, 319–40 (1996).
13. Potts, G. F., Bloom, E. L., Evans, D. E. & Drobos, D. J. Neural reward and punishment sensitivity in cigarette smokers. *Drug Alcohol Depend.* **144**, 245–253 (2014).
14. Redish, a D. Addiction as a computational process gone awry. *Science* **306**, 1944–7 (2004).
15. Daw, N. D. & Doya, K. The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* **16**, 199–204 (2006).
16. Niv, Y., Daw, N., Dayan, P., Computational, G. & Unit, N. How fast to work : Response vigor , motivation and tonic dopamine.

17. Grace, A. & Grace, A. Grace AA , Bunney BS . The control of firing pattern in nigral dopamine neurons : burst firing . *J Neurosci* 4 : 2877-2890. 2877–2890 (1984).
18. Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.* **28**, 309–69 (1998).
19. Apicella, P., Ljungberg, T., Scarnati, E. & Schultz, W. Responses to reward in monkey dorsal and ventral striatum. *Exp. Brain Res.* **85**, 491–500 (1991).
20. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–9 (1997).
21. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
22. Hamid, A. A. *et al.* Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–126 (2015).
23. Adamantidis, A. R. *et al.* Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J. Neurosci.* **31**, 10829–35 (2011).
24. Zhang, H. & Sulzer, D. Frequency-dependent modulation of dopamine release by nicotine. *Nat. Neurosci.* **7**, 581–2 (2004).
25. Wang, L. *et al.* Modulation of dopamine release in the striatum by physiologically relevant levels of nicotine. *Nat. Commun.* 1–9 (2014). doi:10.1038/ncomms4925
26. Koranda, J. L. *et al.* Nicotinic receptors regulate the dynamic range of dopamine release in vivo. *J. Neurophysiol.* **111**, 103–11 (2014).
27. Tan, H., Bishop, S. F., Lauzon, N. M., Sun, N. & Laviolette, S. R. Chronic nicotine exposure switches the functional role of mesolimbic dopamine transmission in the processing of nicotine's rewarding and aversive effects. *Neuropharmacology* **56**, 741–751 (2009).
28. Besson, M. *et al.* Long-term effects of chronic nicotine exposure on brain nicotinic receptors. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8155–60 (2007).
29. Redgrave, P., Gurney, K. & Reynolds, J. What is reinforced by phasic dopamine signals? *Brain Res. Rev.* **58**, 322–339 (2008).
30. Hyland, B. I., Reynolds, J. N. J., Hay, J., Perk, C. G. & Miller, R. Firing Modes of Midbrain Dopamine Cells in Rats. **114**, (2002).
31. Cohen, J. D., McClure, S. M. & Yu, A. J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **362**, 933–42 (2007).
32. Phillips, P. E. M., Stuber, G. D., Heien, M. L. A. V., Mark Wightman, R. & Carelli, R. M. erratum: Subsecond dopamine release promotes cocaine seeking. *Nature* **423**, 461–461 (2003).
33. Robinson, M. J. F., Robinson, T. E. & Berridge, K. C. Incentive Saliency and the Transition to Addiction. *Biol. Res. Addict.* 391–399 (2013). doi:10.1016/B978-0-12-398335-0.00039-X
34. Roitman, M. F. Dopamine Operates as a Subsecond Modulator of Food Seeking. *J. Neurosci.* **24**, 1265–1271 (2004).
35. Beeler, J. a., Frazier, C. R. M. & Zhuang, X. Putting desire on a budget: dopamine and energy expenditure, reconciling reward and resources. *Front. Integr. Neurosci.* **6**, 49 (2012).
36. sutton-barto.PsychRev.81.pdf.
37. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. **191**, 507–20 (2007).
38. Rigoli, F. *et al.* Dopamine increases a value-independent gambling propensity. *Neuropsychopharmacology* **41**, in press (2016).
39. Beeler, J. a, Daw, N., Frazier, C. R. M. & Zhuang, X. Tonic dopamine modulates exploitation of reward learning. *Front. Behav. Neurosci.* **4**, 170 (2010).
40. Addicott, M. a, Pearson, J. M., Wilson, J., Platt, M. L. & McClernon, F. J. Smoking and the bandit: a preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Exp. Clin. Psychopharmacol.* **21**, 66–73 (2013).
41. Naude, J. *et al.* Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking. *Nat Neurosci advance on*, (2016).
42. Faure, P., Tolu, S., Valverde, S. & Naudé, J. The role of nicotinic acetylcholine receptors in regulating Dopamine neuron activity. *Neuroscience* (2014).

doi:10.1016/j.neuroscience.2014.05.040

43. Heyman, G. M. Addiction and choice: theory and new data. *Front. psychiatry* **4**, 31 (2013).
44. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. **191**, 507–20 (2007).
45. Dayan, P. & Niv, Y. Reinforcement learning: The Good, The Bad and The Ugly. *Curr. Opin. Neurobiol.* **18**, 185–196 (2008).
46. Salamone, J. D., Correa, M., Farrar, A. M., Nunes, E. J. & Pardo, M. Dopamine, behavioral economics, and effort. *Front. Behav. Neurosci.* **3**, 13 (2009).
47. Salamone, J. D., Correa, M., Farrar, A. & Mingote, S. M. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)*. **191**, 461–482 (2007).
48. Tolu, S. *et al.* Nicotine enhances alcohol intake and dopaminergic responses through  $\beta 2^*$  and  $\beta 4^*$  nicotinic acetylcholine receptors. *Sci. Rep.* **7**, 45116 (2017).
49. Geisler, S., Derst, C., Veh, R. W. & Zahm, D. S. Glutamatergic afferents of the ventral tegmental area in the rat. *J. Neurosci.* **27**, 5730–43 (2007).
50. Margolis, E. B., Mitchell, J. M., Ishikawa, J., Hjelmstad, G. O. & Fields, H. L. Midbrain Dopamine Neurons: Projection Target Determines Action Potential Duration and Dopamine D2 Receptor Inhibition. *J. Neurosci.* **28**, 8908–8913 (2008).
51. Ducrot, C., Fortier, E., Bouchard, C. & Rompré, P.-P. Opposite modulation of brain stimulation reward by NMDA and AMPA receptors in the ventral tegmental area. *Front. Syst. Neurosci.* **7**, 57 (2013).
52. Wise, R. & Rompre, P. Brain Dopamine And Reward. *Annu. Rev. Psychol.* **40**, 191–225 (1989).
53. Ungless, M. a, Whistler, J. L., Malenka, R. C. & Bonci, a. Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature* **411**, 583–7 (2001).
54. Carlezon, W. A. & Chartoff, E. H. Intracranial self-stimulation (ICSS) in rodents to study the neurobiology of motivation. *Nat. Protoc.* **2**, 2987–2995 (2007).
55. Palmatier, M. I. *et al.* Dissociating the primary reinforcing and reinforcement-enhancing effects of nicotine using a rat self-administration paradigm with concurrently available drug and environmental reinforcers. *Psychopharmacology (Berl)*. **184**, 391–400 (2006).
56. Kenny, P. J. & Markou, A. Nicotine Self-Administration Acutely Activates Brain Reward Systems and Induces a Long-Lasting Increase in Reward Sensitivity. *Neuropsychopharmacology* 1203–1211 (2005). doi:10.1038/sj.npp.1300905
57. Locey, M. L. & Dallery, J. Isolating behavioral mechanisms of intertemporal choice: nicotine effects on delay discounting and amount sensitivity. *J. Exp. Anal. Behav.* **91**, 213–23 (2009).
58. Addicott, M. A., Pearson, J. M., Froeliger, B., Platt, M. L. & Joseph McClernon, F. Smoking automaticity and tolerance moderate brain activation during explore-exploit behavior. *Psychiatry Res. - Neuroimaging* **224**, 254–261 (2014).
59. Adinoff, B. Neurobiologic processes in drug reward and addiction. *Harv. Rev. Psychiatry* **12**, 305–20 (2004).



**Figure 1**      *Intracranial self-stimulation multi-armed bandit task for mice.*

- A.** Mice were unilaterally implanted with a stimulation electrode in the medial forebrain bundle (Paxinos stereotaxic coordinates: AP= +1.4mm; ML= +/- 1.2mm; DV=4.8mm). After recovery mice were familiarized to the multi-armed bandit task: an open-field (0.8-m diameter) with three explicit squares (2x2 cm) forming an equilateral triangle (50-cm side). After 5 days of deterministic setting (DS, *see Methods*) in which mice were rewarded with an ICSS at all three locations, mice were implanted with an osmotic minipump filled with nicotine (NIC, 10mg/kg/d) or saline (SAL).
- B.** Trajectories of one mouse (5 min) at the 1<sup>st</sup> day (left) of training and the 10<sup>th</sup> day (right) in the DS.
- C.** Number of rewards (equal to number of transitions) are significantly different between SAL (N=28) and NIC (N=29) in DS (T-test,  $t = -2.899$ ,  $df = 52.545$ ,  $p=0.0054$ ). No differences were found the proportion of visits over the different locations ( $p_A$ : T-test,  $t = -1.281$ ,  $df = 53.573$ ,  $p=0.206$ ;  $p_B$ : T-test,  $t = 1.102$ ,  $df = 54.405$ ,  $p=0.275$ ;  $p_C$ : T-test,  $t = 0.109$ ,  $df = 51.261$ ,  $p=0.9139$ ).
- D.** Time to goal (average duration from one location to another in seconds) was significantly higher for NIC compared to SAL (T-test,  $t = 3.155$ ,  $df = 45.11$ ,  $p=0.002852$ ), while there was no difference in the percentage of directional changes (T-test,  $t = 1.4431$ ,  $df = 54.646$ ,  $p=0.1547$ ).
- E.** The maximum speed (in centimeters per second) after ICSS delivery was significantly higher for NIC mice compared to SAL (T-test,  $t = -2.9784$ ,  $df = 55.879$ ,  $p=0.00428$ ).

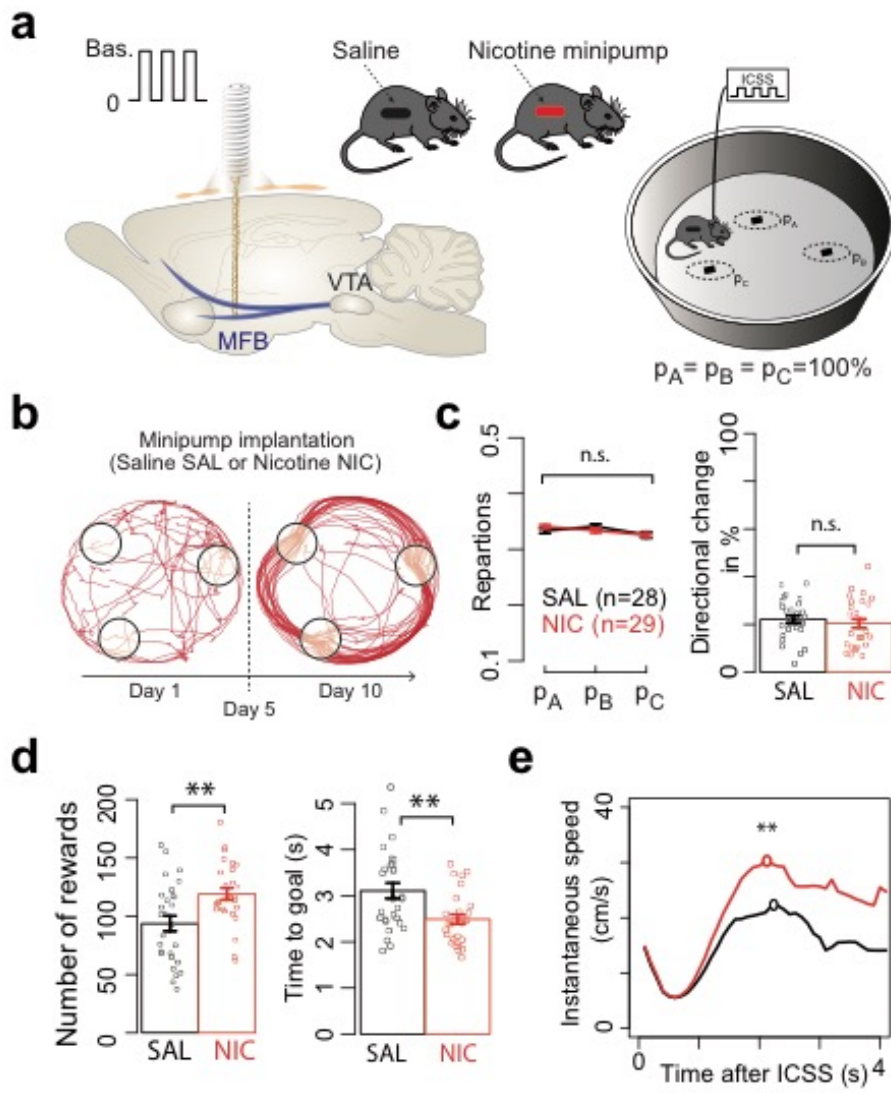


Figure 1



**Figure 2** *Choice behaviour in the probabilistic setting.*

- A.** Representation of the state transitions (gambles) in the probabilistic setting (PS) of the multi-armed bandit task. Gamble 1 (G1) =  $P_{(B|C)}$ , Gamble 2 (G2) =  $P_{(A|C)}$  and Gamble 1 (G1) =  $P_{(A|B)}$ . Trajectories of one SAL mouse and one NIC mouse at the 10<sup>th</sup> day of PS (5 min).
- B.** Both the number of transitions (T-test,  $t = -2.433$ ,  $df = 45.893$ ,  $p=0.01893$ ), as well as the number of rewards (T-test,  $t = -2.7957$ ,  $df = 50.051$ ,  $p=0.007326$ ) were significantly higher for NIC versus SAL mice in the PS. The NIC group also showed an increase in the percentage of directional change (T-test,  $t = -2.5026$ ,  $df = 43.874$ ,  $p=0.01613$ ) in the PS.
- C.** Instantaneous speed after ICSS was different NIC and SAL at maximum speed (T-test,  $t = -3.3786$ ,  $df = 72.799$ ,  $p=0.001173$ ). No difference was observed between the two groups after omissions (T-test,  $t = -1.5912$ ,  $df = 79.138$ ,  $p=0.1155$ ). The maximum speed was not significantly different between rewarding and omitted trials for NIC (T-test,  $t = 1.2888$ ,  $df = 72.44$ ,  $p=0.2016$ ). However, a difference was observed for SAL (T-test,  $t = 3.9876$ ,  $df = 67.294$ ,  $p=0.0001631$ ).
- D.** The proportions of visits over the different probability locations was significantly lower for NIC on  $p_C=0.25$  (T-test,  $t = 2.4124$ ,  $df = 45.65$ ,  $p=0.01993$ ) and higher on  $p_B=0.5$  (T-test,  $t = -2.6236$ ,  $df = 45.321$ ,  $p=0.01181$ ), but no difference was found for  $p_A=1.00$  (T-test,  $t = -1.578$ ,  $df = 52.992$ ,  $p=0.1205$ ).
- E.** A difference between SAL and NIC mice was displayed for G2 ( $P_{(A|C)}$ , T-test,  $t = -2.3657$ ,  $df = 52.971$ ,  $p=0.02169$ ), but not for G1 ( $P_{(B|C)}$ , (T-test,  $t = 0.9966$ ,  $df = 48.095$ ,  $p=0.3239$ ), nor G3 ( $P_{(A|B)}$ , T-test,  $t = -1.178$ ,  $df = 45.669$ ,  $p=0.2449$ ).
- F.** The uncertainty softmax model gave a significantly different parameter for the inverse temperature ( $\beta$ ) between SAL and NIC (T-test,  $t = -2.6537$ ,  $df = 38.81$ ,  $p=0.01148$ ), no difference was found on the uncertainty parameter ( $\phi$ , T-test,  $t = -1.4508$ ,  $df = 51.943$ ,  $p=0.1528$ ).
- G.** The average intensity in  $\mu A$  of the ICSS was similar for SAL and NIC (T-test,  $t = 0.60649$ ,  $df = 54.25$ ,  $p=0.5467$ ).

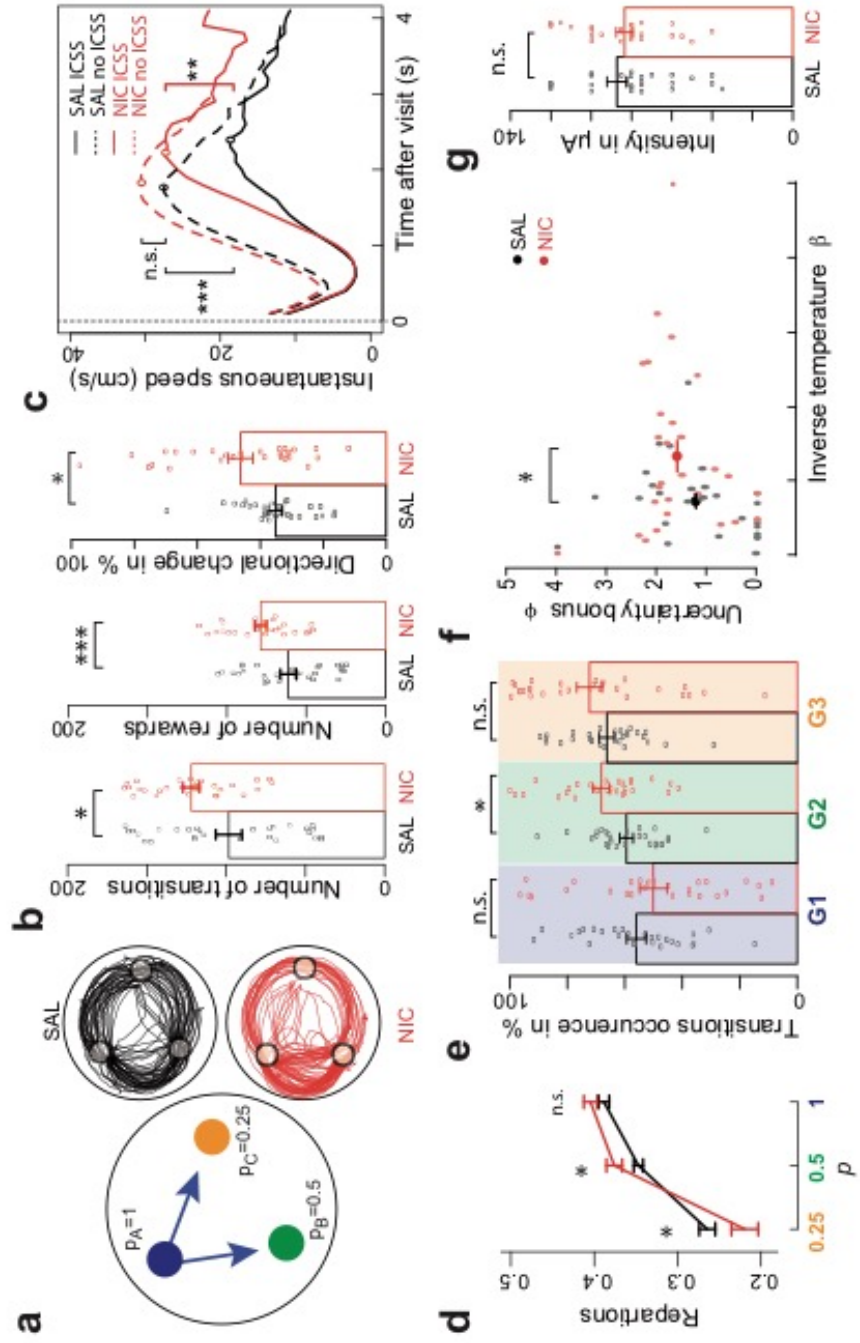


Figure 2

**Figure 3** *Modification of values through bidirectional shifts in the ICSS intensity.*

- A.** The intensity of the ICSS was increased with 20; 40; 60 and 80 $\mu$ A in a random order over 4 days for mice with a SAL minipump (N=12) and compared with baseline measures (10<sup>th</sup> day of PS).
- B.** No significant differences were found between the number of transitions (paired T-test,  $t = 0.031887$ ,  $df = 11$ ,  $p=0.9751$ ) and rewards (paired T-test,  $t = -0.38613$ ,  $df = 11$ ,  $p=0.7068$ ) when comparing baseline with the +80 $\mu$ A condition. However, in the +80 $\mu$ A, SAL mice displayed a significant increase in the percentage of directional change (paired T-test,  $t = -4.4239$ ,  $df = 11$ ,  $p=0.001022$ ).
- C.** The proportion of visits of the different probabilities was significant different for all locations. With a decrease in visits to  $p_C=0.25$  (paired T-test,  $t = 5.4313$ ,  $df = 11$ ,  $p=0.0002066$ ) and a proportional increase for  $p_B=0.5$  (paired T-test,  $t = -6.0943$ ,  $df = 11$ ,  $p=7.8e-05$ ) and  $p_A=1.00$  (paired T-test,  $t = -2.6112$ ,  $df = 11$ ,  $p=0.02421$ ) for for the +80 $\mu$ A condition compared to baseline. This shift translated into a significant increase in the inverse temperature (paired T-test,  $t = 3.7393$ ,  $df = 11$ ,  $p=0.00327$ ) and uncertainty parameter (paired T-test,  $t = 4.8022$ ,  $df = 11$ ,  $p=0.0005514$ ) for the +80 $\mu$ A condition.
- D.** In mice under nicotine exposure (N=5) the ICSS intensity was decreased with -20 $\mu$ A and -40 $\mu$ A in a random order. Intensity decreases were compared with baseline (10<sup>th</sup> day of PS).
- E-F.** Due to a limited sample size (N=5) and statistical power no statistical analyses were performed.

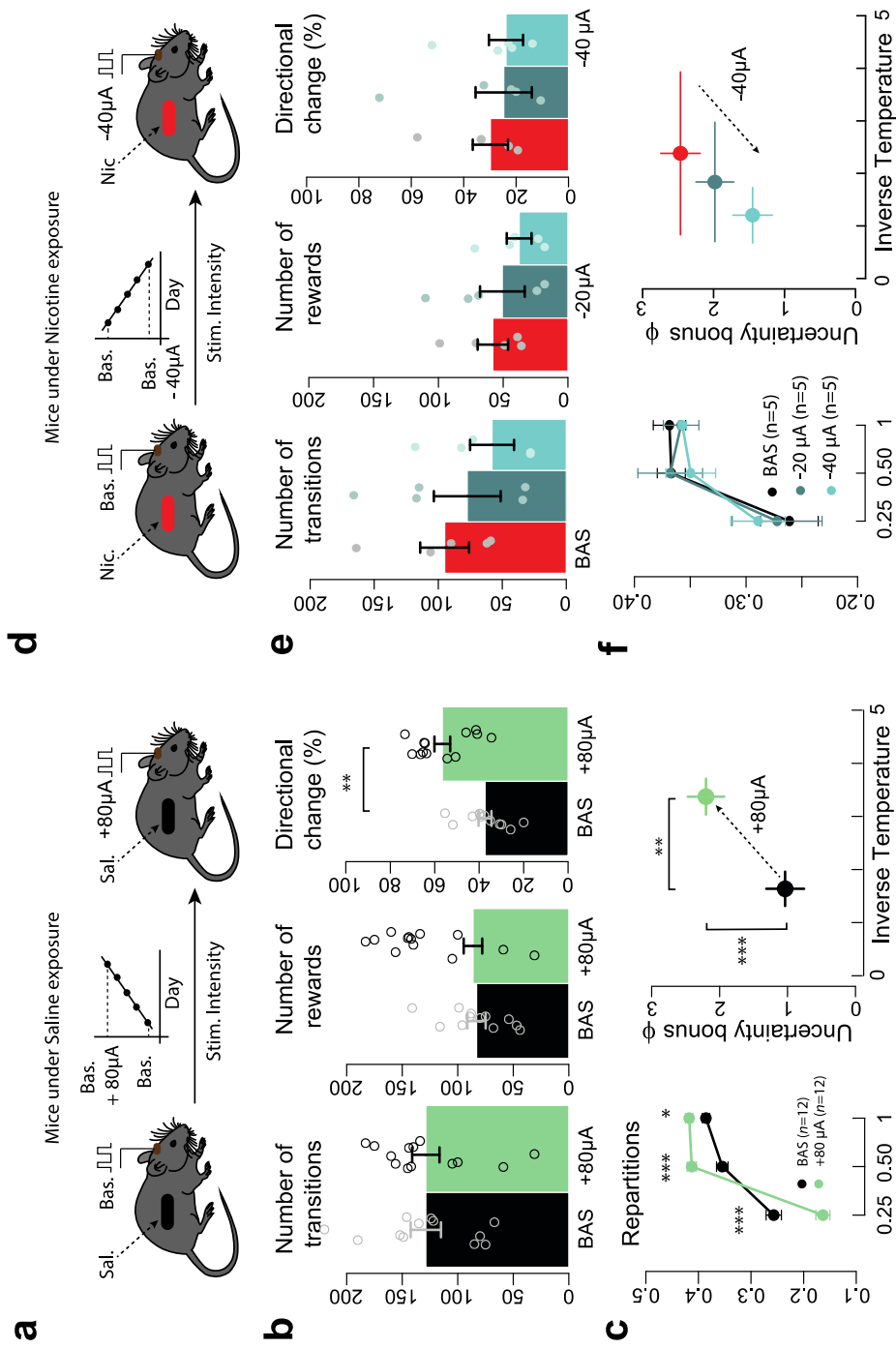


Figure 3

**Figure 4** *Juxtacellular recordings of VTA DA neurons in anaesthetized mice.*

- A.** VTA DA cell frequency and bursting activity (see trace) was compared between 2 sets of conditions: Saline minipump (N=123) versus Nicotine minipump (N=62) and Saline minipump + ICSS (N=78) versus Nicotine minipump + ICSS (N=76) after completion of the PS. All experiments were performed after 24-/+ 2 days of SAL/NIC (10mg/kg/day) exposure.
- B.** A difference in DA cell frequency (Hz) was observed between the SAL and NIC conditions (ANOVA-test,  $F_{(2,18)} = 21.5316$ ,  $p = 5.001e-06$ ), as well as after NIC + ICSS compared to SAL + ICSS (ANOVA-test,  $df = 72.4212$ ,  $p = 5.904e-16$ ). No interaction effect was observed ( $F_{(1,335)} = 1.0224$ ,  $p = 0.3127$ ).
- C.** A significant increase in DA bursting activity was observed after nicotine treatment (ANOVA test,  $F_{(1,335)} = 25.3972$ ,  $p = 7.649e-07$ ), compared to SAL in both groups. However, no effect was found due to iCSS (Anova test,  $F_{(1,335)} = 1.0224$ ,  $p = 0.3127$ ) and no interaction effect was observed (Anova test,  $F_{(1,335)} = 0.6544$ ,  $p = 0.4191$ ).

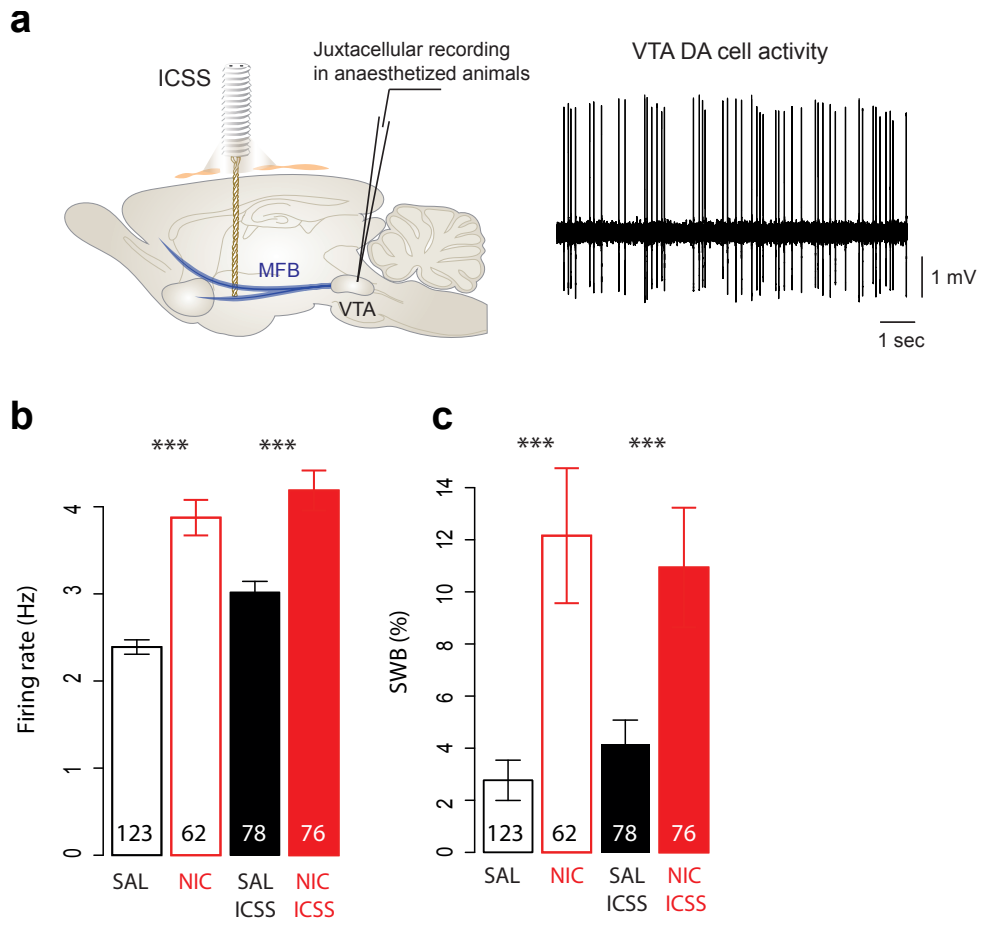


Figure 4

**Figure 5** *Optogenetic manipulation of VTA DA cells tonic activity.*

- A.** DAT-iCre mice (n=4) were injected with an AAV5.EF1a.DIO.CatCh.eYFP virus in the VTA and then simultaneously implanted with an optical fiber in the VTA and an ICSS electrode in the MFB (see methods section).
- B.** All YFP+ cells (in green, white arrows) were also positive to TH immune-staining (red) in midbrain slices preparation of DAT-iCre mouse injected the virus confirming the conditional expression of CatCh in DA cells. The orange arrow indicates a TH+/YFP- neuron.
- C.** 1s continuous photostimulation in VC mode (-60 mV) with blue light (460 nm) evoked an inward current in DANs (n=4, mean=240.9 pA +/- 20.82, left panel). In current-clamp mode, light-trains stimulation (10Hz, 5ms/pulse) evoked a firing with a 100% of spike/pulse (n=4 cells, right panel).
- D.** Post-hoc identification of a TH+/YFP+/NB+ DA neuron recorded in vivo (left). In vivo photostimulation trains with different frequencies (1Hz-10Hz, 5ms pulse) were associated with a probability of spike/ pulse close 1. 20 Hz trains stimulations decreased the probability of spike/pulse.
- E.** Mice were trained in both the DS and PS of the ICSS bandit task without photostimulation. After stable PS behaviour (10<sup>th</sup> day), photostimulation (8Hz, 5ms pulse, blue light:460nm) was delivered 5 min before and during the task on alternating days. 4 ON and 4 OFF days were not averaged for the 4 mice. Differences were found between the paired ON and OFF days for number of transitions (per transition: paired T-test,  $t = -2.911$ ,  $df = 16$ ,  $p=0.01021$  and number of rewards (paired T-test,  $t = -2.6213$ ,  $df = 16$ ,  $p=0.01852$ ).
- F.** No significant differences were found for the repartition to  $p=0.25$  (paired T-test,  $t = 1.9516$ ,  $df = 16$ ,  $p=0.06871$ ),  $p=0.50$  (paired T-test,  $t = -1.9238$ ,  $df = 16$ ,  $p=0.07236$ ), and  $p=1.00$  (paired T-test,  $t = -1.1572$ ,  $df = 16$ ,  $p=0.2642$ ), for paired ON and OFF days. The inverse temperature showed a difference (paired T-test,  $t=2.2111$ ,  $df=16$ ,  $p=0.04194$ ), but the uncertainty parameter did not (paired T-test,  $t=1.6204$ ,  $df=16$ ,  $p=0.1247$ ).

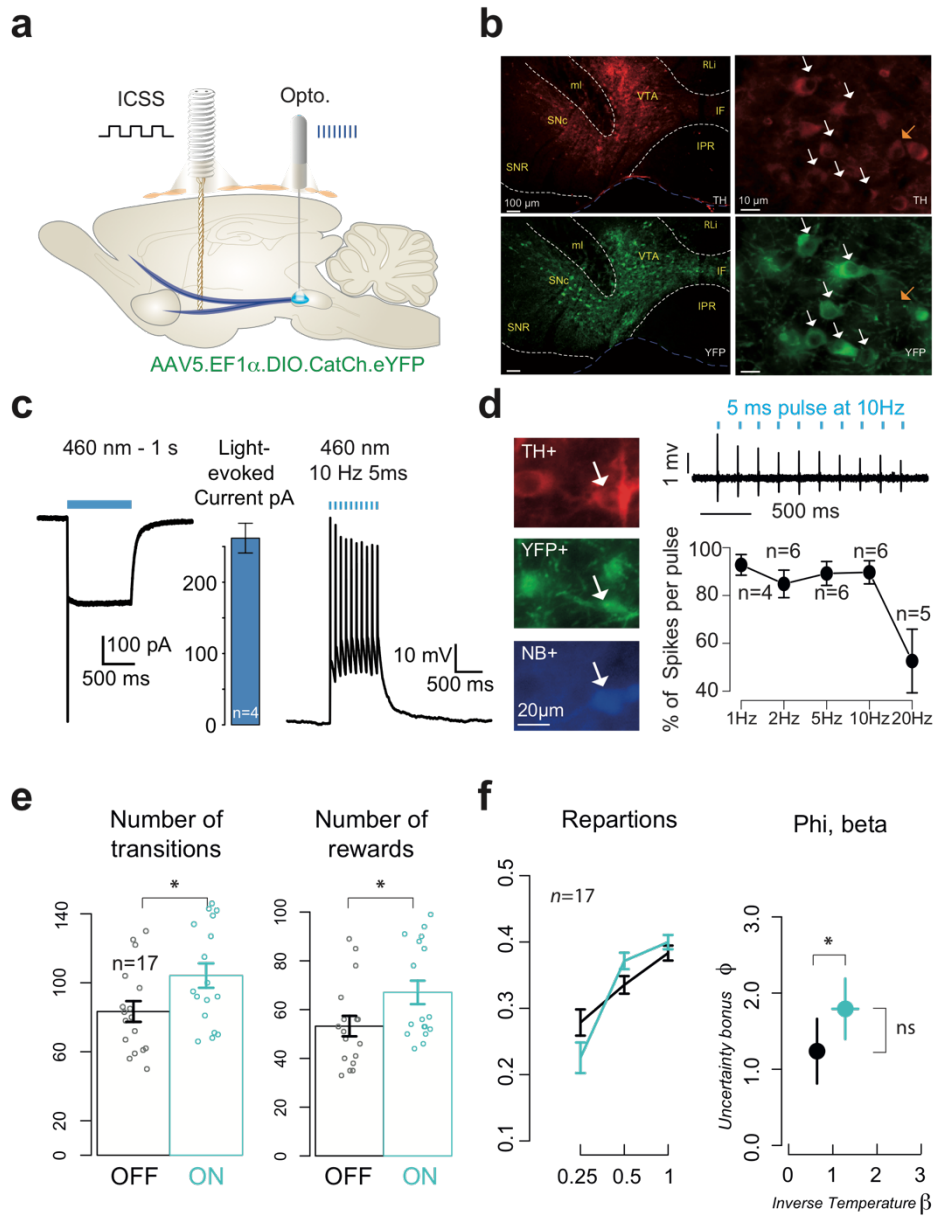


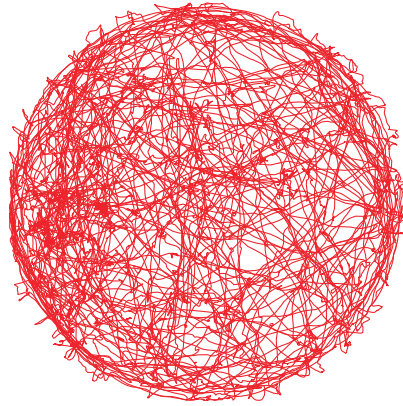
Figure 5



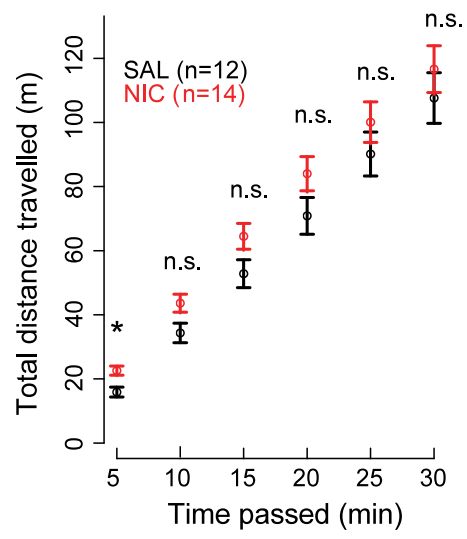
**Supplementary Figure 1** *Open field behaviour after long-term nicotine exposure.*

- A.** Trajectory of a naïve mouse in an open field (30 minutes) after 24 days of nicotine (10 mg/kg/day) treatment.
- B.** Total distance travelled in meters measured every 5 minutes during a 30 minutes' open field exploration. Nicotine mice (N=14) show a larger distance travelled at 5 (T-test,  $t = -2.4154$ ,  $df = 22.074$ ,  $p=0.02444$ ), but not at 10 (T-test,  $t = -1.9357$ ,  $df = 23.691$ ,  $p=0.06492$ ) and 15 minutes (T-test,  $t = -1.7824$ ,  $df = 23.983$ ,  $p=0.8736$ ), compared to Saline mice (N=12). Also, the final distance travelled after 30 minutes was not significantly different (T-test,  $t = -0.85234$ ,  $df = 22.06$ ,  $p=0.4032$ ).

**a**



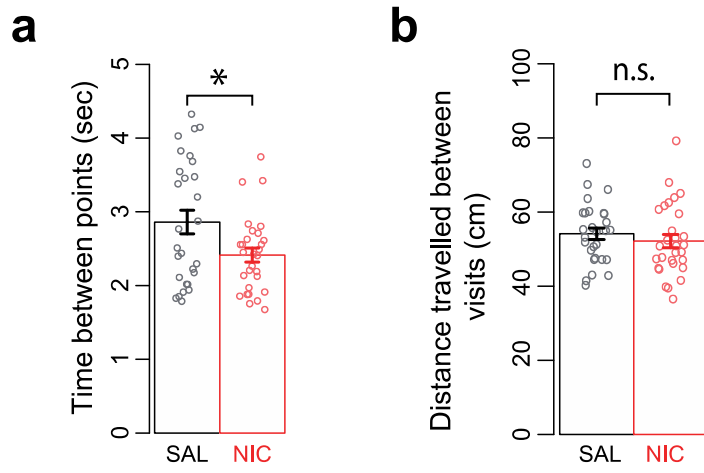
**b**



Supp Figure 1

**Supplementary Figure 2** *Distance travelled and travel time between locations in the PS.*

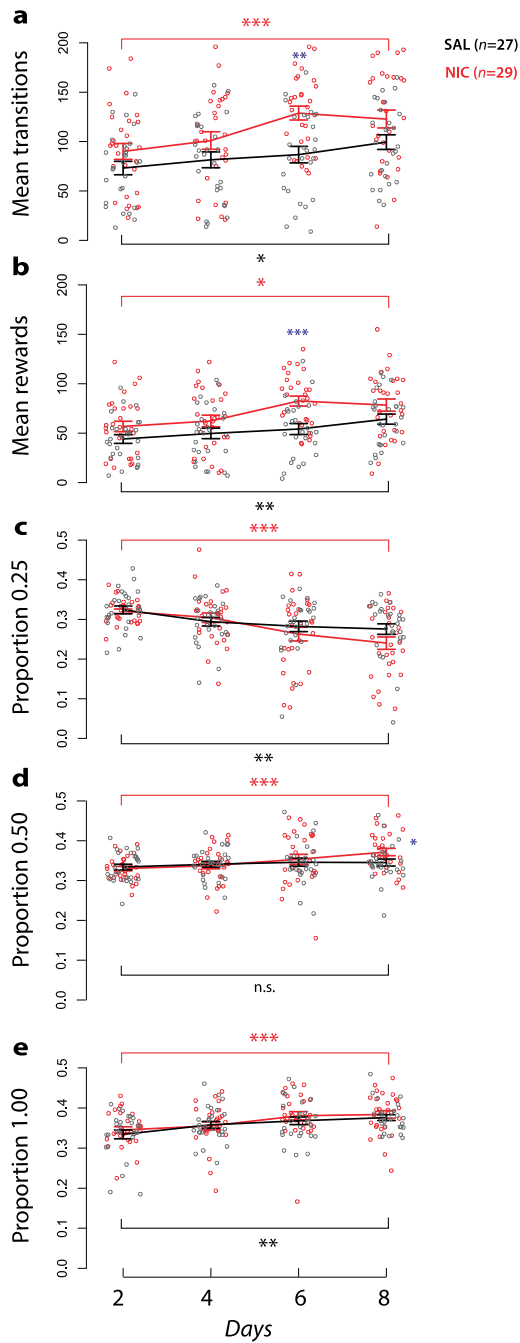
- A.** SAL mice (N=28) have a larger travel time on average between the different locations in the PS, compared to NIC mice (N=29), (T-test,  $t = 2.1074$ ,  $df = 52.622$ ,  $p=0.03987$ ).
- B.** No significant difference was found in the distance travelled on average between the points in the PS for the SAL and NIC mice (T-test,  $t = 1.038$ ,  $df = 51.79$ ,  $p=0.3041$ ).



Supp Figure 2

**Supplementary Figure 3** *Learning in the probabilistic setting.*

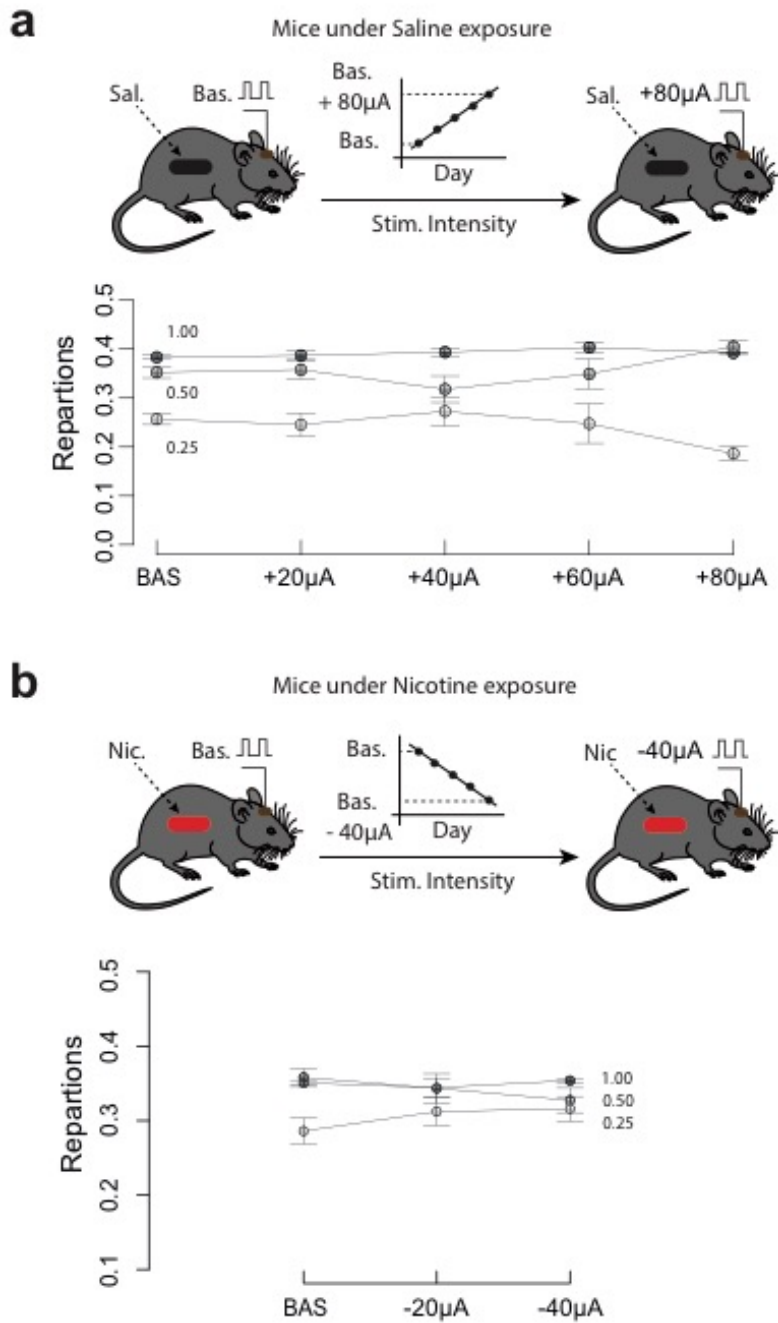
- A.** Mean number of transitions increased significantly between DAY 2 and DAY 8 of the PS for SAL mice (N=27, T-test,  $t = -2.6062$ ,  $df = 51.67$ ,  $p=0.01193$ ) and NIC mice (N=29, T-test,  $t = -2.7073$ ,  $df = 55.101$ ,  $p=0.009015$ ).
- B.** Mean number of rewards increased significantly between DAY 2 and DAY 8 of the PS for SAL mice (T-test,  $t = -3.007$ ,  $df = 50.71$ ,  $p=0.004098$ ) and NIC mice (T-test,  $t = -2.6487$ ,  $df = 54.675$ ,  $p=0.01054$ ).
- C.** The proportion of visits to the  $p=0.25$  significantly decreased between DAY 2 and DAY 8 of the PS for SAL mice (T-test,  $t = 2.9228$ ,  $df = 47.326$ ,  $p=0.005305$ ) and NIC mice (T-test,  $t = 4.8378$ ,  $df = 34.89$ ,  $p=2.634e-05$ ).
- D.** The proportion of visits to the  $p=0.50$  significantly increased between DAY 2 and DAY 8 of the PS for NIC mice (T-test,  $t = -3.9167$ ,  $df = 46.372$ ,  $p=0.0002931$ ), but not for SAL mice (T-test,  $t = -0.97836$ ,  $df = 49.422$ ,  $p=0.3327$ ).
- E.** The proportion of visits to the  $p=1.00$  significantly increased between DAY 2 and DAY 8 of the PS for SAL mice (T-test,  $t = -3.1962$ ,  $df = 44.508$ ,  $p=0.002562$ ) and NIC mice (T-test,  $t = -3.1108$ ,  $df = 54.362$ ,  $p=0.00297$ ).



Supp Figure 3

**Supplementary Figure 4** *Development of shifts in choice with changing ICSS.*

- A.** A gradual shift in choice behaviour implicating increased value-sensitivity is observed in accordance with the increase in ICSS value. Statistics were only applied on the comparison between baseline and +80 $\mu$ A (see Fig. 3).
- B.** A similar gradual shift towards decreased value-sensitivity was observed when the ICSS intensity was decreased for nicotine-treated mice. However, the sample size was too small to perform statistical analyses.



Supp Figure 4





# Discussion



## 9 | Perspectives

My PhD project focuses on nicotine-induced modifications of decision-making through the regulation of the dopaminergic system. Apart from the data I have presented in the previous chapters, I developed a new approach to understand such modifications in a more natural environment and to create a possibility to translate our findings to human behaviour. I have started to collect data with promising results that I will present here in the form of future directions.

### 9.1 Closed economy for mice

The multi-armed ICSS bandit task that was developed in our team has numerous advantages. Our protocol meets a series of requirements that have proven difficult to address altogether: simultaneous tracking of action selection and action execution in an environment with controlled rewards and exploration. And moreover, a possibility to eventually record electrophysiological activity. Furthermore data is used to fit reinforcement learning/decision-making models, allowing a refined analysis of the data. However, these methods also present specificity that can limit generalisation. In this task, animals are in a situation of gambling-like circumstances. In tasks in which rewards are small or far from satiety there could be short delays between serially repeated trials (the succession of gambles) and these settings induce preference for risk, defined as the expected variance of the reward [260]. In the multi-armed bandit task the ICSSs are easily obtainable rewards that are only divided by short distance. On the one hand this promotes motivation, on the other hand this could evoke impulsivity. Furthermore, ICSS is a very interesting, but peculiar reward that may have more effect on the brain system than we are aware of. For this reason, we aim at developing new strategies to test the generalisation of our results induced by nicotine. For that purpose, we have started to develop a closed economy paradigm:

A closed economy in a behavioural context relates to an environment in which mice are continuously "self-sufficient" in obtaining daily food and water, and in which behaviour

is measured automatically without interference of an experimenter [261]. We use classical operant boxes (Med. Associates) for mice with two levers and a food well in between them (see *Figure 9.1a*). The floor of the cage is adapted from grid to a solid aluminum plate and covered in a layer of bedding. Water is provided ad libitum. Mice have access to the two non-retractable levers continuously. This set up allows us to study neuroeconomical phenomena driven by natural motivation (no food restriction; no time restrains). We focus on value discrimination by reversing the location of the reinforcing lever. At any point during the experiment one lever is reinforced with one food pellet at a certain fixed ratio (FR) and the other lever is inactive. The reinforcing lever changes position every 24 hours or 96 hours. Experiments take 5-7 days and during this period the mouse is solitary housed in the operant box. Nicotine was administered acutely (I.P. injections, 500µg/kg), or in drinking water (see *Figure. 9.1*). Under the control conditions, we recorded VTA DA cell activity of WT mice over periods of 24 hours with in vivo electrophysiology (see *Figure 9.2*).

A first consequence of the continuous presence of food is that in a closed economy paradigm the initiation and termination of the consummatory behaviours are defined by the animal. There is no explicit motivating operation, other than the internal state of hunger [262]. For this reason, it is not necessary to use pre-session food or water restriction. In contrast, in an open economy protocol, the commodity may be earned only during the experimental sessions, which is time-limited. Furthermore, even if limited, additional food or water is usually offered outside of the economy, and the animal does not need to work for it. An important facet of closed economies is the study of behavioural patterns; how does an animal devote its time to various activities and how does the time budget relate to choice and effort? In our approach, we focus on value discrimination and quantified choice behaviour in a simple dynamic foraging task. In such a task, the animal has to continuously sample the environment and adjust its behaviour in response to changing rules to obtain rewards. It involves the presentation of 2 levers, one associated with the certain delivery (100 %) of rewards (1 pellet), the other with no delivery. So, at any point during the experiment one lever is reinforced with one food pellet at a certain fixed ratio (FR) and the other lever is inactive. The rewarding lever is changed every 24 hours. We thought that in this context we could evaluate impact of nicotine exposure on more "daily" decisions.

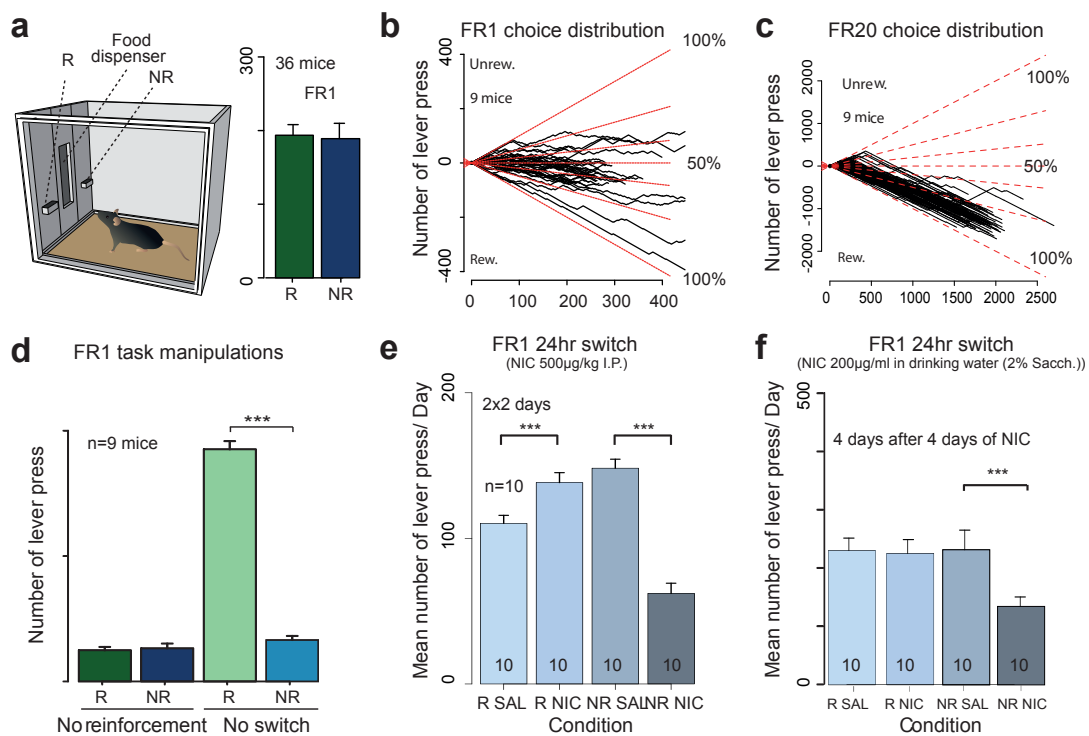


Figure 9.1 – **The effect of nicotine in a closed economy paradigm** **a:** (Left) Setup where animal live 24/24h (Right) Choice distribution between the rewarding (R) and non-rewarding (NR). **b: Sequence of choice in FR1** Mice alternate their choice between the R and NR lever. Cumulative representation of the sequence of choice in an individual mice. Abscissa represent the trial number. Zero indicates moment of the lever switch (R and NR). n=9 mice, 4 alternations per mice. Value of 1 is given for NR lever (Unrew. direction ordinate), -1 for R lever (Rew. direction ordinate). Red lines at top and bottom indicate a choice of 100 % of of respectively NR or R. Here the animal's choice is about 50%. **c: Sequence of choice in FR20** Mice had to press 20 times consecutively to obtain a reward. Mice clearly switch rapidly and focus on the R lever.**d: Lever press distribution** (left) when there is no food reinforcement, (Right) When only one lever is reinforced and there is no switch. **e: Choice distribution under chronic nicotine** with acute IP administration of nicotine **f: Choice distribution under chronic nicotine.** Nicotine was administered in drinking water

To summarize briefly, our first main aim was to demonstrate that in such a paradigm, the lever-press behavior of the mice and its decision-making to press the left or right lever first depends on the effort that is necessary to obtain food. On a FR1 schedule, animals alternate continuously between the rewarded and unrewarded (NR) lever (*see*

*Figure 9.1a-b*). In contrast, when we use a FR20 schedule, animals start to track the rewarded lever and actually discriminate between the levers. Recordings of VTA DA cell activity during the FR1 task reveals that DA cell activity show a phasic activity when a lever is pressed, reinforced or non-reinforced (*see Figure 9.2c*). Since, this corresponds with the non-specific lever presses under FR1, we hypothesize that phasic DA cell activity reflects either effort-dependent or learned reward association. Furthermore, we showed that when none of the levers was reinforced, mice do press significantly less, mimicking exploratory play behaviour. When only one lever is reinforced (i.e there is no daily alternation), mice mainly press the rewarded lever and mostly ignore the unrewarded one (*see Figure 9.1d*). This led us to conclude that under a FR1 schedule, there is motivation to press the lever to obtain food, but mice do not discriminate between the rewarding and non-rewarding lever. We used this behavioural situation to quantify the impact of nicotine in the same environment.

Mice were exposed to different nicotine regimens. Nicotine was administered acutely (I.P. injections, 500µg/kg) or in drinking water (200µg/ml, 2% Saccharine). In both cases (*see Figure 9.1e-f*) our results suggest that nicotine increases value sensitivity. The animal focuses more on the rewarding lever when exposed to nicotine. Overall, this behaviour is similar to what has been observed in the multi-armed ICSS bandit task. This paradigm also provides the possibility to address nicotine's effect on VTA DA cell dynamics.

Under control conditions, we recorded VTA DA cell activity of WT mice over periods of 24 hours with in vivo electrophysiology (*see Figure 9.2a*). In this condition mice were implanted with a double octrode in the VTA (AP=3.1; ML=+/-0.5, DV=4.15). A descending system was attached to the electrodes to change the DV in vivo. We were able to record event-specific DA activations (*see Figure 9.2b*), as well as DA activity levels during dark vs light phase (*see Figure 9.2c*) or during transition between activity (i.e when the mouse moves) and inactivity (characterised by absence of movement) (*see Figure 9.2d*). Our preliminary results suggest a fine-grained temporal modulation of the DA cells activity, a opportunity to link more specifically the state of the DA system to the animals behaviour and contribute to the ongoing debate of the roles of phasic and tonic DA activity [46]. Furthermore, it allows us to quantify the effect of nicotine and its dynamics in a natural environment.

*Overall these results tend to confirm our analyses in the multi-armed ICSS bandit*

*task; nicotine increases value sensitivity, also in a completely different decision-making situation.*



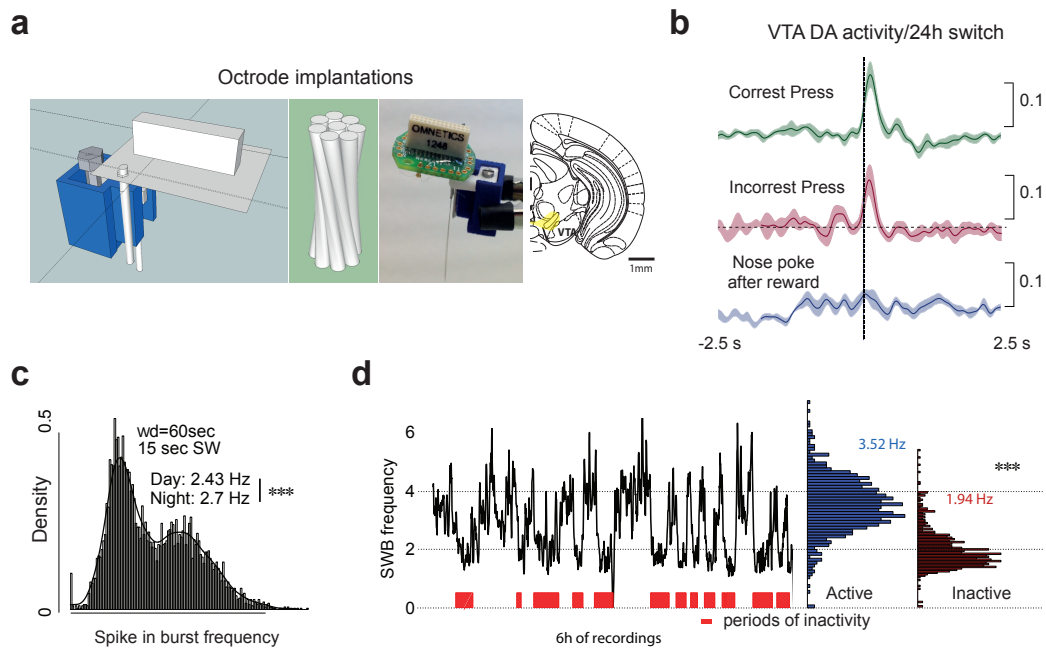


Figure 9.2 – VTA DA Cells recording in close economy paradigm **a**: Octrode implantation methods **b**: Mean activity of VTA DA cells centered around typical events: (Top) Rewarding lever (Middle) Non rewarding lever (Above) Nose Poke. **c**: **Bursting frequency distribution.** Note the bimodal curve. Day night indicate difference in the mean frequency during the Day and the night. **d**: **Bursting frequency fluctuation at a short time scale:** Difference in the bursting activity during active and inactive period

## 9.2 Online 3D Human bandit task

The main question in nicotine research remains how certain situations and findings translate to human behaviour? As a first trial, we have translated the multi-armed ICSS bandit task into a human version that is available online (<http://faurelab.cnrs.fr/three-points/>). We have collected a preliminary set of data (*see Figure 9.3*). When subjects enter the website they have the possibility to choose between three available languages (French, English and Dutch). After a short description of the study and a consent form, participants are directed to a demographic questionnaire in which 1) Age; 2) Gender; 3) Level of education and 5) smoke behaviour is asked. Participants are also asked if they have participated before. When a subject indicates to be a smoker (daily or casual) its forwarded to the Fagerström Questionnaire (as briefly discussed in the introduction, *see Figure 9.4*). Data of 79 (out of 126) participants was included in the preliminary data analyses. Data was excluded when the participant had indicated participation before or when data was incomplete or unreliable based on response times.

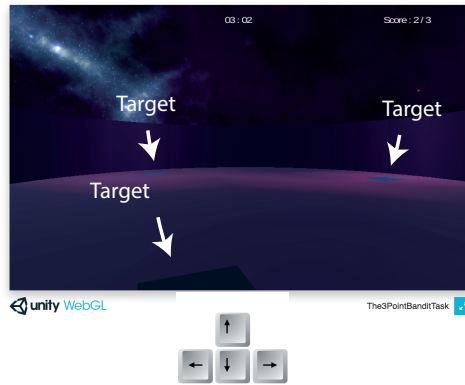
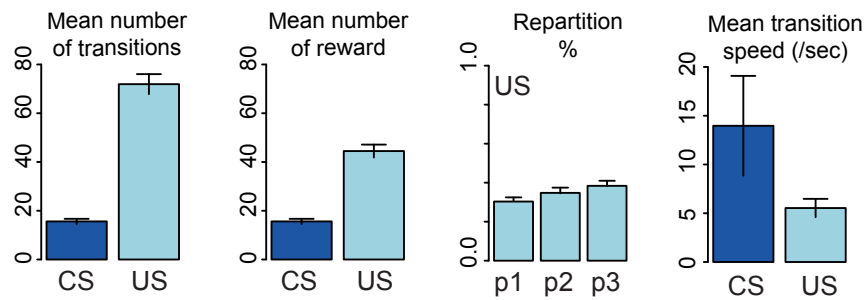

**a****b**

Figure 9.3 – **A.** Screenshot of the virtual "open-field" in which three targets must be visited to obtain rewards in the DS and PS. One can navigate with the arrow keys. **B.** Preliminary data showing that subjects familiarize with the task through increased speed; number of transitions

After the questionnaire participants were directed to the task. The task was developed in a Unity and takes place in a virtual representation of an open field. The task starts with a practice round of 1 minute in the deterministic setting. The participant is asked to navigate through the open field with the arrow keys of the keyboard. Three blue marks are indicated and visitation is rewarded with stars. After familiarization with the task, the probabilistic setting can be played. This setting takes 4 minutes. Similar to the mouse version, the three different marks are rewarded with different probabilities ( $p=0.25;0.50;1.00$ ).

We have created an online version of the task to collect data as a proof of concept and to fine tune the timings and test the feasibility of the game across ages. A large pitfall is the lack of actual rewards, which does not allow for a direct comparison with the mouse version, since the brain stimulation literally drives behaviour. Furthermore, participants cannot develop stable choice behaviour after just one session.

A future application of the human test would be to perform electroencephalography (EEG) recordings while participants virtually navigate through the task.


 **tabac-info-service.fr**  
Quand on sait, c'est plus facile d'arrêter

**ÉVALUATION DE LA DÉPENDANCE CHIMIQUE À LA NICOTINE** :

**QUESTIONNAIRE DE FAGERSTRÖM**

Combien de temps après votre réveil fumez-vous votre première cigarette ?	Dans les 5 premières minutes	3
	Entre 6 et 30 minutes	2
	Entre 31 et 60 minutes	1
	Après 60 minutes	0
Trouvez-vous difficile de vous abstenir de fumer dans les endroits où c'est interdit ?	Oui	1
	Non	0
À quelle cigarette de la journée renoncerez-vous le plus difficilement ?	La première le matin	1
	N'importe quelle autre	0
Combien de cigarettes fumez-vous par jour en moyenne ?	10 ou moins	0
	11 à 20	1
	21 à 30	2
	31 ou plus	3
Fumez-vous à un rythme plus soutenu le matin que l'après-midi ?	Oui	1
	Non	0
Fumez-vous lorsque vous êtes malade, au point de devoir rester au lit presque toute la journée ?	Oui	1
	Non	0
<b>Total</b>		

Tabac-info-service.fr est le site d'information et d'aide à l'arrêt du tabac du Ministère chargé de la Santé et de l'INPES.  
Pour toute question sur l'arrêt du tabac et pour bénéficier d'un suivi personnalisé gratuit par un tabacologue, appelez le 39 89\*.  
\* 0,15 €/min depuis un poste fixe, du lundi au samedi de 8h à 20h.

 **tabac-info-service.fr**  
Quand on sait, c'est plus facile d'arrêter

**DE FAÇON GÉNÉRALE, VOICI COMMENT INTERPRÉTER LE SCORE OBTENU PAR VOTRE PATIENT** :

**SCORE DE 0 À 2** :

Le sujet n'est pas dépendant à la nicotine. Il peut arrêter de fumer sans avoir recours à des substituts nicotiques. Si toutefois le sujet redoute l'arrêt, vous pouvez lui apporter des conseils utiles de type comportementaux (jeter les cendriers, boire un verre d'eau...). Vous pouvez également lui conseiller d'appeler Tabac Info Service au 39 89.

**SCORE DE 3 À 4** :

Le sujet est faiblement dépendant à la nicotine. Il peut arrêter de fumer sans avoir recours à un substitut nicotinique. En cas de manque ou de difficultés passagères (irritabilité, manque, envie très forte...) vous pouvez éventuellement lui conseiller de prendre un substitut nicotinique par voie orale (comprimé à sucer, gomme à mâcher, comprimé sublingual...). Si le sujet redoute l'arrêt, vous pouvez lui conseiller d'appeler Tabac Info Service au 39 89 pour recevoir des conseils et du soutien.

**SCORE DE 5 À 6** :

Le sujet est moyennement dépendant. L'utilisation des traitements pharmacologiques de substitution nicotinique va augmenter ses chances de réussite. Vos conseils seront utiles pour l'aider à choisir la galénique la plus adaptée à son cas.

**SCORE DE 7 À 10** :

Le sujet est fortement ou très fortement dépendant à la nicotine. L'utilisation de traitements pharmacologiques est recommandée (traitement nicotinique de substitution ou bupropion LP ou varenicline). Ce traitement doit être utilisé à dose suffisante et adaptée. En cas de difficulté, orienter le patient vers une consultation spécialisée.

Tabac-info-service.fr est le site d'information et d'aide à l'arrêt du tabac du Ministère chargé de la Santé et de l'INPES.  
Pour toute question sur l'arrêt du tabac et pour bénéficier d'un suivi personnalisé gratuit par un tabacologue, appelez le 39 89\*.  
\* 0,15 €/min depuis un poste fixe, du lundi au samedi de 8h à 20h.

Figure 9.4 – Fagerström Questionnaire [6]

## 10 | Nicotine-induced alterations in value-based decision-making

The two presented papers display two different implications of the nicotinic cholinergic system in value-based decision-making through alterations in the VTA. First, the inactivation and selective re-expression of  $\beta 2$  subunit-containing receptors in the VTA led to the discovery of the implication of VTA  $\beta 2$  subunits in the computation of expected uncertainty. Secondly, the systemic chronic administration of nicotine led to the finding that increased activation of VTA DA neurons promotes enhanced value sensitivity, a behavioural characteristic that was mimicked by photostimulation of VTA DA neurons. As discussed in chapter 2 and 3, nicotine and cholinergic input act on different locations within the mesocorticolimbic dopaminergic system, as well as different locations specifically within the VTA. Through nAChRs, agonists can alter DA activity by playing directly onto the DA cell bodies or through GABA-ergic and glutamatergic VTA interneurons. Furthermore, VTA signaling can be modified by feedback arriving from among others, the VTA projections and mesopontine cholinergic, glutamatergic and GABAergic inputs. In other words, the pathways to and from the VTA form a complex dynamic network that undergoes modifications on different levels in KO animals or when drugs are applied.

The main outcome of the presented work suggests that nicotine increases value sensitivity through nicotine-induced VTA dopamine activity. A growing number of studies have reported a link between addictive substances and drug-induced aberrant decision-making. However, the current literature is inconsistent on how these adverse behavioural adaptations contribute to the different stages of addiction (initiation, maintenance, cessation, [263]). Since the dopaminergic "reward" system is so inherent to the understanding of addiction it is not surprising that much research has been carried out assessing reward sensitivity in the attempt to describe the underlying effects of nicotine. However, a restrictive scope on addiction-related behaviours, such as drug-seeking and consumption, rules out the basic comprehension of global drug-induced neurophysiological and behavioural modifications that may underlie and maintain addiction. Global drug-induced

brain-plasticity and its impact on daily life behaviours are not yet completely understood. However, this broader perspective on addiction is slowly more widely accepted.

Evidence from animal studies suggests that nicotine alters the brain's threshold for ICSS brain stimulation. Huston-Lyons and Kornetsky (1991) demonstrated how the reward threshold of brain stimulation was lowered by nicotine and that this effect could be blocked by the nicotine receptor antagonist, mecamylamine [264]. These findings were confirmed in a study of Kenny and Markou (2006), in which was found that the administration of nicotine (0.03 mg/kg) in rats lowered the reward thresholds. This effect was reversed by the nicotinic receptor antagonist dihydro- $\beta$ -erythroidine [265]. Furthermore, even after nicotine self-administration had ceased nicotine enhanced sensitivity of the brain reward system, although similar experiments displayed that nicotine withdrawal elevated brain stimulation thresholds [24,265]. A straight-forward conclusion is that nicotine increases the sensitivity for brain stimulation, and therefore reward sensitivity. If the physiological effect of repetitive ICSS stimulation on DAN activity modulates this behaviour may be debatable.

Additionally, behavioural work has shown that nicotine has reinforcement-enhancing properties such that nicotine promotes behaviours that result in the delivery of non-pharmacological reinforcers [16,18]. Similarly, behavioural evidence from human studies has demonstrated that a single transdermal dose of nicotine enhances responding to non-drug rewards in non-smokers [266]. These findings suggest that nicotine increases the salience of available stimuli, which results in an incentive to increase response actions.

Our aim was to take this investigation into increased salience and reward sensitivity a step further to study whether it is legitimate to speak of reward sensitivity (change in reward perception) or if we actually deal with an increased contrast (value sensitivity) between available options. In an all-or-nothing paradigm, an increased contrast, can yield impulsive, reward sensitive or risk-taking behaviour. However, when different values (probabilities or magnitudes) are presented over several available actions, it is possible to study how the animal discriminates between different values when it organises its choices. Nicotine favours selection of high rewards and thus animals tend to become exploitive. We

demonstrated this effect in the multi-armed ICSS bandit task, but also in the home-cage environment. However, even in these paradigms we are not able to clearly distinguish between reward perception and value sensitivity (i.e. the politic).

Nicotine and nicotinic receptors can certainly influence a number of decision making parameters (see for example [267], or uncertainty [?]). In contrast to reward sensitivity or value sensitivity one could also suggest that nicotine changes the balance between positive and negative reinforcement due to punishment sensitivity and altered negative feedback processing. This would imply that instead of focusing on the highest reward value available the focus is merely on ignoring the lowest values or aversive consequences. The effect of nicotine on punishment sensitivity and on punishment schedules has been investigated in a small number of animal studies. Responding on a punishment schedule can be maintained by concurrent reinforcement and it is well known that drugs can modify operant responses that are both rewarded and punished simultaneously. For example, minor tranquilisers restore responses which are suppressed by punishment, whereas stimulant drugs such as amphetamine further reduce responses that have already been suppressed by punishment, but increase operant responding when there is no punishment contingency [268]. Morrison (1969) looked at the effects of 0.1 and 0.4mg/kg nicotine on punished behaviour and found that in most animals nicotine further suppressed responding that was suppressed by the electric shock, though it did increase responding in some animals. Since the effect of nicotine in this study was variable and small in magnitude, it was suggested that the conclusions drawn were speculative and that further work should be carried out.

Overall, the development of novel neuroeconomical paradigms should be stimulated to address questions that remain unanswered. The multi-armed ICSS bandit task allows several adaptations [?] that could provide knowledge about behavioural flexibility under nicotine. Also, the closed economy operant box paradigm offers several options. In general, we believe that this work is a welcomed addition to the field of addiction research and with supplementary cellular techniques could reveal even more about the underlying mechanisms of changes in a nicotine addicted brain.





# Bibliography

- [1] M J Stolerman, I P & Jarvis. The scientific case that nicotine is addictive. *Psychopharmacology*, 117(1):2–20, 1995.
- [2] J S Fowler, N D Volkow, G J Wang, N Pappas, J Logan, C Shea, D Alexoff, R R MacGregor, D J Schlyer, I Zezulkova, and A P Wolf. Brain monoamine oxidase A inhibition in cigarette smokers. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24):14065–9, 1996.
- [3] Fabio Marti, Ouafa Arib, Carole Morel, Virginie Dufresne, Uwe Maskos, Pierre-Jean Corringer, Renaud de Beaurepaire, and Philippe Faure. Smoke extracts and nicotine, but not tobacco extracts, potentiate firing and burst activity of ventral tegmental area dopaminergic neurons in mice. *Neuropsychopharmacology*, 36(11):2244–57, oct 2011.
- [4] U.S. Department of Health and Human Services. *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease*. 2010.
- [5] Denise B Kandel, Mei-chen Hu, Pamela C Griesler, and Christine Schaffran. On the development of nicotine dependence in adolescence. *Drug and Alcohol Dependence*, 91:26–39, 2007.
- [6] Todd F Heatherton, Lynn T Kozlowski, Richard C Frecker, and Karl Olov Fagerstrom. The Fagerstrom Test for Nicotine Dependence: A revision of the Fagerström Tolerance Questionnaire. *British Journal of Addiction*, 86(9):1119–1127, 1991.
- [7] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: Dsm-5*. 2013.
- [8] Jed. E. Rose and Edward D. Levin. Inter-relationships between conditioned and primary reinforcement in the maintenance of cigarette smoking. *British Journal of Addiction*, 86(5):605–609, 1991.
- [9] Keith Payne, F Joseph McClernon, and Ian G Dobbins. Automatic affective responses to smoking cues. *Experimental and clinical psychopharmacology*, 15(4):400–409, 2007.
- [10] Cynthia A. Conklin and Stephen T. Tiffany. Applying extinction research and theory to cue-exposure addiction treatments. *Addiction*, 97(2):155–167, 2002.
- [11] Anthony R. Caggiula, Eric C. Donny, Anthony R. White, Nadia Chaudhri, Sheri Booth, Maysa A. Gharib, Alycia Hoffman, Kenneth A. Perkins, and Alan F. Sved. Cue dependency of nicotine self-administration and smoking. *Pharmacology Biochemistry and Behavior*, 70(4):515–530, 2001.

- [12] Michael A. Sayette, Christopher S. Martin, Joan M. Wertz, Saul Shiffman, and Michael A. Perrott. A multi-dimensional analysis of cue-elicited craving in heavy smokers and tobacco chippers. *Addiction*, 96(10):1419–1432, 2001.
- [13] Timothy B Baker, Thomas H Brandon, and Laurie Chassin. Motivational influences on cigarette smoking. *Annual review of psychology*, 55:463–91, jan 2004.
- [14] Eric C. Donny, Nadia Chaudhri, Anthony R. Caggiula, F. Fay Evans-Martin, Sheri Booth, Maysa A. Gharib, Laure A. Clements, and Alan F. Sved. Operant responding for a visual reinforcer in rats is enhanced by noncontingent nicotine: Implications for nicotine self-administration and reinforcement. *Psychopharmacology*, 169(1):68–76, 2003.
- [15] Anthony R. Caggiula, Eric C. Donny, Nadia Chaudhri, Kenneth A. Perkins, F. Fay Evans-Martin, and Alan F. Sved. Importance of nonpharmacological factors in nicotine self-administration. *Physiology and Behavior*, 77(4-5):683–687, 2002.
- [16] Matthew I. Palmatier, F. Fay Evans-Martin, Alycia Hoffman, Anthony R. Caggiula, Nadia Chaudhri, Eric C. Donny, Xiu Liu, Sherri Booth, Maysa Gharib, Laure Craven, and Alan F. Sved. Dissociating the primary reinforcing and reinforcement-enhancing effects of nicotine using a rat self-administration paradigm with concurrently available drug and environmental reinforcers. *Psychopharmacology*, 184(3-4):391–400, 2006.
- [17] Matthew I Palmatier, Xiu Liu, Gina L Matteson, Eric C Donny, R Anthony, and Alan F Sved. Conditioned reinforcement in rats established with self-administered nicotine and enhanced by noncontingent nicotine. *Psychopharmacology*, 195(2):235–243, 2010.
- [18] Nadia Chaudhri, Anthony R. Caggiula, Eric C. Donny, Matthew I. Palmatier, Xiu Liu, and Alan F. Sved. Complex interactions between nicotine and nonpharmacological stimuli reveal multiple roles for nicotine in reinforcement. *Psychopharmacology*, 184(3-4):353–366, 2006.
- [19] Christie D. Fowler and Paul J. Kenny. Intravenous nicotine self-administration and cue-induced reinstatement in mice: Effects of nicotine dose, rate of drug infusion and prior instrumental training. *Neuropharmacology*, 61(4):687–698, 2011.
- [20] Vincent David, Morgane Besson, Jean Pierre Changeux, Sylvie Granon, and Pierre Cazala. Reinforcing effects of nicotine microinjections into the ventral tegmental area of mice: Dependence on cholinergic nicotinic and dopaminergic D1 receptors. *Neuropharmacology*, 50(8):1030–1040, 2006.
- [21] William A. Corrigall and Kathleen M. Coen. Nicotine maintains robust self-administration in rats on a limited-access schedule. *Psychopharmacology*, 99(4):473–478, 1989.
- [22] Shelly S. Watkins, Mark P. Epping-Jordan, George F. Koob, and Athina Markou. Blockade of nicotine self-administration with nicotinic antagonists in rats. *Pharmacology Biochemistry and Behavior*, 62(4):743–751, 1999.
- [23] M. R. Picciotto, M. Zoli, R. Rimondini, C. Léna, E. M. Marubio, L. M. and Pich, and J. P. Changeux. Acetylcholine receptors containing the beta2 subunit are involved in the reinforcing properties of nicotine. *Nature*, 391(6663):173, 1998.
- [24] M P Epping-Jordan, S S Watkins, G F Koob, and A Markou. Dramatic decreases in brain reward function during nicotine withdrawal. *Nature*, 393(6680):76–9, 1998.

- [25] U. Maskos, B. E. Molles, S. Pons, M. Besson, B. P. Guiard, J.-P. Guilloux, A. Evrard, P. Cazala, A. Cormier, M. Mameli-Engvall, N. Dufour, I. Cloëz-Tayarani, A.-P. Bemelmans, J. Mallet, A. M. Gardier, V. David, P. Faure, S. Granon, and J.-P. Changeux. Nicotine reinforcement and cognition restored by targeted expression of nicotinic receptors. *Nature*, 436(7047):103–107, 2005.
- [26] R Exley and S J Cragg. Presynaptic nicotinic receptors: a dynamic and diverse cholinergic filter of striatal dopamine neurotransmission. *British journal of pharmacology*, 153 Suppl:S283–97, mar 2008.
- [27] S Tolu, R Eddine, F Marti, V David, M Graupner, S Pons, M Baudonnat, M Husson, M Besson, C Reperant, J Zemdegs, C Pages, Y A Hay, B Lambolez, J Caboche, B Gutkin, A M Gardier, J P Changeux, P Faure, and U Maskos. Co-activation of VTA DA and GABA neurons mediates nicotine reinforcement. *Mol Psychiatry*, 18(3):382–393, 2013.
- [28] John A. Dani and Steve Heinemann. Molecular and cellular aspects of nicotine abuse. *Neuron*, 16(5):905–908, 1996.
- [29] D. J.K. Balfour. The psychobiology of nicotine dependence. *European Respiratory Review*, 17(110):172–181, 2008.
- [30] G. Di Chiara and A. Imperato. Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences*, 85(14):5274–5278, 1988.
- [31] G Di Chiara. Role of dopamine in the behavioural actions of nicotine related to addiction. *European journal of pharmacology*, 393(1-3):295–314, mar 2000.
- [32] Paul E. M. Phillips, Garret D. Stuber, Michael L. A. V. Heien, R. Mark Wightman, and Regina M. Carelli. Subsecond dopamine release promotes cocaine seeking. *Nature*, 423(6938):461–461, 2003.
- [33] Ethan S Bromberg-Martin, Masayuki Matsumoto, and Okihide Hikosaka. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, 68(5):815–34, dec 2010.
- [34] Jeremiah Y Cohen, Sebastian Haesler, Linh Vong, Bradford B Lowell, and Naoshige Uchida. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, pages 1–6, January 2012.
- [35] B I Hyland, J N J Reynolds, J Hay, C G Perk, and R Miller. Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience*, 114(2), 2002.
- [36] Mathias Pessiglione, Ben Seymour, Guillaume Flandin, Raymond J Dolan, and Chris D Frith. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106):1042–5, aug 2006.
- [37] Wolfram Schultz. Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, 80(1):1–27, 1998.
- [38] Richard S. Sutton and Andrew G. Barto. Reinforcement learning. *Learning*, 3(9):322, 2012.
- [39] David Redish. Addiction as a computational process gone awry. *Science*, 306(5703):1944–7, dec 2004.

- [40] Ronald Keiflin and Patricia H. Janak. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*, 88(2):247–263, 2015.
- [41] W Schultz. Dopamine neurons and their role in reward mechanisms. *Current opinion in neurobiology*, 7(2):191–7, apr 1997.
- [42] K C Berridge and T E Robinson. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain research. Brain research reviews*, 28(3):309–69, dec 1998.
- [43] Shelly B Fligel, Jeremy J Clark, Terry E Robinson, Leah Mayo, Alayna Czuj, Christina a Akers, Sarah M Clinton, Paul E M Phillips, and Huda Akil. A selective role for dopamine in reward learning. *Nature*, 469(7328):53–57, 2011.
- [44] Trevor W. Robbins and Barry J. Everitt. Functions of dopamine in the dorsal and ventral striatum. *Seminars in Neuroscience*, 4(2):119–127, 1992.
- [45] Saleem M Nicola. The flexible approach hypothesis: unification of effort and cue responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *J. Neurosci.*, (49):16585–16600.
- [46] Arif A Hamid, Jeffrey R Pettibone, Omar S Mabrouk, Vaughn L Hetrick, Robert Schmidt, Caitlin M Vander Weele, Robert T Kennedy, Brandon J Aragona, and Joshua D Berke. Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1):117–126, 2015.
- [47] M. F. Roitman. Dopamine Operates as a Subsecond Modulator of Food Seeking. *Journal of Neuroscience*, 24(6):1265–1271, 2004.
- [48] Mark W. Howe, Patrick L. Tierney, Stefan G. Sandberg, Paul E. M. Phillips, and Ann M. Graybiel. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, 500(7464):575–579, aug 2013.
- [49] Mike J F Robinson, Terry E. Robinson, and Kent C. Berridge. Incentive Salience and the Transition to Addiction. *Biological Research on Addiction*, (1):391–399, 2013.
- [50] John E Kelsey and Anzela Niraula. Effects of acute and sub-chronic nicotine on impulsive choice in rats in a probabilistic delay-discounting task. *Psychopharmacology*, 227(3):385–92, jun 2013.
- [51] Yu Ohmura, Taiki Takahashi, and Nozomi Kitamura. Discounting delayed and probabilistic monetary gains and losses by smokers of cigarettes. *Psychopharmacology*, 182(4):508–15, nov 2005.
- [52] El Lesage, S E Aronson, M T Sutherland, T J Ross, B Jo Salmeron, and El A Stein. Neural Signatures of Cognitive Flexibility and Reward Sensitivity Following Nicotinic Receptor Stimulation in Dependent Smokers A Randomized Trial. *JAMA Psychiatry*, pages 1–9, 2017.
- [53] F M Benes. Carlsson and the discovery of dopamine. *Trends in pharmacological sciences*, 22(1):46–47, 2001.
- [54] J Olds and P Milner. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology*, 47(6):419–427, December 1954.

- [55] James Olds. Pleasure Centers' in the Brain. *Sci Am*, 95(4):105–116, 1956.
- [56] Burrhus Frederic Skinner. The Behavior of Organisms: An experimental analysis. *The Psychological Record*, page 486, 1938.
- [57] R. Wise and PP. Rompre. Brain Dopamine And Reward. *Annual Review of Psychology*, 40(1):191–225, 1989.
- [58] P a Garris, M Kilpatrick, M a Bunin, D Michael, Q D Walker, and R M Wightman. Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature*, 398(6722):67–69, 1999.
- [59] R Wise. Addictive drugs and brain stimulation reward. *Current opinion in neurobiology*, 6(2):243–51, apr 1996.
- [60] L. W. Swanson. The projections of the ventral tegmental area and adjacent regions: A combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Research Bulletin*, 9(1-6):321–353, 1982.
- [61] G J Mogenson, D L Jones, and C Y Yim. From motivation to action: functional interface between the limbic system and the motor system. *Progress in neurobiology*, 14(2-3):69–97, 1980.
- [62] Suzanne N. Haber. The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy*, 26(4):317–330, dec 2003.
- [63] Francesco Pistillo, Francesco Clementi, Michele Zoli, and Cecilia Gotti. Nicotinic, glutamatergic and dopaminergic synaptic transmission and plasticity in the mesocorticolimbic system: Focus on nicotine effects. *Progress in Neurobiology*, 124:1–27, 2015.
- [64] Daniel Dautan, Albert S Souza, Icnelia Huerta-Ocampo, Miguel Valencia, Maxime Assous, Ilana B Witten, Karl Deisseroth, James M Tepper, J Paul Bolam, Todor V Gerdjikov, and Juan Mena-Segovia. Segregated cholinergic transmission modulates dopamine neurons integrated in distinct functional circuits. *Nature Neuroscience*, 19(8):1025–1033, 2016.
- [65] Marisela Morales and Elyssa B. Margolis. Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. *Nature Reviews Neuroscience*, 2017.
- [66] Cyriel M A Pennartz, Henk J. Groenewegen, and Fernando H. Lopes da Silva. The nucleus accumbens as a complex of functionally distinct neuronal ensembles: An integration of behavioural, electrophysiological and anatomical data. *Progress in Neurobiology*, 42(6):719–761, 1994.
- [67] J E Lachowicz and D R Sibley. Molecular characteristics of mammalian dopamine receptors. *Pharmacology & toxicology*, 81(3):105–113, 1997.
- [68] Christina Missale, Russel S Nash, Susan W. Robinson, Mohamed Jaber, and Marc G. Caron. Dopamine Receptors: From Structure to Function. *Physiol Rev*, 78(1):189–225, 1998.
- [69] X Zhuang, L Belluscio, and R Hen. G(olf)alpha mediates dopamine D1 receptor signaling. *The Journal of neuroscience*, 20(16):RC91, 2000.
- [70] L Q Jin, H Y Wang, and E Friedman. Stimulated D(1) dopamine receptors couple to multiple Galpha proteins in different brain regions. *Journal of neurochemistry*, 78(5):981–990, 2001.

- [71] Akinori Nishi, Mahomi Kuroiwa, and Takahide Shuto. Mechanisms for the Modulation of Dopamine D1 Receptor Signaling in Striatal Neurons. *Frontiers in Neuroanatomy*, 5(July):43, 2011.
- [72] S. L. Vincent, Y. Khan, and F. M. Benes. Cellular colocalization of dopamine D1 and D2 receptors in rat medial prefrontal cortex. *Synapse*, 19(2):112–120, 1995.
- [73] Susan R Sesack and Anthony a Grace. Cortico-Basal Ganglia reward network: microcircuitry. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 35(1):27–47, jan 2010.
- [74] Satoshi Ikemoto, Bradley S. Glazier, James M. Murphy, and William J. McBride. Role of Dopamine D1 and D2 Receptors in the Nucleus Accumbens in Mediating Reward. *J. Neurosci.*, 17(21):8580–8587, 1997.
- [75] Chun Yun Chang, Guillem R Esber, Yasmin Marrero-Garcia, Hau-Jie Yau, Antonello Bonci, and Geoffrey Schoenbaum. Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nature neuroscience*, 19(December):1–8, 2015.
- [76] Wolfram Schultz. Multiple dopamine functions at different time courses. *Annual review of neuroscience*, 30:259–288, January 2007.
- [77] Elyssa B Margolis, Hagar Lock, Gregory O Hjelmstad, and Howard L Fields. The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *The Journal of Physiology*, 577(Pt 3):907–924, December 2006.
- [78] Mark A. Ungless and Anthony A. Grace. Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends in Neurosciences*, 35(7):422–430, 2012.
- [79] A Grace and S Bunney. The control of firing neurons: Single pattern firing in nigral dopamine. *Journal of Neuroscience*, 4(11):2866–2876, 1984.
- [80] Anthony a Grace, Stan B Floresco, Yukiori Goto, and Daniel J Lodge. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in neurosciences*, 30(5):220–7, may 2007.
- [81] Mark a Ungless, Peter J Magill, and J Paul Bolam. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science (New York, N.Y.)*, 303(5666):2040–2, mar 2004.
- [82] Louis A. Chiodo. Dopamine-containing neurons in the mammalian central nervous system: Electrophysiology and pharmacology. *Neuroscience and Biobehavioral Reviews*, 12(1):49–91, 1988.
- [83] T Y Zhang and M J Meaney. Epigenetics and the environmental regulation of the genome and its function. *Annual Review of Psychology*, 61:439–66, C1–3, 2010.
- [84] Stephan Lammel, Andrea Hetzel, Olga Häckel, Ian Jones, Birgit Liss, and Jochen Roeper. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron*, 57(5):760–73, mar 2008.

- [85] Richard Exley, Nicolas Maubourguet, Vincent David, Raphaël Eddine, Alexis Evrard, Stéphanie Pons, Fabio Marti, Sarah Threlfell, Pierre Cazala, J Michael McIntosh, Jean-Pierre Changeux, Uwe Maskos, Stephanie J Cragg, and Philippe Faure. Distinct contributions of nicotinic acetylcholine receptor subunit  $\alpha 4$  and subunit  $\alpha 6$  to the reinforcing effects of nicotine. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7577–7582, 2011.
- [86] Raphaël Eddine, Sebastien Valverde, Stefania Tolu, Daniel Dautan, Audrey Hay, Carole Morel, Yihui Cui, Bertrand Lambolez, Laurent Venance, Fabio Marti, and Philippe Faure. A concurrent excitation and inhibition of dopaminergic subpopulations in response to nicotine. *Scientific Reports*, 5:8184, 2015.
- [87] P Faure, S Tolu, S Valverde, and J Naudé. Role of nicotinic acetylcholine receptors in regulating dopamine neuron activity. *Neuroscience*, 282:86–100, jun 2014.
- [88] A A Grace and B S Bunney. Intracellular and extracellular electrophysiology identification and characterization. *Journal of Neuroscience*, 10(2), 1983.
- [89] B S Bunney and A A Grace. Acute and chronic haloperidol treatment: comparison of effects on nigral dopaminergic cell activity. *Life sciences*, 23(16):1715–1727, October 1978.
- [90] Stan B Floresco, Anthony R West, Brian Ash, Holly Moore, and Anthony A Grace. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature neuroscience*, 6(9):968–73, 2003.
- [91] S E Hyman and R C Malenka. Addiction and the brain: the neurobiology of compulsion and its persistence. *Nature reviews. Neuroscience*, 2(10):695–703, oct 2001.
- [92] Michela Marinelli and James E McCutcheon. Heterogeneity of dopamine neuron activity across traits and states. *Neuroscience*, (July), jul 2014.
- [93] Vaishnav Krishnan, Ming-Hu Han, Danielle L Graham, Olivier Berton, William Renthal, Scott J Russo, Quincey Laplant, Ami Graham, Michael Lutter, Diane C Lagace, Subroto Ghose, Robin Reister, Paul Tannous, Thomas a Green, Rachael L Neve, Sumana Chakravarty, Arvind Kumar, Amelia J Eisch, David W Self, Francis S Lee, Carol a Tamminga, Donald C Cooper, Howard K Gershenfeld, and Eric J Nestler. Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell*, 131(2):391–404, oct 2007.
- [94] Jacques Barik, Fabio Marti, Carole Morel, Sebastian P Fernandez, Christophe Lanteri, Gérard Godeheu, Jean-Pol Tassin, Cédric Mombereau, Philippe Faure, and François Tronche. Chronic stress triggers social aversion via glucocorticoid receptor in dopaminergic neurons. *Science (New York, N.Y.)*, 339(6117):332–5, jan 2013.
- [95] M Marinelli, C N Rudick, X-T Hu, and F J White. Excitability of dopamine neurons: modulation and physiological consequences. *CNS & neurological disorders drug targets*, 5(1):79–97, February 2006.
- [96] M Marinelli and F J White. Enhanced vulnerability to cocaine self-administration is associated with elevated impulse activity of midbrain dopamine neurons. *Journal of Neuroscience*, 20(23):8876–8885, December 2000.



- [97] F G Gonon. Non-linear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by in vivo electrochemistry. *Neuroscience*, 24(1):19–28, 1988.
- [98] Hsing-Chen Tsai, Feng Zhang, Antoine Adamantidis, Garret D Stuber, Antonello Bonci, Luis de Lecea, and Karl Deisseroth. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science (New York, N.Y.)*, 324(5930):1080–4, 2009.
- [99] Stefania Tolu, Fabio Marti, Carole Morel, Carole Perrier, Nicolas Torquet, Stephanie Pons, Renaud de Beaurepaire, and Philippe Faure. Nicotine enhances alcohol intake and dopaminergic responses through  $\beta_2^*$  and  $\beta_4^*$  nicotinic acetylcholine receptors. *Scientific Reports*, 7(March):45116, 2017.
- [100] Wei-Xing Pan and Brian I Hyland. Pedunculopontine tegmental nucleus controls conditioned responses of midbrain dopamine neurons in behaving rats. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(19):4725–32, may 2005.
- [101] D J Lodge and A A Grace. The laterodorsal tegmentum is essential for burst firing of ventral tegmental area dopamine neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 103(13):5167–5172, March 2006.
- [102] J Mena-Segovia, P Winn, and J P Bolam. Cholinergic modulation of midbrain dopaminergic systems. *Brain research reviews*, 58(2):265–71, aug 2008.
- [103] K Chergui, P J Charley, H Akaoka, C F Saunier, J Brunet, M Buda, T H Svensson, G Chouvet, U Inserm, Cnrs Ura, Centre Hospitalier, Lyon Sud, Grand Revoyet, and Pierre Benite. Tonic Activation of NMDA Receptors Causes Spontaneous Burst Discharge of Rat Midbrain Dopamine Neurons In Vivo. *European Journal of Neuroscience*, 5:137–144, 1993.
- [104] Larry S Zweifel, Jones G Parker, Collin J Lobb, Aundrea Rainwater, Valerie Z Wall, Jonathan P Fadok, Martin Darvas, Min J Kim, Sheri J Y Mizumori, Carlos a Paladini, Paul E M Phillips, and Richard D Palmiter. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18):7281–8, may 2009.
- [105] Anna Y Kuznetsova, Marco A Huertas, Alexey S Kuznetsov, Carlos A Paladini, and Carmen C Canavier. Regulation of firing frequency in a computational model of a midbrain dopaminergic neuron. *Journal of computational neuroscience*, 28(3):389–403, June 2010.
- [106] C A Paladini. GABA(A) and GABA(B) antagonists differentially affect the firing pattern of substantia nigra dopaminergic neurons in vivo. *Synapse*, 32(3):165–176, June 1999.
- [107] James M Tepper and Christian R Lee. GABAergic control of substantia nigra dopaminergic neurons. *Progress in brain research*, 160:189–208, 2007.
- [108] C J Lobb, C J Wilson, and C A Paladini. A Dynamic Role for GABA Receptors on the Firing Pattern of Midbrain Dopaminergic Neurons. *Journal of neurophysiology*, 104(1):403–413, July 2010.
- [109] Olivier Waroux, Laurent Massotte, Livia Alleva, Amaury Graulich, Elizabeth Thomas, Jean François Liégeois, Jacqueline Scuvée-Moreau, and Vincent Seutin. SK channels con-

- trol the firing pattern of midbrain dopaminergic neurons in vivo. *European Journal of Neuroscience*, 22(12):3111–3121, 2005.
- [110] C.A. Sanford A.S. Chung A.D. Guler C. Chavkin R. Lujan & L.S. Zweifel M.E. Soden, G.L. Jones. Disruption of Dopamine Neuron Activity Pattern Regulation through Selective Expression of a Human KCNN3 Mutation. *Neuron*, 19(6):431–444, 2014.
- [111] J Cornwall, J D Cooper, and O T Phillipson. *Experimental Brain Research* 9. pages 157–171, 1990.
- [112] Philip Winn. How best to consider the structure and function of the pedunculopontine tegmental nucleus: Evidence from animal studies. *Journal of the Neurological Sciences*, 248(1-2):234–250, 2006.
- [113] S A Oakman, P L Faris, P E Kerr, C Cozzari, and B K Hartman. Distribution of pontomesencephalic cholinergic neurons projecting to substantia nigra differs significantly from those projecting to ventral tegmental area. *The Journal of neuroscience*, 15(9):5859–5869, 1995.
- [114] D J Lodge and A A Grace. The laterodorsal tegmentum is essential for burst firing of ventral tegmental area dopamine neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 103(13):5167–72, 2006.
- [115] S. J A Lokwan, P. G. Overton, M. S. Berry, and D. Clark. Stimulation of the pedunculopontine tegmental nucleus in the rat produces burst firing in A9 dopaminergic neurons. *Neuroscience*, 92(1):245–254, 1999.
- [116] R.D. Oades & G.M. Halliday. Ventral tegmental (a10) system: neurobiology. 1. anatomy and connectivity. *Brain Research Reviews*, 12:117–165, 1987.
- [117] Jochen Roeper. Dissecting the diversity of midbrain dopamine neurons. *Trends in Neurosciences*, 36(6):336–342, 2013.
- [118] Stephan Lammel, Byung Kook Lim, and Robert C. Malenka. Reward and aversion in a heterogeneous midbrain dopamine system. *Neuropharmacology*, 76(PART B):351–359, 2014.
- [119] Kevin T Beier, Elizabeth E Steinberg, Katherine E Deloach, Stanley Xie, Kazunari Miyamichi, Lindsay Schwarz, Xiaojing J Gao, Eric Kremer, Robert C Malenka, and L Luo. Circuit Architecture of Midbrain Dopamine Neurons Revealed by Systematic Input and Output Mapping. *Cell*, pages 622–634, 2015.
- [120] Thomas S. Hnasko, Nao Chuhma, Hui Zhang, Germaine Y. Goh, David Sulzer, Richard D. Palmiter, Stephen Rayport, and Robert H. Edwards. Vesicular Glutamate Transport Promotes Dopamine Storage and Glutamate Corelease In Vivo. *Neuron*, 65(5):643–656, 2010.
- [121] David H Root, Carlos A Mejias-Aponte, Jia Qi, and Marisela Morales. Role of glutamatergic projections from ventral tegmental area to lateral habenula in aversive conditioning. *Journal of Neuroscience*, 34(42):13906–13910, October 2014.
- [122] Stephan Lammel, Daniela I Ion, Jochen Roeper, and Robert C Malenka. Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron*, 70(5):855–62, jun 2011.

- [123] H. Ji and P. D. Shepard. Lateral Habenula Stimulation Inhibits Rat Midbrain Dopamine Neurons through a GABAA Receptor-Mediated Mechanism. *Journal of Neuroscience*, 27(26):6923–6930, 2007.
- [124] Stephan Lammel, Byung Kook Lim, Chen Ran, Kee Wui Huang, and Michael J Betley. Input-specific control of reward and aversion in the ventral tegmental area. *Nature*, 491(7423):212–217, 2013.
- [125] Colin M. Stopper, Maric T L Tse, David R. Montes, Candice R. Wiedman, and Stan B. Floresco. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron*, 84(1):177–189, 2014.
- [126] Masayuki Matsumoto and Okihide Hikosaka. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, 447(7148):1111–1115, 2007.
- [127] Masayuki Matsumoto and Okihide Hikosaka. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, 459(7248):837–41, jun 2009.
- [128] Arpiar Saunders, Ian a Oldenburg, Vladimir K Berezovskii, Caroline a Johnson, Nathan D Kingery, Hunter L Elliott, Tiao Xie, Charles R Gerfen, and Bernardo L Sabatini. A direct GABAergic output from the basal ganglia to frontal cortex. *Nature*, 05, 2015.
- [129] Shiliang Zhang, Jia Qi, Xueping Li, Hui-Ling Wang, Jonathan P Britt, Alexander F Hoffman, Antonello Bonci, Carl R Lupica, and Marisela Morales. Dopaminergic and glutamatergic microdomains in a subset of rodent mesoaccumbens axons. *Nature neuroscience*, 18(3):386–92, 2015.
- [130] Lauren Faget, Fumitaka Osakada, Jinyi Duan, Reed Ressler, Alexander B. Johnson, James A. Proudfoot, Ji Hoon Yoo, Edward M. Callaway, and Thomas S. Hnasko. Afferent Inputs to Neurotransmitter-Defined Cell Types in the Ventral Tegmental Area. *Cell Reports*, 15(12):2796–2808, 2016.
- [131] Frédéric Brischoux, Subhojit Chakraborty, Daniel I Brierley, and Mark A Ungless. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 106(12):4894–4899, March 2009.
- [132] Christopher Ford. The Role of D2-Autoreceptors in Regulating Dopamine Neuron Activity and Transmission. *Neuroscience*, 510(7503):84–91, 2015.
- [133] Johannes W de Jong, Theresia J M Roelofs, Frédérique M U Mol, Anne E J Hillen, Katharina E Meijboom, Mienke C M Luijendijk, Harrie a M van der Eerden, Keith M Garner, Louk J M J Vanderschuren, and Roger a H Adan. Reducing Ventral Tegmental Dopamine D2 Receptor Expression Selectively Boosts Incentive Motivation. *Neuropsychopharmacology*, 40(9):2085–95, 2015.
- [134] A. J. Lança, K. L. Adamson, K. M. Coen, B. L.C. Chow, and W. A. Corrigall. The pedunculopontine tegmental nucleus and the role of cholinergic neurons in nicotine self-administration in the rat: A correlative neuroanatomical and behavioral study. *Neuroscience*, 96(4):735–742, 2000.
- [135] Antoine R Adamantidis, Hsing-Chen Tsai, Benjamin Boutrel, Feng Zhang, Garret D Stuber, Evgeny a Budygin, Clara Touriño, Antonello Bonci, Karl Deisseroth, and Luis

- de Lecea. Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior., jul 2011.
- [136] Kay M Tye, Julie J Mirzabekov, Melissa R Warden, Emily A Ferenczi, Hsing-chen Tsai, Joel Finkelstein, Sung-yon Kim, Avishek Adhikari, Kimberly R Thompson, Aaron S Andalman, Lisa A Gunaydin, and Ilana B Witten. expression of depression-related behaviour. *Nature*, 493(7433):537–541, 2013.
- [137] Garret D. Stuber<sup>1</sup>, Jonathan P. Britt and Antonello Bonci. Optogenetic modulation of neural circuits that underlie reward seeking. 71(12):1061–1067, 2013.
- [138] Ilana B. Witten, Elizabeth E. Steinberg, Soo Yeun Lee, Thomas J. Davidson, Kelly A. Zalocusky, Matthew Brodsky, Ofer Yizhar, Saemi L. Cho, Shiaoching Gong, Charu Ramakrishnan, Garret D. Stuber, Kay M. Tye, Patricia H. Janak, and Karl Deisseroth. Recombinase-driver rat lines: Tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*, 72(5):721–733, 2011.
- [139] William R. Stauffer, Armin Lak, Aimei Yang, Melodie Borel, Ole Paulsen, Edward S. Boyden, and Wolfram Schultz. Dopamine Neuron-Specific Optogenetic Stimulation in Rhesus Macaques. *Cell*, 166(6):1564–1571.e6, 2016.
- [140] Kelly R. Tan, Cédric Yvon, Marc Turiault, Julie J. Mirzabekov, Jana Doehner, Gwenaél Labouèbe, Karl Deisseroth, Kay M. Tye, and Christian Lüscher. GABA Neurons of the VTA Drive Conditioned Place Aversion. *Neuron*, 73(6):1173–1183, 2012.
- [141] Ruud Van Zessen, Jana L. Phillips, Evgeny A. Budygin, and Garret D. Stuber. Activation of VTA GABA Neurons Disrupts Reward Consumption. *Neuron*, 73(6):1184–1194, 2012.
- [142] Dipesh Chaudhury, Jessica J Walsh, Allyson K Friedman, Barbara Juarez, M Ku, Ja Wook Koo, Deveroux Ferguson, Hsing-chen Tsai, Lisa Pomeranz, J Daniel, Alexander R Nectow, Mats Ekstrand, Ana Domingos, Michelle Mazie, Ezekiel Mouzon, Mary Kay Lobo, Rachael L Neve, Jeffrey M Friedman, Scott J Russo, Karl Deisseroth, Eric J Nestler, and Ming-hu Han. Rapid regulation of depression-related behaviors by control of midbrain dopamine neurons. *Nature*, (7433):532–536.
- [143] Jun-Li Cao, Herbert E Covington, Allyson K Friedman, Matthew B Wilkinson, Jessica J Walsh, Donald C Cooper, Eric J Nestler, and Ming-Hu Han. Mesolimbic dopamine neurons in the brain reward circuit mediate susceptibility to social defeat and antidepressant action. *Journal of neuroscience*, 30(49):16453–8, dec 2010.
- [144] Kohei Tanaka, Tomoyuki Furuyashiki, Shiho Kitaoka, Yuta Senzai, Yuki Imoto, Eri Segi-Nishida, Yuichi Deguchi, Richard M Breyer, Matthew D Breyer, and Shuh Narumiya. Prostaglandin E<sub>2</sub>-mediated attenuation of mesocortical dopaminergic pathway is critical for susceptibility to repeated social defeat stress in mice. *Journal of neuroscience*, 32(12):4319–29, mar 2012.
- [145] Kevin Lloyd and Peter Dayan. Safety out of control: dopamine and defence. *Behavioral and brain functions*, 12(1):15, 2016.
- [146] Caroline E. Bass, Valentina P. Grinevich, Dominic Gioia, Jonathan D. Day-Brown, Keith D. Bonin, Garret D. Stuber, Jeff L. Weiner, and Evgeny A. Budygin. Optogenetic stimulation of VTA dopamine neurons reveals that tonic but not phasic patterns of

- dopamine transmission reduce ethanol self-administration. *Frontiers in Behavioral Neuroscience*, 7(November):1–10, 2013.
- [147] Maria A. Mikhailova, Caroline E. Bass, Valentina P. Grinevich, Ann M. Chappell, Alex L. Deal, Keith D. Bonin, Jeff L. Weiner, Raul R. Gainetdinov, and Evgeny A. Budygin. Optogenetically-induced tonic dopamine release from VTA-nucleus accumbens projections inhibits reward consummatory behaviors. *Neuroscience*, 333:54–64, 2016.
- [148] Annie M Whitaker, Muhammad A Farooq, Scott Edwards, and Nicholas W Gilpin. HHS Public Access. 19(1):69–77, 2016.
- [149] Yu Shin Ding, S. John Gatley, Joanna S. Fowler, Nora D. Volkow, Deepak Aggarwal, Jean Logan, Stephen L. Dewey, Feng Liang, F. Ivy Carroll, and Michael J. Kuhar. Mapping nicotinic acetylcholine receptors with PET. *Synapse*, 24(4):403–407, 1996.
- [150] V Itier and D Bertrand. Neuronal nicotinic receptors: from protein structure to function. *FEBS letters*, 504(3):118–125, 2001.
- [151] Federico Dajas-Bailador and Susan Wonnacott. Nicotinic acetylcholine receptors and the regulation of neuronal signalling. *Trends in Pharmacological Sciences*, 25(6):317–324, 2004.
- [152] Antoine Taly, Pierre-Jean Corringer, Denis Guedin, Pierre Lestage, and Jean-Pierre Changeux. Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nature Reviews Drug Discovery*, 8(9):733–750, 2009.
- [153] Jean-Pierre Changeux and Stuart J Edelstein. Allosteric mechanisms of signal transduction. *Science (New York, N.Y.)*, 308(5727):1424–8, jun 2005.
- [154] Jianlin Wu, John J. Dougherty, and Robert A. Nichols. Dopamine receptor regulation of Ca<sup>2+</sup> levels in individual isolated nerve terminals from rat striatum: Comparison of presynaptic D 1-like and D2-like receptors. *Journal of Neurochemistry*, 98(2):481–494, 2006.
- [155] M R Picciotto, M Zoli, C Léna, A Bessis, Y Lallemand, N Le Novère, P Vincent, E M Pich, P Brûlet, and J P Changeux. Abnormal avoidance learning in mice lacking functional high-affinity nicotine receptor in the brain. *Nature*, 374(6517):65–67, 1995.
- [156] B. Schilström, H. M. Svensson, T. H. Svensson, and G. G. Nomikos. Nicotine and food induced dopamine release in the nucleus accumbens of the rat putative role of  $\alpha 7$  nicotinic receptors in the ventral tegmental area. *Neuroscience*, 85(4):1005–1009, 1998.
- [157] L M Marubio and J Changeux. Nicotinic acetylcholine receptor knockout mice as animal models for studying receptor function. *European journal of pharmacology*, 393(1-3):113–121, 2000.
- [158] Jean-Pierre Changeux. Nicotine addiction and nicotinic receptors: lessons from genetically modified mice. *Nature reviews. Neuroscience*, 11(6):389–401, jun 2010.
- [159] A. R. Tapper. Nicotine Activation of 4\* Receptors: Sufficient for Reward, Tolerance, and Sensitization. *Science*, 306(5698):1029–1032, 2004.
- [160] Ryan M. Drenan, Sharon R. Grady, Paul Whiteaker, Tristan McClure-Begley, Sheri McKinney, Julie M. Miwa, Sujata Bupp, Nathaniel Heintz, J. Michael McIntosh, Merouane

- Bencherif, Michael J. Marks, and Henry A. Lester. In Vivo Activation of Midbrain Dopamine Neurons via Sensitized, High-Affinity  $\alpha 6^*$  Nicotinic Acetylcholine Receptors. *Neuron*, 60(1):123–136, 2008.
- [161] Ryan M Drenan and Henry A. Lester. Insights into the neurobiology of the nicotinic cholinergic system and nicotine addiction from mice expressing nicotinic receptors harboring gain-of-function mutations. *Pharmacological Reviews*, 64(4):869–79, 2012.
- [162] J-P Changeux, D Bertrand, P J Corringer, S Dehaene, S Edelstein, C Léna, N Le Novère, L Marubio, M Picciotto, and M Zoli. Brain nicotinic receptors: structure and regulation, role in learning and reinforcement. *Brain research reviews*, 26(2-3):198–216, May 1998.
- [163] John a Dani and Daniel Bertrand. Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annual review of pharmacology and toxicology*, 47:699–729, jan 2007.
- [164] John A Thompson and Gidon Felsen. Activity in mouse pedunculopontine tegmental nucleus reflects action and outcome in a decision-making task. *Journal of neurophysiology*, 110(12):2817–2829, December 2013.
- [165] Ruby Klink, Alban de Kerchove d’Exaerde, Michele Zoli, and Jean-Pierre Changeux. Molecular and physiological diversity of nicotinic acetylcholine receptors in the midbrain dopaminergic nuclei. *The Journal of neuroscience*, 21(5):1452–1463, 2001.
- [166] Nicolas Champtiaux, Cecilia Gotti, Matilde Cordero-Erausquin, Denis J David, Cédric Przybylski, Clément Léna, Francesco Clementi, Milena Moretti, Francesco M Rossi, Nicolas Le Novère, J Michael McIntosh, Alain M Gardier, and Jean-Pierre Changeux. Subunit composition of functional nicotinic receptors in dopaminergic neurons investigated with knock-out mice. *Journal of Neuroscience*, 23(21):7820–7829, August 2003.
- [167] Cecilia Gotti, Stefania Guiducci, Vincenzo Tedesco, Silvia Corbioli, Lara Zanetti, Milena Moretti, Alessio Zanardi, Roberto Rimondini, Manolo Mugnaini, Francesco Clementi, Christian Chiamulera, and Michele Zoli. Nicotinic acetylcholine receptors in the mesolimbic pathway: primary role of ventral tegmental area  $\alpha 6\beta 2^*$  receptors in mediating systemic nicotine effects on dopamine release, locomotion, and reinforcement. *Journal of neuroscience*, 30(15):5311–5325, 2010.
- [168] Susmita Chatterjee, Nathan Santos, Joan Holgate, Carolina L Haass-koffler, F Woodward Hopf, Viktor Kharazia, Henry Lester, Antonello Bonci, and Selena E Bartlett. The  $\alpha 5$  Subunit Regulates the Expression and Function of  $\alpha 4^*$ -Containing Neuronal Nicotinic Acetylcholine Receptors in the Ventral-Tegmental Area. *PlosOne*, 8(7):1–10, 2013.
- [169] Monica Mameli-Engvall, Alexis Evrard, Stéphanie Pons, Uwe Maskos, Torgny H Svensson, Jean-Pierre Changeux, and Philippe Faure. Hierarchical control of dopamine neuron-firing patterns by nicotinic receptors. *Neuron*, 50(6):911–21, jun 2006.
- [170] D James Surmeier, Jun Ding, Michelle Day, Zhongfeng Wang, and Weixing Shen. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in neurosciences*, 30(5):228–35, may 2007.
- [171] Huibert D Mansvelder, J Russel Keath, and Daniel S McGehee. Synaptic mechanisms underlie nicotine-induced excitability of brain reward areas. *Neuron*, 33(6):905–919, March 2002.

- [172] V I Pidoplichko, M DeBiasi, J T Williams, and J a Dani. Nicotine activates and desensitizes midbrain dopamine neurons. *Nature*, 390(6658):401–4, nov 1997.
- [173] Stéphanie Caillé, Karine Guillem, Martine Cador, Olivier Manzoni, and François Georges. Voluntary nicotine consumption triggers in vivo potentiation of cortical excitatory drives to midbrain dopaminergic neurons. *Journal of neuroscience*, 29(33):10410–5, 2009.
- [174] Wei Li, William M. Doyon, and John A. Dani. Acute in vivo nicotine administration enhances synchrony among dopamine neurons. *Biochemical Pharmacology*, 82(8):977–983, 2011.
- [175] Mariella De Biasi and John A Dani. Reward, addiction, withdrawal to nicotine. *Annual review of neuroscience*, 34:105–30, 2011.
- [176] David C Perry, Martha I. Davila-Garcia, Craig A Stockmeier, and Kenneth J Kellar. Increased Nicotinic Receptors in Brains from Smokers: Membrane Binding and Autoradiography Studies. *J. Pharmacol. Exp. Ther.*, 289(3):1545–1552, 1999.
- [177] L M Marubio, A M Gardier, S Durier, D David, R Klink, M M Arroyo-Jimenez, J M McIntosh, F Rossi, N Champtiaux, M Zoli, and J P Changeux. Effects of nicotine in the dopaminergic system of mice lacking the alpha4 subunit of neuronal nicotinic acetylcholine receptors. *The European journal of neuroscience*, 17(7):1329–1337, April 2003.
- [178] S Pons, Liana Fattore, G Cossu, S Tolu, E Porcu, J M McIntosh, J P Changeux, U Maskos, and W Fratta. Crucial role of alpha4 and alpha6 nicotinic acetylcholine receptor subunits from ventral tegmental area in systemic nicotine self-administration. *Journal of Neuroscience*, 28(47):12318–12327, November 2008.
- [179] Stephanie J. Cragg. Meaningful silences: How dopamine listens to the ACh pause. *Trends in Neurosciences*, 29(3):125–131, 2006.
- [180] Richard Exley, Nicolas Maubourguet, Vincent David, Raphaël Eddine, Alexis Evrard, Stéphanie Pons, Fabio Marti, Sarah Threlfell, Pierre Cazala, J Michael McIntosh, Jean-Pierre Changeux, Uwe Maskos, Stephanie J Cragg, and Philippe Faure. Distinct contributions of nicotinic acetylcholine receptor subunit  $\alpha 4$  and subunit  $\alpha 6$  to the reinforcing effects of nicotine. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7577–7582, 2011.
- [181] Darlene H Brunzell, Karen E Boschen, Elizabeth S Hendrick, Patrick M Beardsley, and J Michael McIntosh. Alpha-conotoxin MII-sensitive nicotinic acetylcholine receptors in the nucleus accumbens shell regulate progressive ratio responding maintained by nicotine. *Neuropsychopharmacology*, 35(3):665–673, 2010.
- [182] Changhai Cui, T K Booker, Roberta S Allen, Sharon R Grady, Paul Whiteaker, Michael J Marks, Outi Salminen, Theresa Tritto, Christopher M Butt, W R Allen, Jerry a Stitzel, J Michael McIntosh, Jim Boulter, Allan C Collins, and Stephen F Heinemann. The beta3 nicotinic receptor subunit: a component of alpha-conotoxin MII-binding nicotinic acetylcholine receptors that modulate dopamine release and related behaviors. *Journal of neuroscience*, 23(35):11045–11053, 2003.
- [183] Margaret E Rice and Stephanie J Cragg. Nicotine amplifies reward-related dopamine signals in striatum. *Nature Neuroscience*, 7(6):583–584, 2004.

- [184] Hui Zhang and David Sulzer. Frequency-dependent modulation of dopamine release by nicotine. *Nature neuroscience*, 7(6):581–2, jun 2004.
- [185] Sarah Threlfell, Tatjana Lalic, Nicola J. Platt, Katie A. Jennings, Karl Deisseroth, and Stephanie J. Cragg. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron*, 75(1):58–64, 2012.
- [186] U Maskos. The cholinergic mesopontine tegmentum is a relatively neglected nicotinic master modulator of the dopaminergic system: relevance to drugs of abuse and pathology. *British journal of pharmacology*, 153 Suppl:S438–45, mar 2008.
- [187] S Wonnacott. Gates and filters: unveiling the physiological roles of nicotinic acetylcholine receptors in dopaminergic transmission. *British journal of pharmacology*, 153 Suppl:S2–4, 2008.
- [188] C Morel, Liana Fattore, S Pons, Y A Hay, F Marti, B Lambolez, M De Biasi, M Lathrop, W Fratta, U Maskos, and Philippe Faure. Nicotine consumption is regulated by a human polymorphism in dopamine neurons. *Molecular Psychiatry*, 19(8):930–936, August 2014.
- [189] Fani Koukouli, Marie Rooy, Dimitrios Tziotis, Kurt A Sailor, Heidi C O’Neill, Josien Levenga, Mirko Witte, Michael Nilges, Jean-Pierre Changeux, Charles A Hoeffler, Jerry A Stitzel, Boris S Gutkin, David A DiGregorio, and Uwe Maskos. Nicotine reverses hypofrontality in animal models of addiction and schizophrenia. *Nature Medicine*, 23(3):347–354, 2017.
- [190] Shannon G. Matta, David J. Balfour, Neal L. Benowitz, R. Thomas Boyd, Jerry J. Buccafusco, Anthony R. Caggiula, Caroline R. Craig, Allan C. Collins, M. Imad Damaj, Eric C. Donny, Phillip S. Gardiner, Sharon R. Grady, Ulrike Heberlein, Sherry S. Leonard, Edward D. Levin, Ronald J. Lukas, Athina Markou, Michael J. Marks, Sarah E. McCallum, Neeraja Parameswaran, Kenneth A. Perkins, Marina R. Picciotto, Maryka Quik, Jed E. Rose, Adrian Rothenfluh, William R. Schafer, Ian P. Stolerman, Rachel F. Tyndale, Jeanne M. Wehner, and Jeffrey M. Zirger. Guidelines on nicotine dose selection for in vivo research. *Psychopharmacology*, 190(3):269–319, 2007.
- [191] Cynthia L Gentry, Lincoln H Wilkins, and Ronald J Lukas. Effects of Prolonged Nicotinic Ligand Exposure on Function of Acetylcholine Receptors. *Pharmacology*, 304(1):206–216, 2003.
- [192] J R Pauly, J W Stitzel, M J Marks, and A C Collins. An autoradiographic analysis of cholinergic receptors in mouse brain. *Brain research bulletin*, 22(2):453–459, 1989.
- [193] Mohammed Shoaib, Charles W. Schindler, and Steven R. Goldberg. Nicotine self-administration in rats: Strain and nicotine pre-exposure effects on acquisition. *Psychopharmacology*, 129(1):35–43, 1997.
- [194] S L Yates, M Bencherif, E N Fluhler, and P M Lippiello. Up-regulation of nicotinic acetylcholine receptors following chronic exposure of rats to mainstream cigarette smoke or alpha 4 beta 2 receptors to nicotine. *Biochemical pharmacology*, 50(12):2001–8, 1995.
- [195] X. Zhang, J. Y. Tian, A. L. Svensson, Z. H. Gong, B. Meyerson, and A. Nordberg. Chronic treatments with tacrine and (-)-nicotine induce different changes of nicotinic and muscarinic acetylcholine receptors in the brain of aged rat. *Journal of Neural Transmission*, 109(3):377–392, 2002.



- [196] B Buisson and Daniel Bertrand. Chronic exposure to nicotine upregulates the human  $(\alpha)4((\beta)2$  nicotinic acetylcholine receptor function. *The Journal of neuroscience*, 21(6):1819–1829, March 2001.
- [197] Eduardo J. Molinari, Osvaldo Delbono, Maria L. Messi, Muthukrishnan Renganathan, Stephen P. Arneric, James P. Sullivan, and Murali Gopalakrishnan. Up-regulation of human  $\alpha7$  nicotinic receptors by chronic treatment with activator and antagonist ligands. *European Journal of Pharmacology*, 347(1):131–139, 1998.
- [198] Steven L Parker, Yitong Fu, Kathleen McAllen, Jianhong Luo, J Michael McIntosh, Jon M Lindstrom, and Burt M Sharp. Up-regulation of brain nicotinic acetylcholine receptors in the rat during long-term self-administration of nicotine: disproportionate increase of the  $\alpha6$  subunit. *Molecular pharmacology*, 65(3):611–22, 2004.
- [199] Albert Lai, Neeraja Parameswaran, Mirium Khwaja, Paul Whiteaker, Jon M Lindstrom, Hong Fan, J Michael McIntosh, Sharon R Grady, and Maryka Quik. Long-Term Nicotine Treatment Decreases Striatal  $\alpha6$  \* Nicotinic Acetylcholine Receptor Sites and Function in Mice. *Molecular Pharmacology*, 67(5):1639–1647, 2005.
- [200] M. Mugnaini, M. Garzotti, I. Sartori, M. Pilla, P. Repeto, C. A. Heidbreder, and M. Tessari. Selective down-regulation of  $[125I]Y0$ - $\alpha$ -conotoxin MII binding in rat mesostriatal dopamine pathway following continuous infusion of nicotine. *Neuroscience*, 137(2):565–572, 2006.
- [201] Maureen E M Benwell, D. J K Balfour, and J. M. Anderson. Evidence that Tobacco Smoking Increases the Density of  $[3H]$ Nicotine Binding Sites in Human Brain. *Journal of Neurochemistry*, 50(4):1243–1247, 1988.
- [202] Breese, M J Marks, J Logel, C E Adams, B Sullivan, a C Collins, and S Leonard. Effect of smoking history on  $[3H]$ nicotine binding in human postmortem brain. *The Journal of pharmacology and experimental therapeutics*, 282(1):7–13, 1997.
- [203] A L Brody, R E Olmstead, E D London, J Farahi, J H Meyer, P Grossman, G S Lee, J Huang, E L Hahn, and M A Mandelkern. Smoking-induced ventral striatum dopamine release. *Am.J Psychiatry*, 161:1211–1218, 2004.
- [204] Anitha P. Govind, Paul Vezina, and William N. Green. Nicotine-induced upregulation of nicotinic receptors: Underlying mechanisms and relevance to nicotine addiction. *Biochemical Pharmacology*, 78(7):756–765, 2009.
- [205] C P Fenster, T L Whitworth, E B Sheffield, M W Quick, and R A J Lester. Upregulation of Surface  $\alpha4\beta2$  Nicotinic Receptors Is Initiated by Receptor Desensitization after Chronic Exposure to Nicotine. *Journal of neuroscience*, 19(12):4804–14, 1999.
- [206] Raad Nashmi, Mary E Dickinson, Sheri McKinney, Mark Jareb, Cesar Labarca, Scott E Fraser, and Henry a Lester. Assembly of  $\alpha4\beta2$  nicotinic acetylcholine receptors assessed with functional fluorescently labeled subunits: effects of localization, trafficking, and nicotine-induced upregulation in clonal mammalian cells and in cultured midbrain neurons. *Journal of neuroscience*, 23(37):11554–11567, 2003.
- [207] Patricia C Harkness and Neil S Millar. Changes in conformation and subcellular distribution of  $\alpha4\beta2$  nicotinic acetylcholine receptors revealed by chronic nicotine treatment and expression of subunit chimeras. *Journal of neuroscience*, 22(23):10172–81, 2002.

- [208] Fan Wang, Mark E Nelson, Alexander Kuryatov, Felix Olale, John Cooper, Kent Keyser, and Jon Lindstrom. Chronic Nicotine Treatment Up-regulates Human  $\alpha 3\beta 2$  but Not  $\alpha 3\beta 4$  Acetylcholine Receptors Stably Transfected in Human Embryonic Kidney Cells. *The journal of biological chemistry*, 273(44):28721–28732, 1998.
- [209] Pierre Jean Corringer, Jérôme Sallette, and Jean Pierre Changeux. Nicotine enhances intracellular nicotinic receptor maturation: A novel mechanism of neural plasticity? *Journal of Physiology Paris*, 99(2-3):162–171, 2006.
- [210] J Ramirez-Latorre, C R Yu, X Qu, F Perin, A Karlin, and L Role. Functional contributions of  $\alpha 5$  subunit to neuronal acetylcholine receptor channels. *Nature*, 380(6572):347–351, 1996.
- [211] H D Mansvelder and D S McGehee. Long-term potentiation of excitatory inputs to brain reward areas by nicotine. *Neuron*, 27(2):349–57, aug 2000.
- [212] Satoshi Fujii, Zhanxin Ji, Naonori Morita, and Katumi Sumikawa. Acute and chronic nicotine exposure differentially facilitate the induction of LTP. *Brain Research*, 846(1):137–143, 1999.
- [213] Daniel Saal, Yan Dong, Antonello Bonci, and Robert C Malenka. Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neuron*, 37(4):577–82, feb 2003.
- [214] Luis M. Tuesta, Christie D. Fowler, and Paul J. Kenny. Recent advances in understanding nicotinic receptor signaling mechanisms that regulate drug self-administration behavior. *Biochemical Pharmacology*, 82(8):984–995, 2011.
- [215] Agnes J Jasinska, Todd Zorick, Arthur L Brody, and Elliot A Stein. Neuropharmacology Dual role of nicotine in addiction and cognition : A review of neuroimaging studies in humans. *Neuropharmacology*, 84:111–122, 2014.
- [216] S Dehaene, M Kerszberg, and J P Changeux. A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24):14529–14534, 1998.
- [217] Craig D C Bailey, Mariella De Biasi, Paul J Fletcher, Evelyn K Lambe, Mariella De Biasi, Paul J Fletcher, and Evelyn K Lambe. The Nicotinic Acetylcholine Receptor  $\alpha 5$  Subunit Plays a Key Role in Attention Circuitry and Accuracy. *The Journal of neuroscience*, 30(27):9241–9252, 2010.
- [218] Jie Gao, Bao Ling Adam, and Alvin V. Terry. Evaluation of nicotine and cotinine analogs as potential neuroprotective agents for Alzheimer’s disease. *Bioorganic and Medicinal Chemistry Letters*, 24(6):1472–1478, 2014.
- [219] Morgane Besson, Stefania Guiducci, Sylvie Granon, Jean-Philippe Guilloux, Bruno Guiard, Christelle Reperant, Philippe Faure, Stéphanie Pons, Giuseppe Cannazza, Michele Zoli, Alain M Gardier, and Uwe Maskos. Alterations in  $\alpha 5^*$  nicotinic acetylcholine receptors result in midbrain- and hippocampus-dependent behavioural and neural impairments. *Psychopharmacology (Berl)*, 233(18):3297–3314, September 2016.
- [220] Michael D Greicius, Ben Krasnow, Allan L Reiss, and Vinod Menon. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 100(1):253–8, 2003.

- [221] Martijn P Van Den Heuvel, C W Mandl, S Kahn, and Hilleke E Hulshoff Pol. Functionally Linked Resting-State Networks Reflect the Underlying Structural Connectivity Architecture of the Human Brain. *Human Brain Mapping*, 3141:3127–3141, 2009.
- [222] Marina R Picciotto. Nicotine as a modulator of behavior: beyond the inverted U. *Trends in pharmacological sciences*, 24(9):493–9, sep 2003.
- [223] Stephen J Heishman, Bethea a Kleykamp, and Edward G Singleton. Meta-analysis of the acute effects of nicotine and smoking on human performance. *Psychopharmacology*, 210(4):453–69, jul 2010.
- [224] Katerina Zoe Kolokotroni, Robert J Rodgers, and Amanda a Harrison. Acute nicotine increases both impulsive choice and behavioural disinhibition in rats. *Psychopharmacology*, 217(4):455–73, oct 2011.
- [225] Jürgen Gallinat, Eva Meisenzahl, Leslie K. Jacobsen, Peter Kalus, Jeffrey Bierbrauer, Thorsten Kienast, Henning Witthaus, Karolina Leopold, Frank Seifert, Florian Schubert, and Mario Staedtgen. Smoking and structural brain deficits: A volumetric MR investigation. *European Journal of Neuroscience*, 24(6):1744–1750, 2006.
- [226] Manoranjan S D D’Souza and Athina Markou. The ‘Stop’ and ‘Go’ of Nicotine Dependence: Role of GABA and Glutamate. *Cold Spring Harb Perspect Med*, 3:1–18, 2013.
- [227] Edward D Levin, F Joseph McClernon, and Amir H Rezvani. Nicotinic effects on cognitive function: behavioral characterization, pharmacological specification, and anatomic localization. *Psychopharmacology (Berl)*, 184(3-4):523–539, March 2006.
- [228] Sandra Kalmijn, Martin P J Van Boxtel, Monique W M Verschuren, Jelle Jolles, and Lenore J Launer. Cigarette Smoking and Alcohol Consumption in Relation to Cognitive Performance in Middle Age. *American Journal of Epidemiology*, 156(10):936–944, 2002.
- [229] Angela J Yu and Peter Dayan. Uncertainty, neuromodulation, and attention. *Neuron*, 46(4):681–92, 2005.
- [230] Nathaniel D Daw and Kenji Doya. The computational neurobiology of learning and reward. *Current opinion in neurobiology*, 16(2):199–204, apr 2006.
- [231] Antonio Rangel, Colin Camerer, and P Read Montague. Neuroeconomics: The neurobiology of value-based decision-making. *Autism speaks*, 9(7):545–556, 2008.
- [232] Kenway Louie and Paul W. Glimcher. Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, 1251(1):13–32, 2012.
- [233] Wolfram Schultz. Dopamine reward prediction- error signalling: a two-component response. *Nature*, 17(3):183–195, 2016.
- [234] Mark D Humphries, Mehdi Khamassi, and Kevin Gurney. Dopaminergic Control of the Exploration-Exploitation Trade-Off via the Basal Ganglia. *Frontiers in neuroscience*, 6(February):9, jan 2012.
- [235] Richard S Sutton and Andrew G Barto. *Reinforcement Learning. An Introduction*. MIT Press, 1998.
- [236] D Kahneman and A Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 1979.

- [237] Paul J.H. Schoemaker. The Expected Utility Model: Its variants, purposes, evidence and limitations. *Journal of economic literature*, XX:529–563, 1982.
- [238] Merideth a Addicott, John M Pearson, Jessica Wilson, Michael L Platt, and F Joseph McClernon. Smoking and the bandit: a preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Experimental and clinical psychopharmacology*, 21(1):66–73, mar 2013.
- [239] Leonie de Visser, Judith R. Homberg, Manuela Mitsogiannis, Fiona D. Zeeb, Marion Rivalan, Aurélie Fitoussi, Vasco Galhardo, Ruud van den Bos, Catherine A. Winstanley, and Françoise Dells-Hagedorn. Rodent versions of the Iowa gambling task: Opportunities and challenges for the understanding of decision-making. *Frontiers in Neuroscience*, 5(OCT):1–21, 2011.
- [240] Dionisio a Amodeo, Joshua h Jones, John a Sweemey, and Michael E. Ragozzino. Differences in BTBR T+tf/J and C57BL/6J mice on probabilistic reversal learning and stereotyped behaviours. *Behavioural brain research*, 227(1):64–72, 2012.
- [241] Keith B. J. Franklin George Paxinos. The mouse brain in stereotaxic coordinates. 2004.
- [242] Hervé Le Mouellic, Yvan Lallemand, and Philippe Brûlet. Homeosis in the mouse induced by a null mutation in the Hox-3.1 gene. *Cell*, 69(2):251–264, 1992.
- [243] Jérémie Naudé, Stefania Tolu, Malou Dongelmans, Nicolas Torquet, Sébastien Valverde, Guillaume Rodriguez, Stéphanie Pons, Uwe Maskos, Alexandre Mourot, Fabio Marti, and Philippe Faure. Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking. *Nature Neuroscience*, 19(3), 2016.
- [244] PN Tobler, Christopher D Fiorillo, and W Schultz. Adaptive coding of reward value by dopamine neurons. *Science*, 203(March):1642–1645, 2005.
- [245] Armin Lak, William R Stauffer, and Wolfram Schultz. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6):2343–8, feb 2014.
- [246] C D Fiorillo, P N Tobler, and W Schultz. Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. *Science*, 299(March):1898–1902, 2003.
- [247] E. A. Witte, M. C. Davidson, and R. T. Marrocco. Effects of altering brain cholinergic activity on covert orienting of attention: Comparison of monkey and human performance. *Psychopharmacology*, 132(4):324–334, 1997.
- [248] Peter Killeen. The matching law. *Journal of the Experimental Analysis of Behavior*, 17(3):489–495, 1972.
- [249] Nathaniel D Daw, John P O’Doherty, Peter Dayan, Ben Seymour, and Raymond J Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–9, jun 2006.
- [250] Mehdi Keramati, Amir Dezfouli, and Payam Piray. Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS computational biology*, 7(5):e1002055, may 2011.
- [251] Jc Gittins, Jc Gittins, Dm Jones, and Dm Jones. A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika*, 66(3):561–565, 1979.

- [252] Joscha Bach and Stefan Edelkamp. *Advances in Artificial Intelligence*. 2011.
- [253] Y Niv, Nathaniel D Daw, and Peter Dayan. Choice values. *Nature Neuroscience*, January 2006.
- [254] W K Bickel, a L Odum, and G J Madden. Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, 146(4):447–54, oct 1999.
- [255] Forest Baker, Matthew W. Johnson, and Warren K. Bickel. Delay discounting in current and never-before cigarette smokers: Similarities and differences across commodity, sign, and magnitude. *Journal of Abnormal Psychology*, 112(3):382–392, 2003.
- [256] Adam Johnson, Matthijs AA van der Meer, and A David Redish. Integrating hippocampus and striatum in decision-making. *Current opinion in neurobiology*, 17(6):692–697, December 2007.
- [257] Matthew L Locey and Jesse Dallery. Isolating behavioral mechanisms of intertemporal choice: nicotine effects on delay discounting and amount sensitivity. *Journal of the experimental analysis of behavior*, 91(2):213–23, mar 2009.
- [258] Anne Jackson, Sarah Silk, Yazeed Buhidma, and Mohammed Shoaib. Varenicline, the clinically effective smoking cessation agent, restores probabilistic response reversal performance during withdrawal from nicotine. *Addiction Biology*, (2011), 2016.
- [259] El Lesage, S E Aronson, M T Sutherland, T J Ross, B Jo Salmeron, and El A Stein. Neural Signatures of Cognitive Flexibility and Reward Sensitivity Following Nicotinic Receptor Stimulation in Dependent Smokers A Randomized Trial. *JAMA Psychiatry*, (April):1–9, 2017.
- [260] Sarah R Heilbronner and Benjamin Y Hayden. Contextual factors explain risk-seeking preferences in rhesus monkeys. *Frontiers in neuroscience*, 7(February):7, jan 2013.
- [261] Clare M. Mathes Neil E. Rowland, Cheryl H. Vaughan and Anaya Mitra. Feeding behavior, obesity, and neuroeconomics. *Physiological Behavior*, 28(93):97–109, 2008.
- [262] William Timberlake and Blaine F Peden. Closed Economies. *Psychology*, 1(1):35–60, 1987.
- [263] Hamed Ekhtiari, Padideh Nasser, Fatemeh Yavari, Azarkhsh Mokri, and John Monterosso. *Neuroscience of drug craving for addiction medicine: From circuits to therapies*, volume 223. Elsevier B.V., 1 edition, 2016.
- [264] David Huston-Lyons and Conan Kornetsky. Effects of nicotine on the threshold for rewarding brain stimulation in rats. *Pharmacology, Biochemistry and Behavior*, 41(4):755–759, 1992.
- [265] Paul J Kenny and Athina Markou. Nicotine Self-Administration Acutely Activates Brain Reward Systems and Induces a Long-Lasting Increase in Reward Sensitivity. *Neuropsychopharmacology*, pages 1203–1211, 2005.
- [266] Ruth S. Barr, Diego A. Pizzagalli, Melissa A. Culhane, Donald C. Goff, and A. Eden Evins. A Single Dose of Nicotine Enhances Reward Responsiveness in Nonsmokers: Implications for Development of Dependence. *Biological Psychiatry*, 63(11):1061–1065, 2008.
- [267] P. Read Montague, Brooks King-Casas, and Jonathan D. Cohen. Imaging Valuation Models in Human Choice. *Annual Review of Neuroscience*, 29(1):417–448, 2006.

- [268] G. T. Pollard and J. L. Howard. The Geller-Seifter conflict paradigm with incremental shock. *Psychopharmacology*, 62(2):117–121, 1979.